UNIVERSITY OF ŽILINA

TRANSCOM 2013

10-th EUROPEAN CONFERENCE
OF YOUNG RESEARCHERS AND SCIENTISTS

under the auspices of

Dušan Čaplovič
Minister of Education, Science, Research and Sport of the Slovak Republic

&

Tatiana Čorejová
Rector of the University of Žilina

SECTION 3

INFORMATION AND COMMUNICATION TECHNOLOGIES

ŽILINA June 24 - 26, 2013
SLOVAK REPUBLIC

Edited by Michal Kochláň, Anton Lieskovský

ISBN: 978-80-554-0692-3

TRANSCOM 2013
10th European conference of young researchers and scientists

TRANSCOM 2013, the 10th international conference of young European researchers, scientists and educators, aims to establish and expand international contacts and co-operation. The 10th international conference TRANSCOM is jubilee. It will be held in the year when the University of Žilina celebrates 60 years since her constitution (1953 – 2013).The main purpose of the conference is to provide young researchers and scientists with an encouraging and stimulating environment in which they present results of their research to the scientific community. TRANSCOM has been organised regularly every other year since 1995. Between 160 and 400 young researchers and scientists participate regularly in the event. The conference is organised for postgraduate students and young researchers and scientists up to the age of 35 and their tutors. Young workers are expected to present the results they had achieved.

The conference is organised by the University of Žilina. It is the university with about 13 000 graduate and postgraduate students. The university offers Bachelor, Master and PhD programmes in the fields of transport, telecommunications, forensic engineering, management operations, information systems, in mechanical, civil, electrical, special engineering and in social sciences.

SECTIONS AND SCIENTIFIC COMMITTEE

1. TRANSPORT AND COMMUNICATIONS TECHNOLOGY.

Scientific committee: Bugaj M. (SK), Drozdziel P. (PL), Janáček J. (SK), Jánošíková Ľ. (SK), Kampf R. (CZ), Palúch S. (SK), Rievaj V. (SK)

2. ECONOMICS AND MANAGEMENT.

Scientific committee: Blašková M. (SK), Borkowski S. (PL), Graźulis V. (LT), Hittmár Š. (SK), Kucharčíková A. (SK), Kurtev I. (BG), Lyakin A. (RU), Rostášová M. (SK), Tomová A. (SK)

3. INFORMATION AND COMMUNICATION TECHNOLOGIES.

Scientific committee: Dado M. (SK), Klimo M. (SK), Kršák E. (SK), Matiaško K. (SK), Simunic D. (HR), Spalek J. (SK), Švadlenka L. (CZ), Vaculík J. (SK), Vaculík M. (SK), Vrček N. (HR)

4. ELECTRIC POWER SYSTEMS. ELECTRICAL AND ELECTRONIC ENGINEERING.

Scientific committee: Altus J. (SK), Brandštetter P. (CZ), Bury P. (SK), Čápová K. (SK), Dobrucký B. (SK), Janoušek L. (SK), Luft M. (PL), Szychta E. (PL), Špánik P. (SK), Vittek J. (SK)

5. MATERIAL ENGINEERING. MECHANICAL ENGINEERING TECHNOLOGIES.

Scientific committee: Adamczak S. (PL), Bolibruchová D. (SK), Cotetiu R. I. (RO), Dobrzanski L. A. (PL), Guagliano M. (I), Hadzima B. (SK), Konečná R. (SK), Kunz L. (CZ), Kuric I. (SK), Lukáč P. (CZ), Meško J. (SK), Neslušan M. (SK), Palček P. (SK), Takacs J. (H), Trojanová Z. (CZ), Ungureanu N. S. (RO)

6. MACHINES AND EQUIPMENT. APPLIED MECHANICS.

Scientific committee: Gerlici J. (SK), Chudzikiewicz A. (PL), Jandačka J. (SK), Medvecký Š. (SK), Mičieta B. (SK), Ochodek T. (CZ), Sága M. (SK), Žmindák M. (SK)

7. CIVIL ENGINEERING.

Scientific committee: Bujňák J. (SK), Grzeszczyk S. (PL), Ižvolt L. (SK), Malachova A. (RU), Melcer J. (SK), Nikolic R. (RS), Plášek O. (CZ), Segalini A. (I)

8. SOCIAL SCIENCES.

Scientific committee: Bálint V. (SK), Banáry B. (SK), Cabanová V. (SK), Dzhalladova I. (UKR), Jakubíková B. (SK), Khusainov D. (UKR), Kráľová Z. (SK), Marčoková M. (SK), Růžičková M. (SK), Schmeidel E. (PL)

9. SECURITY ENGINEERING. FORENSIC ENGINEERING.

Scientific committee: Kasanický G. (SK), Kohút P. (SK), Navrátil L. (CZ), Poledňák P. (CZ), Sventeková E. (SK), Šimák L. (SK), Weyde M. (D), Zagorecki A. (UK)

ORGANIZING COMMITTEE

CHAIRPERSONS Čelko Ján, Bokůvka Otakar

EXECUTIVE SECRETARY Vráblová Helena

MEMBERS Antal M., Baďura Š., Barabáš J., Belan J., Blinova E., Brída P., Brumerčík F., Brůna M., Buc D., Bulej V., Cíba J., Dobeš P., Dulina Ľ., Ďuranová L., Ďuriníková M., Fanfarová A., Galliková J., Grajcaríková P., Hrbček J., Hyben B., Janík P., Januš P., Jůnová E., Kocifajová L., Kochláň M., Kubjatková M., Kuzmová M., Kvet M., Lieskovský A., Masárová G., Meško M., Michalková M., Nemček B., Nosek R., Pácha M., Plevková L., Pučková K., Sapieta M., Semanová Š., Sýkora P., Šimková I., Štefániková Ľ., Štúňová M., Šušlik Ľ., Titko M., Vaško A., Ždánsky J., Žoltá L.

SECTION 3 INFORMATION AND COMMUNICATION TECHNOLOGIES

REVIEWERS:

Adamec Bohumil
Achimská Veronika
Baďura Štefan
Brída Peter
Cenek Petr
Gokhan Ilk
Grondžák Karol
Gubiš Peter
Hrbček Jozef
Hudec Róbert
Chmulík Michal
Janota Aleš
Jarina Roman
Jurečka Matúš
Kamencay Patrik
Klimo Martin
Kortiš Peter
Kramárová Katarína
Krejcar Ondrej
Madleňák Radovan
Machaj Juraj
Miček Juraj
Paralič Martin
Pirník Rastislav
Počta Peter
Rástočný Karol
Segeč Pavel
Schwartz Ladislav
Ševčík Peter
Šimák Vojtech
Škvarek Ondrej
Václavková Monika
Vaculík Juraj
Vestenický Peter
Zábovská Katarína
Zábovský Michal
Zaitseva Elena
Ždánsky Juraj

Note: **Authors are responsible for language contents of their papers.**

CONTENTS

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Data Collection and Processing in the Queuing System

*Lukáš Achimský
*University of Transport, Faculty of Operation and Economics of Transport and Communications,
Department of Communications, Univerzitná 1, 01026 Žilina, Slovak Republic,
lukas.achimsky@fpedas.uniza.sk

Abstract. The article describes the methodology of data collection necessary for queuing model construction. It describes the derivation of the basic parameter - arrival intensity of the measured values, as well as statistical verification of hypothesis testing the equality of distributions.

Keywords: Arrival intensity, Exponential distribution, Simulation, Equality of distributions.

1. Introduction

Bank XY organized a competition for the best thesis on its operation. Based on the competition, simulation of one of the bank branches was conducted. One of the tasks necessary for the implementation of the simulation was to collect real information at the bank branch on the arrival intensity of branch customers, to determine the distribution of customers arrival and to determine the parameters of this distribution.

When simulating the operation of bank branches, it is necessary to gain information about arrival intensity of the bank customers. [2]

1.1. Place of collection

The data collection was carried out by observing the particular bank branch within a week. Observed data were traced and entered in the prepared sheet. The whole measurements were conducted with the approval of the bank's management. During the observations, we found that customer service was provided weekdays from 9 a.m. to 5 p.m. through three counters and through two counters during the weekend and weekdays from 5 p.m. to 9 p.m. It follows that the intensity of arrivals of the customers tends to be different in those times. It is necessary to take into account both situations while constructing the model due to realistic simulation. The arrival intensity will be described by two different graphs, where the distribution will have different parameters. Assuming that the intensity of arrivals of customers is the realization of exponential distribution in general, we will realize only one test of equality of distributions of values of arrival intensity measured during the entire period.

2. Collection and Calculating

The arrival intensity of the bank customers throughout the period is described in the following graph that informs about the interval of arrival frequency of customers.

The histogram indicates that it is a realization of exponential distribution, starting at $A = 0$ (because the interval between two arrivals of customers will not be less than 0 minutes). The parameter μ can be derived from the number of customers who came to the bank during the reporting period.

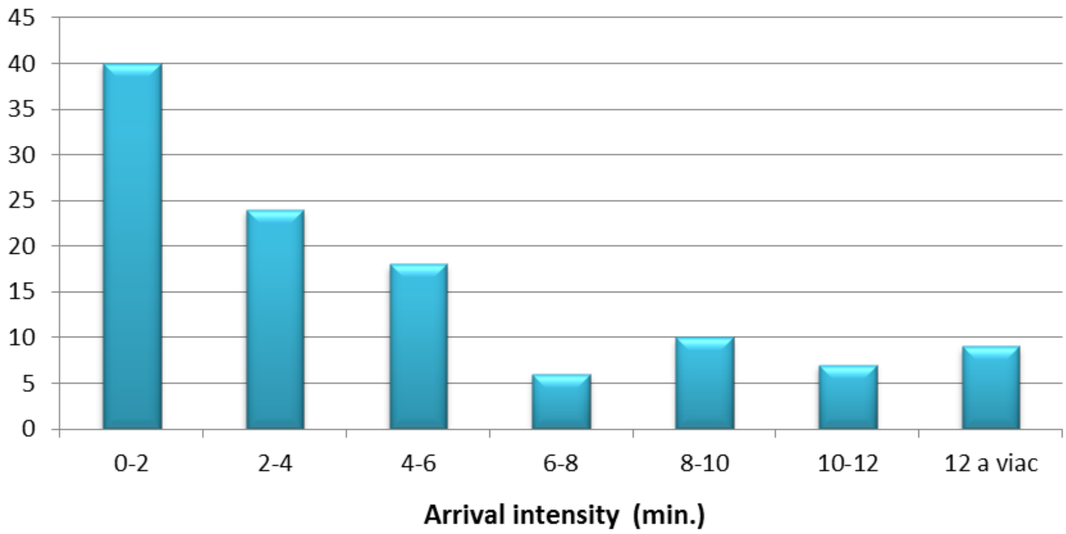


Fig. 1. Total arrival intensity.

However, we will ignore this calculation, because customers were sometimes arriving with a large time lag, which ultimately caused the extreme increase in the value of this parameter and reduced the reporting ability of reality. Arrivals of these customers are included in the last uncertain interval and we will assign it to the range of the previous interval. For this reason, the last interval of arrival of customers will be between 12 and 14 minutes and middle interval will be 13, despite the fact that some arrivals included in this range are higher than 14 minutes. Parameter μ will be estimated by using the arithmetic mean, where the values xi will be the middle intervals of histogram indicated on the x-axis and frequencies ni will be the height of intervals of the histogram indicated on the y-axis. During the reporting period, 114 customers came to the bank and the parameter has the value μ = 4.6. The test of equality of distributions with exponential distribution is described in Table 5. Based on created histogram, we can conclude that arrival intensity is driven by exponential distribution.

Intervals	n_i	x_i	$x_i n_i$	π_i	$\pi_i n$	χ^2
0 – 2	40	1	40	0,351	39,977	0,000
2 – 4	24	3	72	0,228	25,958	0,148
4 – 6	18	5	90	0,148	16,855	0,078
6 – 8	6	7	42	0,096	10,945	2,234
8 – 10	10	9	90	0,062	7,107	1,178
10 – 12	7	11	77	0,040	4,615	1,233
12 – 14	9	13	117	0,075	8,545	0,024
∑	**114**	**X**	**528**	**1,000**	**114,000**	**4,895**

Tab. 1. The test of equality of distributions. [3]

$H_0: f(x) = g(x)$

$H_1: f(x) \neq g(x)$

We reject the null hypothesis, if k> kα,v. In this case, the value of test statistics K is 4.895. At the significance level of 5% is kα,v (k 0.05, 5) = 11.07 so k < kα,v, thereby we can accept the null hypothesis of the equality of distributions.

Arrival intensity distributions between the main and limited operation are described in the following graph. Both distributions have parameter A = 0, because of the above-mentioned reason that arrivals between customers are not less than 0 minutes. Graph of the arrival intensity during the limited operation challenges the assumption of an exponential distribution. It is mainly for the reason that the operation is carried out in non-standard times to perform banking operations, but because of small sample size as well. However, we have proved that the arrival intensity is controlled by the exponential distribution; thereby we can consider this distribution as exponential as well. The parameter μ of these distributions can be analogically estimated as in case of the total arrival intensity, and thus the arrival intensity during a peak traffic is $\mu = 3.94$ and for a limited operation $\mu = 6.85$.

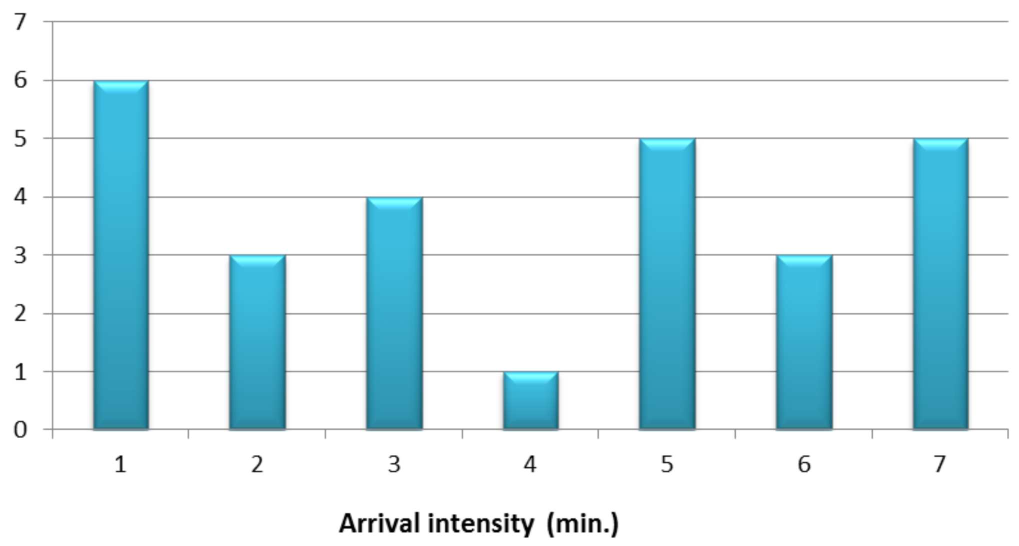


Fig. 2. Arrival intensity in case of 2 counters.

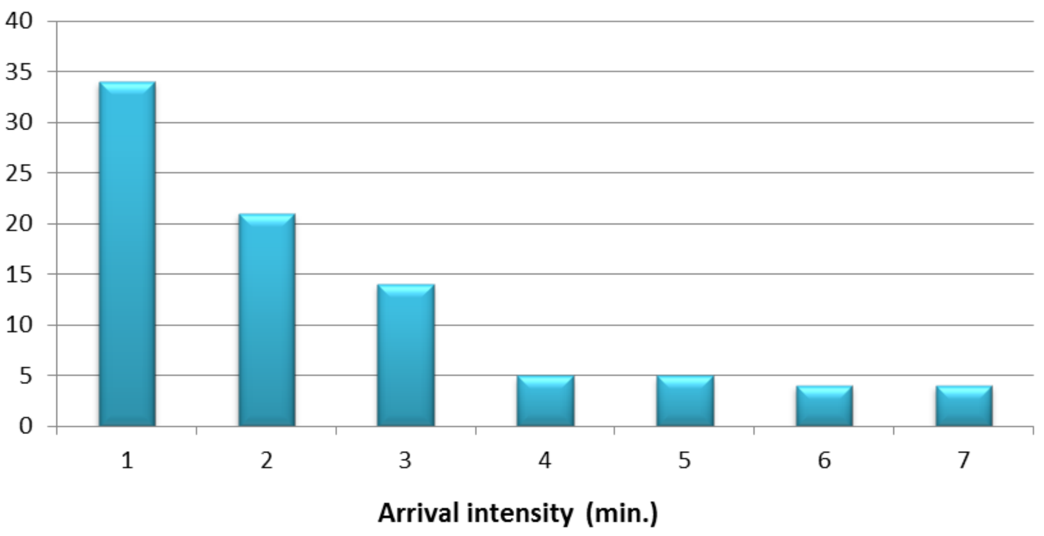


Fig. 3. Arrival intensity in case of 3 counters.

3. Conclusion

In the article, the hypothesis testing the equality of distributions of arrival intensity of the bank customers as one of the main parameters in the construction of queuing simulation model with exponential distribution has been statistically verified. It can be assumed that the intensity of arrivals of customers while solving various tasks in queuing is governed by exponential distribution with different means and possible moves.

References

[1] BRUNOVSKÝ, P.: 2009. Stochastické metódy operačnej analýzy 2006. [online] [citované k 15.4.2012] Dostupné na: <http://www.iam.fmph.uniba.sk/skripta/brunovsky -smoa/smoa_text_08-09.pdf>

[2] JABLONSKÝ, J.: 2002. Operační výzkum. Praha: PROFESIONAL PUBLISHIG, prvé vydanie, 2002. 320 s. ISBN 80-86419-42-8.

[3] KANDEROVÁ, M., ÚRADNÍČEK, V.: 2007. Štatistika a pravdepodobnosť pre ekonómov - II. časť. Banská Bystrica: OZ Financ, 2007. ISBN 978-80-969535-1-6.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Ontology Based E-mail Communication Visualization

*Martin.Antal, *Michal.Zábovský
*University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, {Martin.Antal, Michal.Zabovsky}@fri.uniza.sk

Abstract. In this paper we describe building of ontology based structure of email communication used for analytical purposes. Such structure can be processed by analytical methods used for ontology based networks or graphs. The main advantage of proposed solution is flexible and large scale analytical approach which can be used for visualization of large structures over the current information systems or internet. Our contribution is a solution which combines created ontology with visualization techniques on e-mail communication data where user can easily identify what e-mail address (which person) is important and what is the role of the person in modeled community.

Keywords: ontology, network, visualization, visual analysis, graph theory.

1. Introduction

Communication is the most fascinating possibility in current information systems. The majority of communication is performed by using traditional electronic mail. On the other hand a social network phenomenon defines challenging issue for analyzing traditional email hierarchy by new, network oriented way [1].
Gruber [2] defines ontology as a formal specification of the concepts and the relationships in the domain of interest. That means the ontology is a conceptual schema of objects and relations between these objects. It describes and stores the problem in a way people and software agents can understand it. Data visualization improves understanding of complex data and structures. It is very effective since human brain is tailored to extract huge amount of information by visual perception. Both methods form basic assumption of our work described in this paper.

The paper is structured as follows: Section 2 explains the overall ontology model used in our work. Section 3 discusses methods for visualization of developed network and Section 4 concludes the paper.

2. Ontology model

In computer science and information science, ontology formally represents knowledge as a set of concepts within a domain, and the relationships between pairs of concepts. It can be used to model a domain and support reasoning about entities. Ontology is then a formal, explicit specification of a shared conceptualization.

To create an ontology model we must identify *concepts* (objects) and *roles* (relations) between the concepts in the area of interest [2]. Each concept is then described in details by *attributes*. Ontology theory recognizes two types of attributes/properties: data properties and object properties. *Data properties* stand for attributes with concrete values and *object properties* are references to another objects.

In e-mail communication are defined two important concepts: e-mail and person. Consequently, we have defined another e-mail data properties (subject, receive date, in reply to header/flag) and object properties (receiver, sender, carbon copy, blind carbon copy , reply to e-mail). Every object property references to the instance of concept person. For concept person we have identified only data properties: e-mail address, name. This ontology model can be expanded more in depth and for our purpose is sufficient.

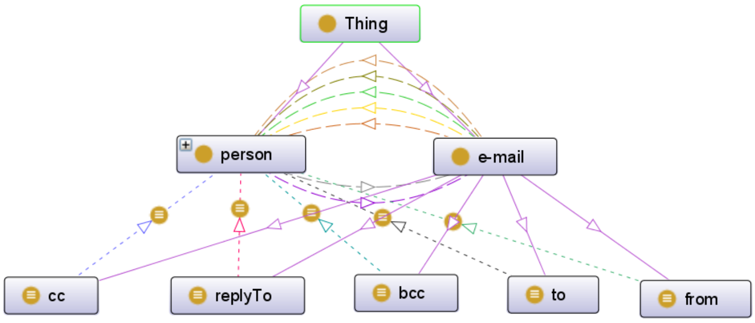


Fig. 1. Ontology model of e-mail communication

Fig. 1 shows the ontology model. The Thing concept represents the concept of all concepts (superconcept). Concepts e-mail and person are classes, the rest of concepts (to, from, cc, bcc, replyTo) are equal classes to class person and define roles they play in conceptual model. Full links between the concepts define hierarchy of information, for example the concept "to" is subclass of the concept "e-mail" and dotted links between concepts represents object properties (i.e. the top dotted link between e-mail and person stands for object property "hasTo" which means that an instance of class e-mail has reference to an instance of class person and this person has role "to").

3. E-mail communication visualization

Ontology model (Fig. 1) was implemented in Protegé software and is visualized by using included plugin called OntoGraf. For the basic overview of the conceptual schema such plugin is sufficient. Our conceptual schema is small but we need to display a lot of instances (more than thousands). OntoGraf is not developed for efficient work with such amount of instances and browsing in these instances is uncomfortable and time-consuming. Thus we must use different approach for visualizing such network structure.
To browse graph structures, specific methods are available. Broad range of measures that characterize graphs, from simple measures, such as the number of vertices and edges that tell the size and sparsity of a graph, to vertex degrees, which tell how locally well-connected each vertex is can be used. Other measures include the geodesic distances in a graph or centrality measures that give a measure of how central in the overall graph each vertex is; for example, PageRank and HITS are measures used to order web page importance as returned from a search engine.

For visualization of graph structures we used tool called Gephi which allows interactive visualization and development of own solution by using its well developed API [3]. Working with two kinds of software brings the problem of incompatible file format. Both tools work with acyclic graphs in the different way. Protegé stores data in ontology format languages as RDF, OWL and Gephi recognizes data in graph file formats as gexf, geml, or dot. It was necessary to convert .owl file format to .gexf file format. To do so, we added new functionality to our application which downloads e-mails from server and stores them into different formats: RDF, OWL, csv, gexf and dot.

After data conversion we applied steps shown on Fig. 2.

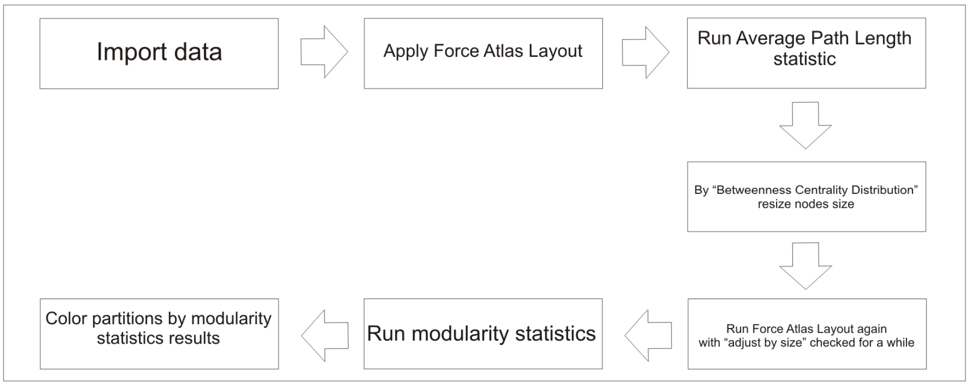


Fig. 2. Steps applied in Gephi

Initially, in the visualization process, data is displayed in a square area and the nodes overlap each other. To eliminate such situation we applied Force Atlas Layout [3]. This method tries to spatialize graph and allows good readability to the viewers. Method is applicable on graphs which nodes varies from 1 to 10 000. If the limit of ten thousands is exceeded improved method Force Atlas 2 [3] can be involved which allows to display nodes up to one million. Next step is running average path length statistic to find important vertices in our data. Details of used algorithm are in [4]. By the ratings of nodes(acquired from statistic) we adjust the size of nodes. After this step we can see different size of nodes but some nodes are still overlapping others. To solve this problem we modified parameters of Force Atlas method and defined Adjust by Size property to be. We are now in the situation when we identified the prominence of the vertices but we still have no idea if or how the network is partitioned. To find partitions or clusters we use a simple heuristic method to unfold the community structure based on modularity optimization. Principle of this method is to iterate network in two phases. In the first phase every node is assigned to one community. Then for each node *i* and his neighbors *j* is calculated value which defines the gain of modularity of replacing *i* from its community and placing it into the *j* 's community. The maximum positive gain is expected for switching else no change is made. The first phase ends when no further improvement can be achieved by repeating this process. In the second phase a new graph is built where communities found in the first phase are nodes and weight of edge between nodes is a sum of weights of linked two nodes-communities. Iteratively these two phases are applied until no more changes are done and the maximum modularity is reached. In detail the method is described in [5]. We can run this method directly in Gephi and use the modularity results to color the vertices and edges by community memberships(the result is shown on Fig. 3).

4. Conclusion

In order to verify our process, we have downloaded private e-mail communication. As an input in Gephi we have a small directed graph with 1326 vertices, where 973 of them represent e-mails and 353 represent specific e-mail addresses (people). Vertices were linked by 3523 edges. Every e-mail node can be linked only with person nodes and vice-versa, so there is no edge between the two of the same node types which implies from ontology model structure. The result of steps described in previous section is shown on Fig. 3.

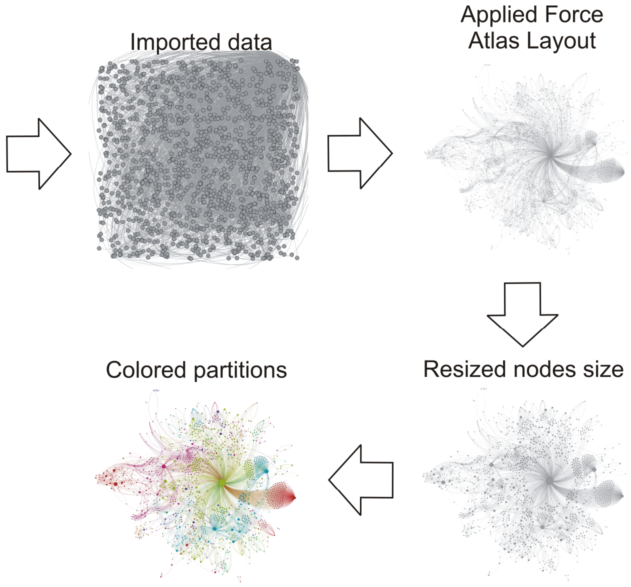


Fig. 3. Steps of process in Gephi

The difference between input and output is significant. From messy and disordered network we obtained a colorful eye-taking graph where user, after turn on showing node labels, can fast identify important communication nodes in groups, between groups or global.

Graph based solution appears to be useful for identification of information-centric clusters inside communities. In practice, just thousands of records can be processed (hundreds or thousands). We were able to produce large datasets represented by graphs but visualization methods were difficult to adopt for large datasets over million edges.

In the future work we will implement additional graph and ontology based metrics in proposed solution to improve analytical and visualization power of the solution. We are expecting analysis of networks with millions of vertices and tents millions of edges [6].[1]

This work was partially supported by the Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic for the Structural Funds of EU, under project ITMS: 26220120050.

References

[1] Zábovský, M.; Matiaško, K.; Zábovská, K.: *Database exploration using metrics and visualization.* Zastosowania Internetu : praca zbiorowa. - Dąbrowa Górnicza: Wyższa Szkoła Biznesu, 2012. - ISBN 978-83-62897-08-7. - S. 159-168.

[2] Gruber, T. R.: *A translation approach to portable ontology specifications.* KNOWLEDGE ACQUISITION, vol. 5, 1993.

[3] Bastian, M.; Heymann, S.; Jacomy, M.: *Gephi: An Open Source Software for Exploring and Manipulating Networks.* 2009.
URL http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154

[4] Brandes, U.: *A faster algorithm for betweenness centrality**. Journal of Mathematical Sociology,vol. 25, c. 2, 2001: s. 163–177.

[5] Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; et al.: *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, vol. 2008, c. 10, 2008: str.P10008

[6] Takáč, L.; Zábovský, M.: *Data analysis in public social networks [Analýza dát vo verejných sociálnych sieťach]*. Present day trends of innovations 2. - Łomża: The State Higher School of Computer Science and Business Administration, 2012. - ISBN 978-83-60571-23-1. - S. 110-115.

[1] Available at http://snap.stanford.edu/data/soc-pokec.html

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Effectiveness of Digital Signature Schemes for Use in Vehicular Communication within Intelligent Transportation Systems

Beňuš, J., Franeková, M., Bubeníková, E.
University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitna 8215/1, 01026 Žilina, Slovak Republic, {Jan.Benus, Maria.Franekova, Emilia.Bubenikova}@fel.uniza.sk

Abstract. The paper deals with a problem of security of vehicular communications with orientation to selection of effectiveness digital signature schemes for assuring authorization of transmission messages in C2C communications. For practical realization the virtual equipment via SW VirtualBox and OpenSSL were used. The results of time relations of the digital signatures realization for three schemes (RSA, DSA and ECDSA) are described and mentioned according to the key length for process of generation and verification of signatures with application to an intelligent car.

Keywords: intelligent transportation system, vehicular network, authorization, digital signature, simulation

1. Introduction

Currently, there are efforts to integrate all the services offered by the ITS (*Intelligent Transportation Systems*) to one large unit [1]. To fulfill this goal, it is important to create communication interconnections between the intelligent vehicles with respect to the standards valid for the wireless communication. For this purpose the VANET (*Vehicular Ad Hoc Networks*) networks are used, where the authentication of authenticity of the offered services is necessary to ensure. In Fig. 1 the principle of communication between the vehicles on each other - C2C (*Car-to-Car*), or between a vehicle and a road side station - C2I (*Car-to-Infrastructure*) is shown.

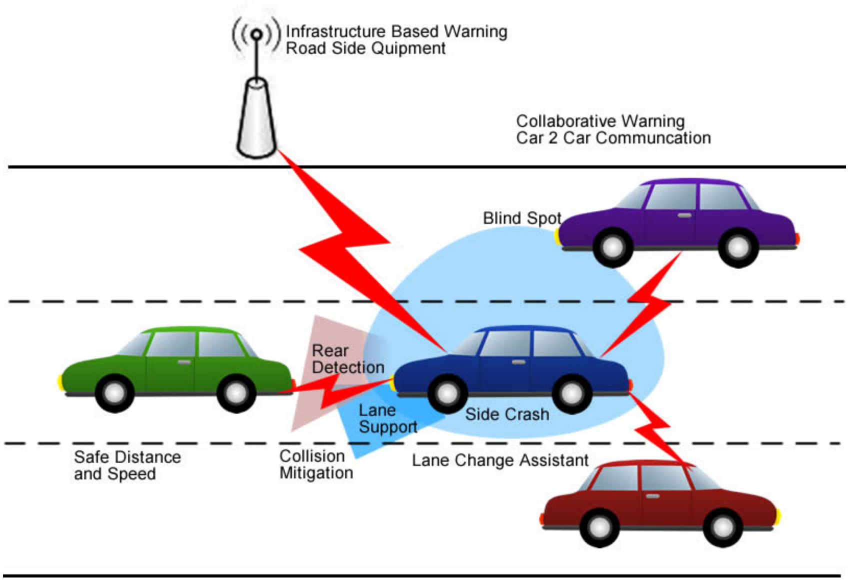


Fig. 1. The principle of communication between the vehicles through the VANET network.

The progressive implementation of the road traffic control applications that use the VANET networks can improve almost all fields of the road traffic. For the communication between road vehicles several new technologies are currently developed [3]. It is the WAVE (*Wireless Access in Vehicular Environments*), CALM (*Communications Architecture for Land Mobile Environment*) and C2C-CC (*Car-to-Car Communication Consortium*) technology, however, for the safety-aimed applications the C2C technology is the most preferred one. The main requirements on the safe communication between road vehicles are the following: message and integrity authentication, nonrepundation of message, recency, access control and the confidentiality of a message. These services are in most cases basically performed by cryptographic techniques. The authors focus only on the assurance the nonrepundation and recency of a message that uses the digital signature cryptographic techniques combined with the time stamps.

2. The Principles of Digital Signature Schemes within Vehicular Networks

Within the scope of authorized communication between the vehicles, a digital signature technique based on asymmetrical cryptography using the PKI (*Public Key Infrastructure*) has been carried, which results from the presumption of CA (*Certification Authorities*) existence [4]. Every CA is responsible for its region assigned at the road infrastructure (it has mostly a stochastic character) and it ensures the identification of every moving node (i.e. vehicle) that is registered within this associated CA, the administration of short-term keys (K) and the certificate distribution ($Cert$). To allow the communication between nodes from different regions, the CA has to provide its certificates to the other CAs (so-called *Crosscertification*). The certificates are released by the CA according to a unique vehicle ID, after the node was registered at the offline mode. After the registration, the distributed certificate has a long-time relevance. Before the sending of a safety-relevant message (in our application - *beaconpacket*), a digital signature is generated in the safety unit of a vehicle V_1 using its private key SK_{V1}. The signature is a function of the message M and the message header H as well. The cryptographic number (the signature), that was created, is added to the message together with the certificate $Cert$, which is connected to an i^{th} anonymous public key of the sender VK^{i}_{V1} that is certified by a associated CA. On the vehicle V_2 side, the received certificate is validated at first (if this has not been already done before) and the received digital signature is verified using the i^{th} public key VK^{i}_{V1} of the vehicle V_1, that is being periodically downloaded by the vehicle V_2 (the other vehicles respectively). At the same time, *geostamp* information, that is located within the message header H, is verified. After these procedures have been done, the safety-relevant message is either accepted, or not. The digital signature generation process by the vehicle V_1 can be expressed mathematically:

$$V_1 \to * : M, H, Sign_{SK^{i}_{V1}}[(M, H) \mid T], Cert_{VK^{i}1}. \tag{1}$$

Where:

M represents the sent safety-relevant message,
H represents the message header,
SK^{i}_{V1} is a temporary private key of the vehicle V_1 in the i^{th} instant of time,
VK^{i}_{V1} is a temporary public key of the vehicle V_1 in the i^{th} instant of time,
T is a time stamp,
$Cert$ is a temporary certificate of the vehicle V_1 (for the anonymous public key VK^{i}_{1}),
$*$ represents the number of receivers (in the case of sending the message to more vehicles).

Current certificate of the vehicle V_1, which is valid in the i^{th} time instant for an anonymous public key of the vehicle V_1 (VK^{i}_{1}), comprises of:

$$Cert^{i}_{V1}\left[VK^{i}_{1}\right] = VK^{i}_{V1} \mid Sign_{SK-CA}\left[VK^{i}_{V1} \mid ID_{CA}\right]. \tag{2}$$

Where:

$Sign_{\text{SK-CA}}$ is a certificate signed by associated certification authority, while the signature is based on the private key SK-CA of the certification authority,
ID_{CA} represents a unique identification number associated certification authority.

Currently, in the commercial sector, a following digital signature schemes based on asymmetrical cryptography are used:

- The digital signature scheme with RSA algorithm.
- The DSA (*Digital Signature Algorithm*) scheme with modified El Gamal algorithm.
- The ECDSA (*Elliptic Curve Digital Signature Algorithms*) scheme with elliptic curves algorithm.

It is important to notice that the execution of the phases of digital signature process based on asymmetrical cryptography is a difficult computational problem for required security level. This can cause time delay of receiving an authorized message at the authorized transmission between several nodes within the road infrastructure and a delayed response to this event in consequence. Therefore it is necessary to use computationally safe digital signature schemes that satisfy requirements on the calculation of signature generation and verification speed. The mathematical description of each schemes can be found e.g. in [5].

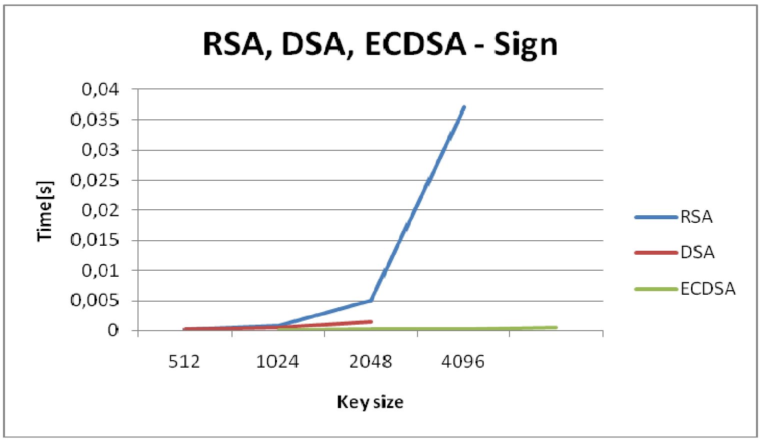
3. Results of SW Execution

We created a simple simulation to demonstrate the difficulty of computation of chosen digital signatures. We assume that the capacity of calculation processor within the On Board Unit (OBU) of a vehicle won't be lower than the capacity of Intel Dual Core processor with frequency of 2,3GHz. We performed the calculations by a virtual device that we created to this purpose using the VirtualBox software and additional software, needed for the proper OpenSSL working. We focused on the RSA, DSA and ECDSA digital signatures. I the future (year 2030) the equivalent key lengths (in bits) for assuring the same level of security with using three types of digital signatures are the following (selection for one level of security): DSA (L=2048, N=224; RSA (k=2048), ECDSA (f=224-255). Within the simulation, the message with predefined length was signed by private key with measured length and the number of signatures during the 10 second interval has been noted. Consequently, the message verification using public key during the 10 second interval has been done. The results are presented in Tab. 1, which contains (from the left): the algorithm name, the key length, the number of messages signed during 10 seconds, the number of messages verified during 10 seconds, the average time of signing one message and the average time of verification of one message in seconds. The graphical results for three selected digital signatures schemes (RSA, DSA and ECDSA) is shown in Fig. 2a) - generation and Fig. 2b) - verification.

	Key	**Private**	**Public**	**Sign [s]**	**Verify [s]**
RSA	512	73148	789777	0.000137	0.000013
RSA	1024	13272	254362	0.000747	0.000039
RSA	2048	2045	64246	0.004873	0.000155
RSA	4096	268	17040	0.037068	0.000574
DSA	512	74480	68644	0.000134	0.000145
DSA	1024	24869	21805	0.000401	0.000459
DSA	2048	6469	5545	0.001533	0.001802
ECDSA	160	92305	24595	0.0001	0.0004
ECDSA	192	73776	18892	0.0001	0.0005
ECDSA	224	57669	14097	0.0002	0.0007
ECDSA	256	47598	10836	0.0002	0.0009
ECDSA	384	22111	4551	0.0005	0.0022
ECDSA	521	11311	2122	0.0009	0.0047

Tab. 1. The results of measured time for generation and verification of digital signature.

a)



b)

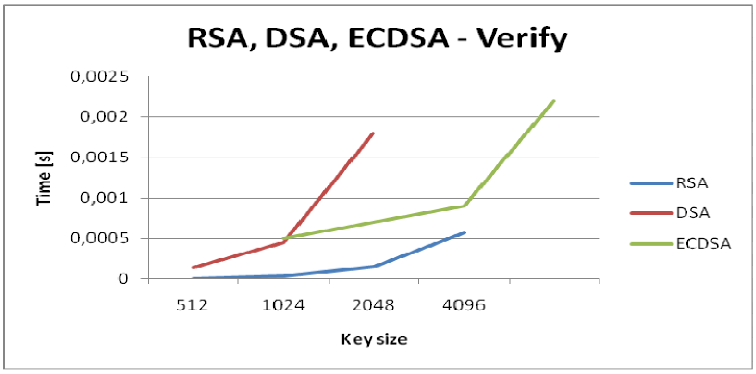


Fig. 2. Example of time relations in RSA, DSA, ECSA digital schemes **.**

4. Conclusion

From the obtained results it is obvious, that in comparison to the ECDSA scheme, the RSA scheme is faster at the signature verification process, but it becomes less effective because of its speed of message signature. The ECDSA appears to be the most effective digital signature scheme, because even if the key lengths are higher (hence the security level is higher), the computational requirements to sign a message are very low and the requirements to verify a message are co-equal to the RSA scheme. ECDSA scheme is recommended to use in secure vehicular architecture [4].

Acknowledgement

This work has been particularly supported by the Educational Grant Agency of the Slovak Republic (KEGA) Number: 024ŽU-4/2012: Modernization of technology and education methods orientated to area of cryptography for safety critical applications.

Bibliography

[1] BUBENÍKOVÁ, E.: Processing of image information in transport applications. In: Slovak. Písomná časť dizertačnej skúšky, ŽU v Žiline, 2012

[2] http://www.sevecom.org/

[3]PAPADIMITRATOS, P. at. al.: Architecture for Secure and Private Vehicular Communications. In: IEEE Communications. Magazine, Nov. 2008. p.1-6, ISBN 1-4244-1178-5

[4] SeVeCom, "Secure Vehicular Communications: Security Architecture and Mechanisms for V2V/V2I, Delivarable 2.1," 2007–2008. In: http://www.sevecom.org

[5] FRANEKOVÁ, M.: Mathematical apparatus for safety evaluation of cryptography and safety codes used in safety related communication system. In: TST 2011, Katowice-Ustrón, Poland, October 2011, selected papers: Springer-Verlag. Berlin Heidelgerg CCIS 104 - Communications in Computer and Information Science. ISBN 978-3-642-24659-3. - s. 126-135.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

QoS Parameters Enhancement by Using Directional MAC Protocols in MANET

*Tomas Cakan, *Vladimir Wieser
*University of Transport, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitna 8215/1, 01026 Žilina, Slovak Republic, {tomas.cakan, vladimir.wieser}@fel.uniza.sk

Abstract. Mobile ad hoc networking involves peer-to-peer communication in a network with dynamically changing topology. The IEEE 802.11 Distributed Coordination Function (DCF) is the dominant MAC protocol for MANET. However, the DCF method could not utilize shared channel efficiently and suffers from various problems. Therefore, in recent years, various methods have been developed to address these problems and accordingly increase the overall network throughput, using the directional antennas. In this paper, a Directional MAC protocol and Directional Virtual Carrier Sensing protocol for directional antennas control was implemented. These two protocols enable efficient directional transmission and reception of packets. The simulation results show that the chosen QoS parameters of the implemented directional MAC protocols were improved compared with IEEE 802.11 MAC.
Keywords: MANET, directional antenna, NS-2, DVCS, D-MAC.

1. Introduction

Wireless mobile Ad hoc networks (MANET) are playing a major role in up-to-date communication networks. Increasing need of business environment for providing cost effective and mobile technology for Internet connection has led to massive implementation of wireless technology. MANET network can be set up or deployed anywhere and anytime because it poses very simple infrastructure setup and none or minimal central administration. These networks are mainly used by community users such as military, researchers, business, students, and emergency services.

Wireless MANET networks typically consist of nodes equipped with omni-directional antennas and used “classical” IEEE 802.11 MAC protocol. Omni-directional antennas spread radio signals in all directions and neighbors around a transmitter or receiver are prevented from transmitting to avoid collisions. The disadvantage of wireless networks with omni-directional antenna lies in limited capacity caused by high interference and low spatial reuse. The signal is received by all nodes within its range. Since the sender intends to send information to just a specific receiver, it is not necessary for all neighboring nodes to receive the signal. As a consequence, the wireless channel is not efficiently used and the receiver gets only a small part of the energy.

Directional antenna offers several benefits in comparison with omni-directional antennas. With directional antennas, transmitter can concentrate most of its power towards the destination and reduce interference to nodes in the vicinity. This leads to extended communication range, increased spatial reuse and less interference to other ongoing transmissions [1].

However, to fully exploit directional antennas it is needed effective directional MAC protocol to be used. A lot of researches are engaged to design an effective directional MAC protocol. We have used two directional MAC protocols. The first one is Directional MAC (D-MAC) protocol and second one is Directional Virtual Carrier Sensing (DVCS) mechanism. These two protocols are compared with each other and with “classical” IEEE 802.11 protocol (omni-directional antenna).

In our work, the simulation model of dynamic MANET in NS-2 simulator environment was created. For simulation evaluation we have chosen average throughput, end to end delay and packet error rate as quality of service (QoS) parameters.

2. Related Work

The IEEE 802.11 [2] standard is the most used MAC protocol in wireless ad hoc networks. It is based on Carrier Sense Multiple Access mechanism with Collision Avoidance (CSMA/CA). The spatial reuse of IEEE 802.11 standard is low because it uses omni-directional antenna. Omni-directional antennas radiate signals in all directions resulting in a circular transmission/reception pattern. The radiation of energy in all directions other than the intended direction not only generates unnecessary interference to other nodes, but also decreases the potential range of transmissions.

2.1. Directional MAC

The Directional MAC (D-MAC) protocol is proposed in many publications [3]. This protocol uses directional transmissions only. It transmits RTS, CTS, ACK and DATA packets directionally. The information about nodes position is assumed to be known beforehand.

2.2. Directional Virtual Carrier Sensing Protocol

DVCS mechanism [4] allows the MAC protocol to determine direction specific channel availability. DVCS does not require any location information about nodes. This protocol is operational in MANET with its independent configuration. The use of RTS and CTS control frames is optional. These control frames can reduce data frame collisions due to the hidden terminal problem. This paper assumes the using of these control frames. Three primary capabilities are added to the IEEE 802.11 MAC protocol for directional communication with DVCS: AOA Caching, Beam locking and unlocking, DNAV settings.

3. Simulation Environment

In this section, simulation model of MANET network was developed in NS-2 network simulator [5]. Two directional MAC protocols were implemented to this simulator. First one is D-MAC protocol and second one is DVCS protocol. These two protocols are compared with the IEEE 802.11 MAC protocol (omni-directional antenna). Simulation model consisted of 30 wireless nodes randomly placed in an area of 1000 x 1000 m. In the mobility scenario, the random waypoint model was set as the mobility model in which each node chooses a random destination. For traffic generation, 5 and 10 CBR (Constant Bit Rate) flows with packet size of 512 bytes were used. Simulation parameters are summarized in Table I. In this work, the average throughput, average end to end delay and packet error rate were used.

Parameter	**Value**	**Parameter**	**Value**
Test area	1000x1000m	Antenna type	Omni-directional
MAC protocol	IEEE 802.11		Directional
	D-MAC	Directional beamwidth	30 degrees
	DVCS	Transmitter power	281.8 mW
Propagation model	Two ray ground	Carrier sense threshold	-70 dBm
Routing protocol	AODV	Traffic type	CBR
Simulation time	100 seconds	Packet size	512 bytes

Tab. 1. Simulation parameters.

4. Simulation Results

4.1. Average Throughput

Figure 1 shows simulation results of average network throughput for the 5 and 10 data flows. The lowest value of average throughput for both data flows was achieved by using classical IEEE 802.11 MAC protocol (omni-directional antenna). DVCS protocol enhances throughput about 250% than that of the IEEE 802.11 MAC protocol and up to 158% over the throughput of the DMAC protocol. From results it is obvious that the highest value was reached with using DVCS mechanism.

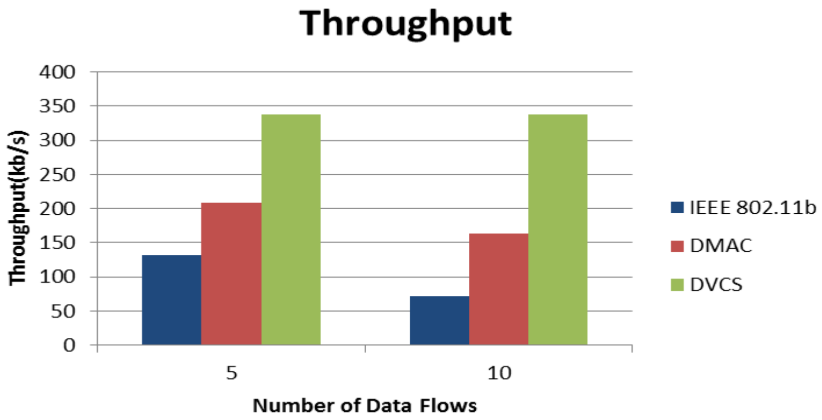


Fig. 1. Average values of throughput for different MAC algorithms and different number of data flows.

4.2. Average Packet Error Rate

Figure 2 shows the values of packet error rate (PER) for two different directional MAC protocols, compared with “classical” IEEE 802.11 MAC protocol. The simulations were created for two different numbers of data flows. From results we can see that the best result reaches the DVCS mechanism. (For 5 data flows 17,52% and for 10 data flows 17,49% PER values were reached). The highest value of PER 86,93% for 10 data flows reached IEEE 802.11 MAC protocol.

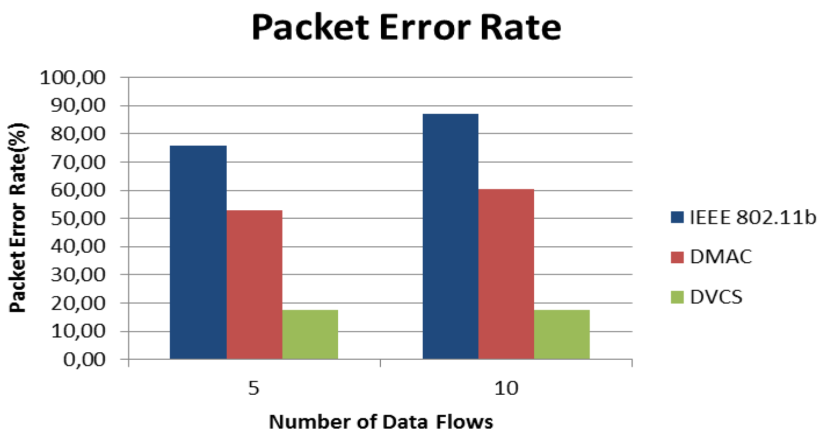


Fig. 2. Values of packet error rate for different MAC algorithms and different number of data flows.

4.3. Average End to End Delay

Figure 3 shows the average values end to end delay for 5 and 10 data flows. Results show that, the highest values of delay for both data flows was reached with IEEE 802.11 MAC protocol. The best value for both data flows was achieved by using DVCS mechanism.

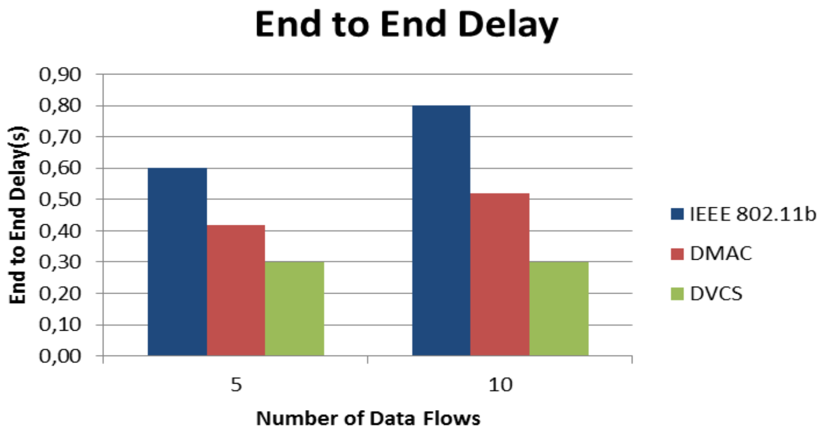


Fig. 3. Average values of end to end delay for different MAC algorithms and different number of data flows.

5. Conclusion

In this paper two directional MAC protocols, Directional MAC protocol and Directional virtual carrier sensing protocol were implemented. These two protocols were compared with “classical” IEEE 802.11 MAC protocol. Comparison was based on changing number of data flows for randomly chosen pair of nodes. The simulation model of MANET network with 30 nodes was created in NS-2 simulator. For simulation evaluations the average throughput, average packet error rate and average end to end delay were chosen.

On the basis of simulation result we can say that DVCS protocol enhances throughput about 250% than that of the IEEE 802.11 MAC protocol and up to 158% over the throughput of the DMAC protocol. This enhancement in network throughput is achieved by increasing the number of simultaneous transmissions and largely by minimized interference.

The results show that by using directional antennas it is possible to increase the network capacity by enhancing QoS parameters. However, efficient directional MAC protocol is needed to full exploit potential of directional antennas. The best results in all cases were obtained by the using of DVCS protocol.

As future work, we plan to investigate the issue of power control over directional antennas for full exploitation of these antennas.

Acknowledgement

This paper was supported by the Scientific Grant Agency VEGA in the project No. 1/0704/12.

References

[1] ZEMIN, W., ZHENGLUN, *A Survey on Directional Antenna Networking*. Wireless Communications Networking and Mobile Computing (WiCOM) 2011, pages 1 – 4, 23 – 25 September 2011.

[2] TAY, C. Y., CHUA, C. K., *A Capacity Analysis for the IEEE 802.11 MAC Protocol*. Wireless Networks 7, pages 159 – 171, March 2001.

[3] NASIPURI, A., YOU, J. YE., HIROMOTO, E, R. *A MAC Protocol for Mobile Ad Hoc Networks Using Directional Antennas*. Wireless Communications and Networking Conference, pages 1214 – 1219, vol.3, September 2000.

[4] TAKAI, M., MARTIN, J., REN, A., BAGRODIA, R. *Directional Virtual Carrier Sensing for Directional Antennas in Mobile Ad Hoc Networks*. MobiHoc 02, June 2002.

[5] The Network Simulator ns-2. [Online]. Available at: http://www.isi.edu/nsnam/ns/.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Novel Model for Application Layer Denial of Service Attacks Detection

[*]Veronika Ďurčeková
[*]University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 2, 01026 Žilina, Slovak Republic, {veronika.durcekova}@fel.uniza.sk

Abstract. This paper focuses its attention on application layer Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks detection, which is a current computer security issue that needs to be solved. Application layer security has gained quite a big importance with the rapid growth of web-based applications. Thus there is a need for developing flexible, reliable and automated security mechanisms, which would be able to detect and respond to threats in real time. This paper highlights detection possibilities and current trends in application layer DoS/DDoS attacks detection and further proposes a novel detection model. Proposed model represents a combination of older detection mechanism with anomaly-based detection technique as a new and perspective detection method.

Keywords: Denial of Service Attacks, Signature-based Detection, Anomaly-based Detection, Threshold Values Estimation.

1. Introduction

This paper focuses its attention on application layer Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks detection. Main purpose of DoS and DDoS attacks is to prevent legitimate and authorized hosts from using a service. Distributed DoS attack is a specific version of DoS attack, which goal is to increase attack intensity by using numerous maliciously misused computers. The topic of DoS and DDoS attacks is really actual problem of computer security area, when there are lots of attacks executed quite frequently these days. These attacks are targeting especially popular and frequently used application servers. The goal is not to steal or misuse sensitive data but to create network congestion or to overload the application server by generating a large amount of traffic addressed to the victim. DoS and DDoS attacks can be performed at different layers, based on what type of resources does the attacker want to exceed. We put our attention to attacks performed at application layer, especially associated with HTTP protocol. Application layer DoS and DDoS attacks are a new sophisticated strategy on how to overload network devices. The main difference between application DDoS and network layer oriented DDoS is in exploiting vulnerabilities of application protocols and not in consuming network bandwidth or overleaping the number of possible parallel connections. This means that in the case of executing application layer DoS or DDoS attack, connections estimated on the network and transport layer have to be established correctly. Thus successful realization of application layer DoS attack requires less number of attacking traffic because the goal is to reach resource limits of a concrete service, which is always lower than the total amount of possible TCP or UDP connections. This makes the detection of application attacks much more complicated because it is difficult to differentiate attacking traffic from the legitimate one.

2. Application Layer DoS Attacks Detection Possibilities and Current Trends

Most DoS and DDoS detection related research was focused on the IP layer, because there are lot of parameters, which can be monitored. These mechanisms attempt to detect attacks by analyzing specific features like arrival rate or header information. However, only a little work has been done in the area of DoS/DDoS attack detection directly on the application layer. Among first detection techniques that were invented to defeat application layer attacks were proposed:

- Client Puzzle Protocol
- Ingress Filtering
- Intrusion Detection Systems or
- Threshold Values.

These older methods did not actually solve the application DoS/DDoS problem clearly enough. Effectiveness of these methods was disputable. [1,2]Today, there is a novel approach for application DoS/DDoS attack detection based on two different principles, which are:

- Signature-based attack detection and
- Anomaly-based attack detection.

1.1. Signature-based Attack Detection

This method for Application layer DoS/DDoS attacks detection is based on monitoring statistical changes. The first step for these methods is to choose a parameter of incoming traffic and model it as a random sequence during normal operation. Next step is to monitor and model attacking traffic and based on the statistical evaluation of previously known attacking data to create a signature of attacking traffic. This detection method can subsequently identify an attack if the monitored traffic matches known characteristics, or in another words signature of malicious activity. Unfortunately, for attackers is it in practice relatively easy to vary the type and content of attack traffic, which makes it difficult to design accurate and universal signatures for DoS attacks [3]. Signature-based detection can be used to detect communication between attackers and their zombie computers through IRC (Internet Relay Channel) but in many cases this communication is encrypted, rendering signature-based detection ineffective. This limits the effectiveness of signature based detection for DoS attacks. Biggest disadvantage of signature-based detection is that it cannot detect unknown attacks, either because the database of known attacks is out of date or because no signature is available yet.

1.2. Anomaly-based attack detection

Another approach is used by anomaly-based detection, which can identify an attack if the monitored traffic behavior does not match the normal traffic profile that was built using a training data. Anomaly-based detection has become a major focus of research due to its ability to detect new attacks, including application DoS attacks. [4] Building a normal profile is most times the first step for all anomaly-based detection techniques. Since there is no clear definition of what is normal, statistical modeling plays a crucial role in constructing the normal profile. An anomaly-based detection method works by building a reference statistical model, which describes the normal behaviour or usage patterns of the monitored system. This detection approach consequently compares the current input with the reference model with the use of similarity metric. A significant difference is then marked as an anomaly. The main advantage of anomaly-based IDS is the ability to detect previously unknown attacks or modifications of well-known attacks, as soon as they are invented. [5,6] Common challenge for all anomaly-based intrusion detection systems is that it is difficult or almost impossible for the training data to provide all types of normal traffic behavior.

3. Hybrid Model for Application DoS/DDoS Attacks Detection

This paper proposes a novel approach in the manner of combining older detection method with current trends in this area. As older detection technique is used Threshold limit estimation, which is combined with the newest perspective detection method, concretely anomaly-based detection.

Principal flow chart, which represents hybrid detection model proposed in this paper, can be seen in the Figure 1.

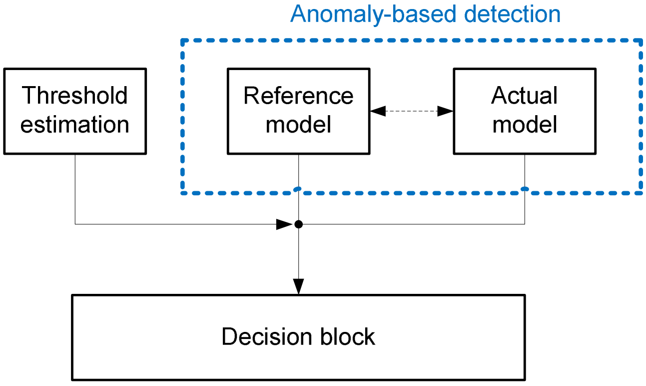


Fig. 1. Hybrid detection model for application layer DoS/DDoS attacks detection.

As it was stated before, proposed model consists of older detection technique and novel anomaly detection. Threshold limit estimation is considered as older technique for application DoS attacks detection. In the process of Threshold estimation a concrete workstation is monitored and scanned. The principle is in setting a limit that can be handled and serviced correctly by corresponding web server. Limitations can be at different layers or parts of the web server. First limitation is related to Operating System installed on the server and its maximum number of parallel processes that can be performed simultaneously. Second limitation that needs to be checked is related to hardware capabilities of the server, concretely Random Access Memory (RAM) size since in the official apache documentation is stated that the single biggest hardware issue affecting webserver performance is RAM. Last, but usually most affecting limitation for Threshold value estimation is software limit. When talking about Apache webserver, which is point of interest in this paper, software limits depend on the type of Multi-Processing Module (MPM) which is configured in the main apache configuration file. This paper deals with apache MPM configuration set to Prefork, with the respect to default values. In such a scenario, apache software limiting factor is MaxClients directive which sets the limit on the number of simultaneous requests that will be served. Any connection attempts over the MaxClients limit will normally be queued up to a number based on ListenBacklog directive. Default value of MaxClients directive for non-threaded Prefork MPM is 256. [7] Gathered OS, HW and SW limitation values are subsequently compared and lowest one is finally set as the definite Threshold value, which can be also named as Saturation Level of appropriate web server. Next step after this estimation is to monitor current real time traffic and monitor whether the actual traffic does or does not exceed this Saturation Level.

Second part of proposed model is anomaly-based detection, which consist of two separate blocks:

- Reference model and
- Actual model.

Reference model is built based on previously captured traffic on the inbound web server interface. There are many statistical parameters that can be monitored. Detection model presented in this paper focuses especially on:

- Intensity of incoming HTTP requests flow
- Request method rate
- Service rate
- Cumulative number of incoming requests per predefined time interval
- Time difference between two received requests
- Mean value and dispersion of upper stated attributes

Previous stated attributes are firstly monitored in a proper period of time and consequently statistically evaluated. Based on these results a reference mathematical description or model of normal traffic behavior and standard web server service features is created.
After the Reference model is built precisely, proposed detection mechanism is able to switch to real-time mode, where there is sniffed actual traffic appearing on the inbound web server interface. Exactly the same statistical attributes that were counted in the phase of Reference model creation will be now counted real-time with the goal to create actual or current mathematical model of incoming traffic. This Actual model is renewed and recounted with a specified time period, which is related also with the actualization of Reference model. Recalculation and actualization of Reference model has to be done to provide correctness of attack detection and low false positive rate.

Last functional block of proposed detection mechanism is a Decision block where values gathered in the Reference and Actual statistical model are compared. Diversion between those values is evaluated and based on the amount of difference, a parcial decision of suspicious or not suspicious traffic is taken. Next part of the Decision block is the comparison of cumulative amount of actual incoming traffic with the Threshold values stated in separate block. In this phase it is important to monitor how much does the cumulative traffic lead to reach Saturation limit. The final decision of attack presence is taken based on consideration whether both of previously taken parcial decisions were attack positive or not.

4. Conclusion

In this paper focuses its attention on the topic of application layer DoS and DDoS attacks detection as a current issue in the area of computer security. Traditional detection techniques and newest trends in detection are briefly described. This paper further proposes a novel hybrid model for application layer HTTP DoS and DDoS attacks detection which is based on the combination of older detection technique with the newest detection trend. Detection principles are described. This hybrid detection model is still under development thus changes in the decision block can still appear.

References

[1] DURCEKOVA, V., SCHWARTZ, L., FRANEKOVA, M. *Evolution of IDS with the focus on application layer DoS attacks detection.* RTT 2012: 14th international conference: Slovakia, [Vrátna], 12-14 September 2012, ISBN 978-80-554-0569-8. - S. 103-106.

[2] DURCEKOVA, V., SCHWARTZ, L., SHAHMEHRI, N. *Sophisticated denial of service attacks aimed at application layer.* ELEKTRO 2012: 9th international conference: Žilina - Rajecké Teplice, Slovakia, May 21st-22nd. ISBN 978-1-4673-1178-6.

[3] XIE, Y., YU, S. Z. *Monitoring the Application-Layer DDoS Attacks for Popular Websites.* Networking, IEEE/ACM Transactions, vol.17, no.1, Feb. 2009, pp.15-25.

[4] XIE, Y., YU, S. Z. *A Novel Model for Detecting Application Layer DDoS Attacks*. Computer and Computational Sciences, 2006. IMSCCS '06. First International Multi-Symposiums, vol.2, 20-24 June 2006, pp.56-63.

[5] ANKALI, S. B., ASHOKA, D. V. *Detection Architecture of Application Layer DDoS Attack for Internet*. Int. J. Advanced Networking and Applications, Volume: 03, Issue: 01, 2011, pg. 984-990

[6] XIANG, G., MIN, W. *Applying Semi-supervised Cluster Algorithm for Anomaly Detection*. Information Processing (ISIP), 2010 Third International Symposium on , vol., no., pp.43-45, 15-17 Oct. 2010

[7] The Apache Software Foundation, Projects and Apache Documentation, http://projects.apache.org/docs/index.html

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

General Sound Pattern Detection Using Elementary Sound Models Database

Róbert Gubka, Michal Kuba
University of Žilina, Faculty of Electrical Engineering, Department of Telecommunication and Multimedia, Univerzitna 1, 01026 Žilina, Slovak Republic, {robert.gubka, michal.kuba}@fel.uniza.sk

Abstract. This paper describes a method for general audio pattern modeling and retrieval based on database of elementary sound models. For creation of this database, we were mainly inspired by automatic speech recognition and keyword spotting system, where larger models are created as a concatenation of pre-trained low-state models of basic speech units. In the paper, specification of the elementary sound model and the methodology of elementary sound models database creation is described and experimental results on selected sound events detection and noise robustness are presented.

Keywords: acoustic pattern, statistical modeling, elementary model, events detection.

1. Introduction

With the rapidly increasing number of digital multimedia documents and audio-visual data comes also a need for effective and robust methods of content-based data analysis. Audio classification, sound events detection, and pattern searching algorithms can be applied in many different sectors, e.g. for automatic file indexing and categorization in multimedia database management, for traffic condition evaluation in intelligent transport systems, or for significant audio events spotting in surveillance and environment inspection systems [1].

Current research in field of audio content analysis is mostly focused on audio classification problem, where unknown data sample or segment of audio record is assigned to one of the pre-trained sound classes, mostly adopting statistical modeling [2],[3] or machine learning methods [4]. Although such type of classification provide good results, the number of sound classes within specific application is limited and the classification can be very coarse. Adding a new sound class to the system requires a sufficient amount of training examples and computationally demanding training process.

Unlike these methods, our proposal is focused on pattern searching rather than classification, so that instead of using pre-trained sound classes, our system work as query-by-example, with specific sound events and audio patterns. For our implementation, we have been inspired by speech recognition and keyword spotting systems based on phoneme models [5], where longer sentences are modeled as a concatenation of pre-trained phoneme models. Our effort was to propose and create a finite database of low-state models estimated from a huge amount of audio data using methods of cluster analysis. These models serve as basic acoustical units for general sound modeling.

2. Elementary Sound Models

In our concept of general audio pattern modeling, the elementary sound is defined as a segment of acoustic signal, within the temporal, spectral, and statistical parameters are static and homogeneous. Individual elementary sounds differ in these characteristics - acoustic observations within the specific elementary sound are very similar and diverse among different elementary sounds. In comparison with speech recognition systems, elementary sound can be understood as a analogy to basic acoustical-linguistic unit (phoneme). As any sentence in particular language can be modeled with finite number of phoneme models, are assumption was any general sound can be modeled as a concatenation of a short elementary sound models from finite database. Our effort was to test this assumption and construct a database of elementary sound models.

2.1. Model Specification

For statistical description of elementary sounds, we adopted the well-known hidden Markov models. A hidden Markov model is an effective parametric representation for a time-series of observations, such as feature vectors measured from natural sounds [6].

A continuous-density hidden Markov model with N states consists of a set of parameters that generally comprises the N-by-N transition matrix, the initial state distribution, and the parameters of the state densities (weights, means and diagonal variances of the state).

Based on our previous research, we adopted a semi-continuous-density output function, where each state of the model shares the same means and variances obtained by cluster analysis. The resulting density function of particular state is determined only by the unique vector of weighting coefficients. In addition, we also omit the transition probabilities, assuming them equal. This simplification facilitates both the training and the decoding process and decreases the computational and memory requirements.

2.2. Database of Elementary Sound Models

Most important part of our proposal is the creation of elementary sound models database. In process of creating the database, we adopted procedures and techniques of data mining, cluster analysis, and Viterbi alignment. The methodology of database creation comprise following steps:

1. To make our system universal and flexible, a huge amount of various sound examples was collected. This sound database consists of short sequences of different sound categories, environmental noises, machinery sounds, music, animal and human sounds, and speech of different languages.
2. Collected data were processed and the mel-frequency cepstral coefficients, together with their first and second order time derivatives, were extracted from 30 ms windows with 10 ms overlap.
3. To obtain the Gaussian mixture components for output density function, processed data vectors were clustered into defined number of clusters using the k-means algorithm. In each cluster, the mean vector and diagonal covariance matrix were computed.
4. Preliminary estimation of weighting coefficients vectors of the elementary sound models was made from one second long segments of audio stream. From each segment, a 3-state model was trained. Because of the huge number of obtained models, their weighting vectors were clustered and their number was reduced.
5. All of the estimated models were added to the decoder search space. Audio stream created by random concatenation of sound examples from whole sound database was created and decoded into sequence of elementary models. Parameters of models were subsequently iteratively optimized, according to assigned sound sequences, with the Baum-Welch algorithm, using the HMM Toolbox [7]. Models without training data were discarded and the process of decoding and training was repeated with newly estimated models.

After specified number of iterations, the number of models was reduced and did not change any further. Finally, we ended up with 379 of 3-state models with 64 Gaussian mixture components, forming the database of elementary sound models.

2.3. Audio Pattern Modeling and Decoding

The usage of the created database of elementary sound models requires a robust sequential decoder. For this purpose, we adopted algorithm for keyword spotting, originally proposed in [8]. The decoder is based on Viterbi algorithm with propagation of accumulated score within the looped network of units and fillers and computing the confidence of each unit. The units represents searched acoustical patterns, created as a concatenation of elementary sound models according to training examples. As the fillers, we adopted the elementary sound models themselves.

3. Experimental Results

Performance of proposed method on selected tasks were evaluated using standard precision and recall metrics, defined as follows:

$$Precision = \frac{N_{true}}{N_{total}}; \quad Recall = \frac{N_{true}}{N_{target}}; \quad F = \frac{2.Precision.Recall}{Precision + Recall} \qquad (1)$$

where N_{true} denotes the number of true positives detections, N_{total} the number of both true and false positive detections, and N_{target} the number of target detections.

For first experiment, we adopted the AudioDat sound database, designed in [9] for audio classification purposes. The database consists of sound examples divided into six different sound classes listed in Tab.1. Achieved results were evaluated on the level of acoustical observations (frames). For best results, a proper level of confidence threshold was set for each unit respectively.

SOUND CLASS	PRECISION [%]	RECALL [%]	F-measure [%]
applause	69.04	62.54	62.95
crying	93.06	87.83	90.11
laughing	34.69	79.93	42.16
music	92.57	94.21	93.25
noise	84.53	91.44	87.33
speech	96.88	96.80	96.81

Tab. 1. Experimental results on AUDIODAT database.

The second experiment was focused on noise robustness evaluation in task of gunshot detection. Results for six SNR levels are listed in Tab. 2. The testing records were created by adding white noise of corresponding level to the signal. Figure 1 shows the precision vs. recall graph for models trained using different training examples. Good noise robustness was achieved by the selected audio features (MFCC + derivatives) and selecting appropriate confidence threshold with particle swarm optimization algorithm. Although the recall level may seem to be low, it must be pointed out that each model was trained using only one example and the results were evaluated on the level of frames against a hand annotation. The overall detection rate achieved about 80 %.

SNR	***Clear***	***26 dB***	***15dB***	***6 dB***	***4 dB***	***2 dB***	***-1.2 dB***
P [%]	97.66	76.92	82.22	80.07	83.74	80.38	78.23
R [%]	65.04	67.16	63.40	51.73	51.56	53.37	49.84
F [%]	75.20	70.09	70.79	62.55	63.75	63.67	60.20

Tab. 2. Experimental results on gunshot detection and noise robustness.

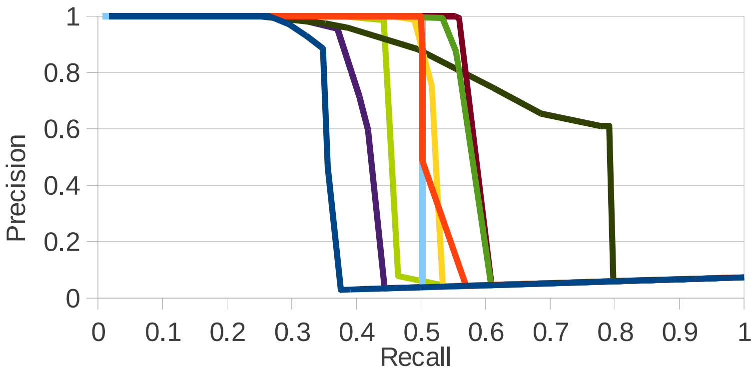


Fig. 1. Precision vs. recall on gunshot detection for different models.

4. Conclusions

In this article we evaluated the possibility of using the elementary sounds as a basic acoustical units for audio modeling in audio pattern recognition task. Preliminary results on artificial records proved our assumption, that different general sound examples can be modeled as a concatenation of elementary sound models from finite database. Experimental results on selected sounds achieved average precision of 78% and recall of 85%. Better results were obtained for sound examples with obvious time-progression characteristics, such as music and speech. Conversely, for “chaotic” sounds, like applause and laughing, the obtained results were significantly lower. Experimental results on noise robustness in gunshot detection task shows, that our proposed method is suitable even for noisy records.

For our future work, we plan to adopt different audio features selected by optimization algorithm and apply the unsupervised clustering method for determination of the optimal number of Gaussian components for mixture output density function and the number of elementary sound models.

References

[1] E. Wold, T. Blum, D. Keislar, J. Wheaton, *Content-based classification, search, and retrieval of audio*, IEEE Multimedia, vol.3, no. 3, Sept. 1996,pp. 27-36

[2] R. Cai, Lie Lu, A. Hanjalic, Hong-Jiang Zhang, Lian-Hong Cai, *A flexible framework for key audio effects detection and auditory context inference*, IEEE Transactions on Audio, Speech, and Language Processing, vol.14, no.3, May 2006, pp. 1026- 1039

[3] A. Mesaros, T. Heittola, A. Eronen, T. Virtanen, *Acoustic event detection in real life recordings*, in 18th European Signal Processing Conference, 2010

[4] Nello Cristianini, John Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press; 1 edition (March 28, 2000), ISBN-13: 978-0521780193

[5] Wollmer, M.; Eyben, F.; Schuller, B.; Rigoll, G., *Robust vocabulary independent keyword spotting with graphical models*, Automatic Speech Recognition & Understanding, 2009. ASRU 2009. IEEE Workshop on , vol., no., pp.349,353, Nov. 13 2009-Dec. 17 2009

[6] L. Rabiner and B.-H. Juang. *Fundamantals of Speech Recognition*. Prentice Hall, 1993.

[7] K. Murphy. (June 2005). Hidden Markov Model (HMM) Toolbox for Matlab.

Available: http://www.cs.ubc.ca/murphyk/Software/HMM/hmm.html

[8] J. Nouza, J. Silovsky, *Fast keyword spotting in telephone speech*, Radioengeneering 2009, vol. 18, no. 4, Dec. 2009, pp. 665-670

[9] R. Jarina, M. Paralič, M. Kuba, J. Olajec, A. Lukáč, M. Dzurek, *Development of a Reference Platform for Generic Audio Classification*, In Proc. 9th Int. Workshop on Image Analysis for MultimediaInteractive Services, WIAMIS’08., Klagenfurt, Austria, 2008

Transcom 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Open Hardware and Open Software Platforms in Logical Systems Education

*Adam Jaroš
*University of Žilina, Faculty of Management Science and Informatics, Department of Technical Cybernetics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, adam.jaros@fri.uniza.sk

Abstract. Popular open hardware Arduino and Raspberry Pi microcontroller boards are excellent choice for education of basic of digital logical systems. Together with classical approach when students verify their logical design on breadboards with integrated circuit and discrete components it extends their experiences. In addition an open software Google Android OS mobile platform can easily be used to create first real world applications. On lower level hardware represents connection with real world through wide range of sensors. On higher level Android application controls servos, LEDs and other electronic devices through analog and digital signals. Hardware module (Arduino, Raspberry Pi) and Android device (tablet, phone) can be linked together with USB cable or Bluetooth or Wi-Fi.

In this paper I present introduction of open hardware modules – Arduino and Raspberry Pi and other supplementary tools appropriate for Logical system education especially for simpler implementation of Moore and Mealy automata. Presented platforms can serve as extension for open software Android OS.

Keywords: Arduino, Raspberry Pi, Google Android OS, finite state automata, logical systems, education, Moore, Mealy.

1. Introduction to Microcontroller with Arduino

What is Arduino? It means three things – a physical piece of hardware, a programming environment and a community [1], it is shown on Fig. 1. Philosophy is to offer as simple and powerful open source physical computing platform as possible. It is “open source” hardware open and free to inspect and modify and at the same time community-build with all its examples as “playground”.

Arduino offers wide hardware variety. There are different modules [2] of different sizes and different number of I/O (Input-Output) signals.

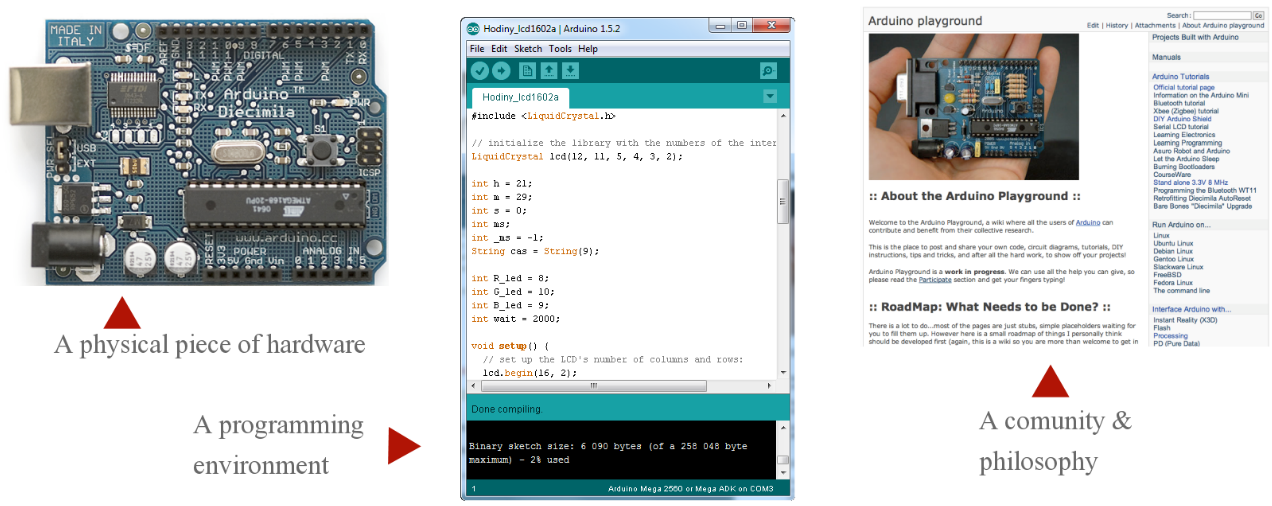


Fig. 1. A philosophy of Arduino platform based on AVR ATmega is to simplify programming as possible – connect Arduino to PC via USB cable; load example in programming IDE and upload compiled code to board on *one click*.

Without going into details Arduino offers quite a few digital and analog I/O pins as well as hardware and software serial ports. This enables direct connection of many simple I/O sensors and actuating units and also complex modules, named “shields” (Bluetooth, Wi-Fi, Data logger, Graphical LCD display, etc.) as shown on Fig. 2.

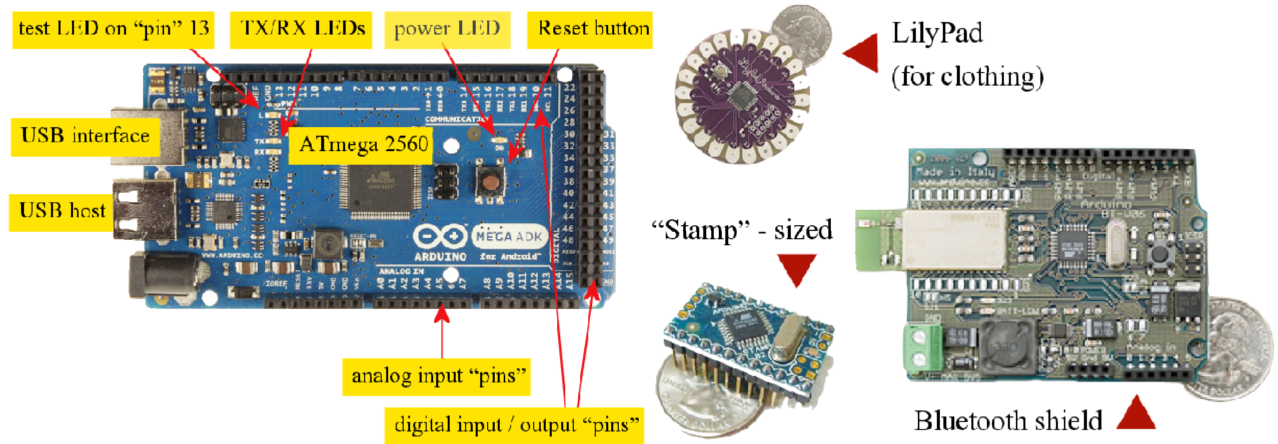


Fig. 2. Arduino MEGA ADK (*left*) has USB host for direct cable connection with Android device. Stamp sized and LilyPad are specialized version of basic module. Bluetooth module (*right*) is one of many “shields” which extend Arduino platform. Quite a few shields can be connected on top of Arduino board on each other.

Arduino capabilities in short, details can be found in [3], model Arduino MEGA ADK:

- 256 kBytes of Flash program memory of which 8 kBytes is used by boot loader,
- 8 kBytes of Static RAM (SRAM), 4 kBytes EEPROM,
- 16 MHz crystal oscillator, power supply +5V / TTL compatible,
- 54 Inputs and Outputs, PWM up to 14, USB Host, USB Interface, 4 Serial ports,
- Completely stand-alone; doesn’t need a computer once is programmed.

Arduino programming software is like text editor. Simplified IDE, show on Fig. 3, can view, write and edit programs called “sketches” (extension *.ino*) and then program it into hardware. After Moore or Mealy state automata is verified on breadboard then it is time to implement it as “sketch” on Arduino platform; connect it through I/O pins with real world inputs and outputs and verify its function. We can also create intelligent *tester* of automata for its automatic verification.

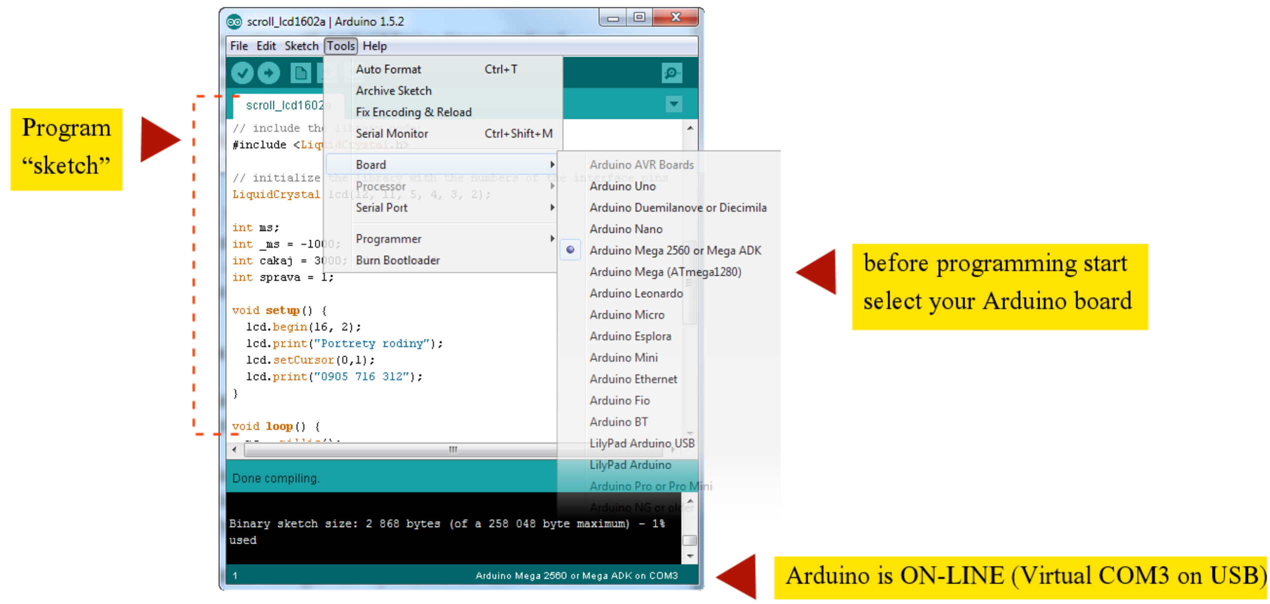


Fig. 3. IDE for Arduino board family programming is learned with easy. After become familiar with code syntax which is similar to C student can explore libraries of additional accessories like LCD display or Data logger, etc.

1.1. Arduino platform support tools – Fritzing and Processing

Community around Arduino has grown in last two years very rapidly. To offer maximum of platform there was created two useful tools.

First of two is Fritzing [4]. This program allows users to document their prototypes, share them with others, teach electronics in a classroom, and layout and manufacture professional pcbs (printed circuit board), shown on Fig. 4. All drawings are colorful and can be used in interactive education process. Program is still actively developed by open source community.

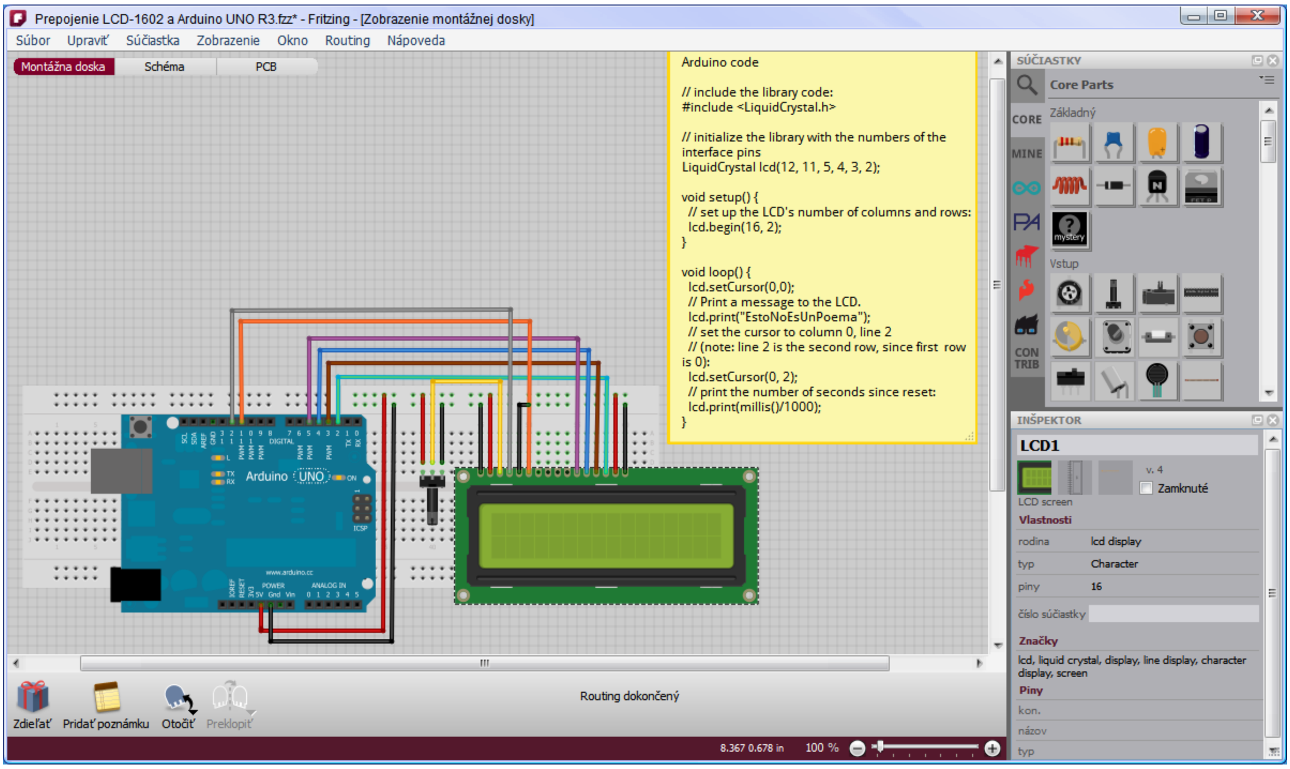


Fig. 4. Fritzing contains all basic Arduino boards. Enable drawing wires and most common electronic components. In addition all components are fully vector and editable. To create new component is quite straightforward.

Second useful tool is Processing [5]. It is an open source programming language and environment for people who want to create images, animations, and interactions. Initially developed to serve as a software sketchbook and to teach fundamentals of computer programming within a visual context. As expected it runs on Windows, Linux and Mac OS X.

Processing can fully *substitute* an Android device and its relative complicated programming. Its power is in simple but powerful language syntax. To simply draw time slope of more logical signals in *real time* is main advantage of such tool, show on Fig. 5.

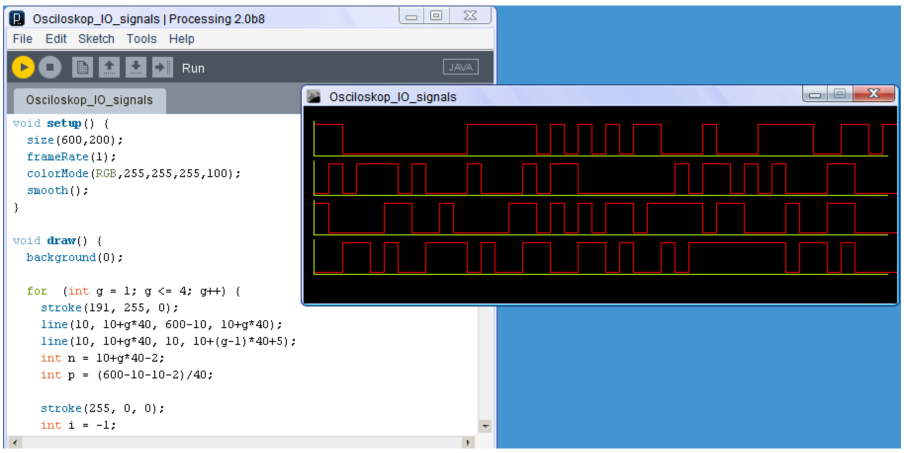


Fig. 5. Processing IDE use the same philosophy of minimalistic design as Arduino IDE. Graph of four logical signals.

2. Raspberry Pi – open hardware computer

Raspberry Pi Foundation [6] and community provide an affordable platform for generations of youngsters to get into the wonderful world of computing in a truly meaningful way [7, 8, and 9], on Fig. 6. Most popular operating systems are Raspbian, Risc OS, Raspbmc and FreeBSD.

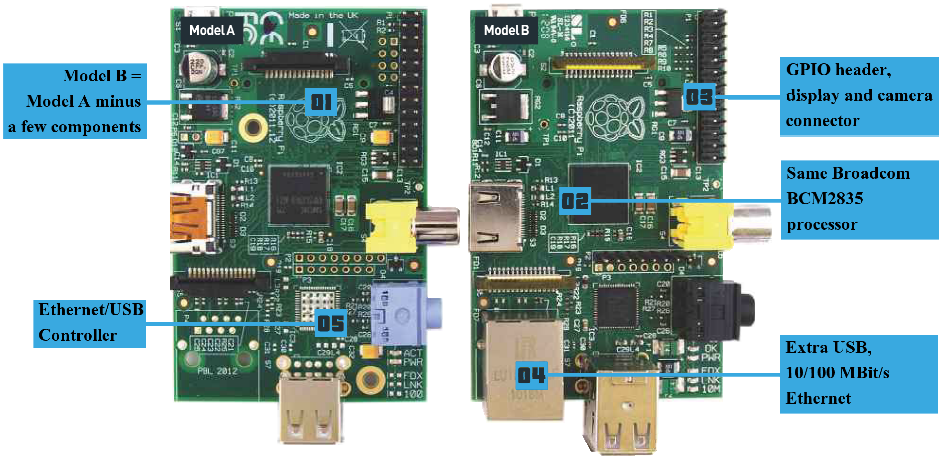


Fig. 6. Comparison of two models A and B of Raspberry Pi open hardware platform.

3. Conclusion

Arduino and Raspberry Pi represents a real *proof* what is possible in area of open hardware. Many high school, universities and hobbyist realized that open hardware is vital and can serve well in many scenarios [10, 11, and 12].

I believe that teachers and students in education of basic of logical systems can easily learn principles and at relatively low cost get large *benefit*. This also can be utilized in Android programming courses in area like control of *intelligent buildings* [13, 14].

Acknowledgement

This work was supported by grant VEGA 1/1099/11: Modeling and simulation of Dynamical Interaction in Environment of Driver/Car/Traffic Situation.

References

[1] Web page: *Arduino – Open hardware*. Url: http://arduino.cc/en [March 25, 2013]

[2] Web page: *Adafruit.com – Arduino and accessories*. Url: http://adafruit.com/ [March 25, 2013]

[3] Web page: *Arduino ADK – tech specification*. Url: http://arduino.cc/en/Main/ArduinoBoardADK [March 25, 2013]

[4] Web page: *Fritzing – prototype tool*. Url: http://fritzing.org/ [March 25, 2013]

[5] Web page: *Processing – programming tool*. Url: http://processing.org/ [March 25, 2013]

[6] Web page: *Raspberry Pi – Manufacturer*. Url: http://www.raspberrypi.org/ [March 25, 2013]

[7] SCHMIDT, M. *Raspberry Pi. A Quick-Start Guide.* The Pragmatic Programmers, 2012, ISBN 978-1-937785-04-8.

[8] RICHARDSON, M., WALLACE, S. *Getting Started with Raspberry Pi.* O'Reilly, 2012, ISBN 978-1-449-34421-4.

[9] MEMBREY, P., HOWS, D. *Learn Raspberry Pi with Linux.* Apress, 2013, ISBN 978-1-4302-4821-7.

[10] WINCHER, D. *Learn Electronics with Arduino.* Apress: Technology in Action, 2012, ISBN 978-1-4302-4366-6.

[11] PURDUM, J. *Beginning C for Arduino.* Apress: Technology in Action, 2012, ISBN 978-1-4302-4776-0.

[12] BORMER, M. *Beginning Android ADK with Arduino.* Apress, 2012, ISBN 978-1-4302-4197-3.

[13] FALUDI, R. Building Wireless Sensor Networks. O'Reilly, 2011, ISBN 978-0-596-80773-3.

[14] RILEY, M. *Programming Your Home. Automate with Arduino, Android and Your Computer.* The Pragmatic Programmers, 2012, ISBN 978-1-93435-690-6.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Translation Plans in Container Terminal Simulation

*Michal Kocifaj
*University of Žilina, Faculty of Management Science and Informatics, Department of Transportation Networks, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, Michal.Kocifaj@fri.uniza.sk

Abstract. The container terminals are one of the most important and most complex parts of the supply chain. At the present, many new terminals are built and the existing ones are extended. A simulation study can save a lot of resources, which could otherwise be lost in unnecessary or inappropriate terminal equipment. There are many simulation tools which supports the simulation of container terminals. All of these tools provide some type of definition of typical terminal operations. We present a possible approach to the definition of these operations using translation plans. The translation plans simplify description of modelled terminal, and can be applied as an alternative to the currently used flowcharts. They can be utilized not only in simulations of container terminals, but also in modelling of warehouses or manufacturing systems.

Keywords: simulation, container terminal, translation, flowchart.

1. Introduction

The container transport has grown dramatically since the first regular sea container service began in 1961 [1]. Nowadays, more than 60 % of goods transported over the sea are containerized and the importance of container terminals is considerably increasing as well. Many new terminals are built and numerous existing ones are going to be extended.

Building or extension of any bigger container terminal is hardly conceivable without preliminary comprehensive study during design of the terminal. If we underestimate the importance of the design phase, many unexpected situations encounter, and future interventions in terminal infrastructure and operation can be very costly. Operation research or simulation study can predict these problems. If we choose either of these methods, we have to describe the functionality of the modelled system (container terminal, in our case). We typically create either a corresponding mathematic model, or a simulation model.

In this paper, we address only the description of translation process for the simulation purpose. Description of translation process using flowcharts is implemented in many simulation tools, including e. g. simulation package Rockwell Automation's Arena [2] or simulation tool Villon [3]. We have developed translation plans as a support tool or as an alternative for these flowcharts.

The remainder of this paper is organized as follows. The next section describes the creation, composition and usage of translation plans, and their comparison with flowcharts. The viability of the proposed approach is shown on a test model of container terminal in the third section. The conclusion summarizes the advantages of translation plans and reveals their possible improvements and extensions.

2. Description of Translation Process

The translation process in simulation tools is typically described by flowcharts. We have developed an alternative to them called translation plans. Both approaches are used to define the sequence of operations, which have to be done to transfer the container from its current position to its destination. We will briefly compare these approaches on a train unloading operation depicted on Fig. 1 (focus is given on a single container transfer). The container is unloaded from the train by crane and transferred to the temporary storage. Then, it is picked up by a reach stacker and transferred to the final storage. This process corresponds in reality to the unloading of empty containers and can be described by a flowchart shown on Fig. 2 or by a translation plan listed in Tab. 1.

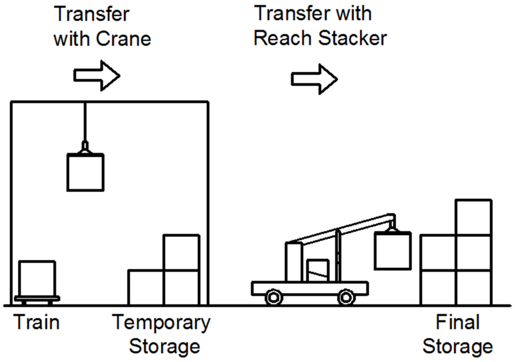


Fig. 1. Two-step translation of container.

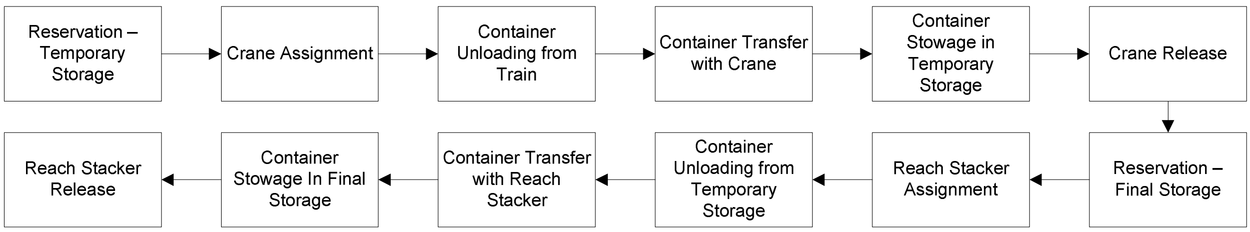


Fig. 2. Description of the unloading operation from Fig. 1 using flowchart.

<table>
<tr><td>Start: Train Track</td><td>Target: Final Storage</td></tr>
<tr><td>Usable for Both Directions: yes</td><td>Chains Count: 1</td></tr>
<tr><td colspan="2">Chain 1 Steps:
1. Translation with Crane.
2. Stowage into Temporary Storage.
3. Translation with Reach Stacker.</td></tr>
</table>

Tab. 1. Simplified translation plan for situation depicted on Fig. 1.

In our translation plans, we have to define the start and target positions of the translation operation, and one or more ways (chains) how to transfer the container from the start to the target. Each chain consists of one or more steps. Additionally, we define if the plan can be used in a reversed order, i. e. from target to start. The benefits of our approach are:

- The translation process is defined only once for each feasible start-target connection.
- The description using flowcharts can be problematic, when they are assigned to vehicles or vessels with many freight kinds (in sense of dimensions, weights, etc.). The flowchart would in this case contain many branches for particular kinds of freight. In contrast, the translation plans are defined in a linear way, and the most suitable execution plan is chosen for each container automatically by the logic of simulation tool.
- The activities in flowcharts are considered to be executed sequentially and they are not started before all preceding activities are done. In our approach, the information about the entire translation process is known in advance and can be used to cooperate and optimize the whole translation process.
- Our implemented simulation module can create simple translation plans automatically, and the user does not have to define them at all (suitable for simple models).

In next paragraphs, we will define the individual parts of the translation plan. The start position of a plan is a place, from which the container will be retrieved. The target position is a place, where the container will be transferred. We determined two possible types of start and target position: storage, and track or road. The start and target position can contain one object or set of objects of the same type. All possible combinations of start and target positions are listed with a brief description in Tab. 2.

Start	Target	Description
Storage	Track / Road	Vehicle loading.
Track / Road	Storage	Vehicle unloading.
Storage	Storage	Displacement of hindering container inside the storage. The start and the target storage are typically the same.
Track / Road	Track / Road	Direct translation between two vehicles.

Tab. 2. Combinations of start and target positions.

Each translation plan contains also at least one translation chain. The translation chain defines the sequence of steps, which have to be done in order to transfer the container from the start to the target position. We defined four step types, listed in Tab. 3. The user defines a set of objects for the first three step types. The simulation logic then automatically chooses the most suitable objects from these sets (for each container individually).

Step type	Description
Transfer with Handling Equipment	Transfer with handling equipment, such as gantry crane or reach stacker. User defines a set of usable equipment.
Transport with Automated Guided Vehicle (AGV)	Transport with AGV. User defines a set of AGVs.
Stowage in Temporary Storage	Stowage of container into temporary storage. This storage serves only as an intermediate destination and all stored containers must be transferred away as soon as possible. It is typically used if straight transfer from start to target is not technically feasible. User defines a set of storages.
Own Power Translation	The container is unloaded or loaded without help of any terminal equipment.

Tab. 3. Defined step types in translation chain.

The simplest chain consists of a single *transport with handling equipment* or *own power translation* step. The other two types of steps must be accompanied by the *transport with handling equipment* steps, because the container itself cannot jump between storages and AGVs. Some examples of translation chains will be given in the third section of this paper.

The user of simulation model can define multiple chains for a single plan. In this case, optional conditions have to be assigned to the individual chains. Then, the first chain, which conditions are met, is chosen. The conditions can comprise time of the day, properties of transferred container and so forth (example: different chains can be used for 40- and 20-feet containers). If none constraints are specified, then the first defined chain is always chosen, the others in the same plan are ignored.

If the translation plan is used in reversed order (i.e. from target to start), the steps in particular chain are automatically executed in reversed sequence.

3. Experiments

The proposed approach has been incorporated into existing simulation tool Villon [4, 5]. This tool allows the user to create detailed simulation models of rail and road transport as well as simulation of rail-road container terminals. Villon implements a storage operations logic, which assigns the proper storage place or container for each incoming translation request. The translation plans has been used to describe the feasible translation operations and course of these operations.

The translation plans have been tested on a model of container terminal, depicted on Fig. 3. The test terminal handles rail and road transportation. Trains are handled on four parallel tracks by a crane. Roads are shown as continuous lines and possible movement direction is indicated with arrows. There are three available storages: full containers (served by crane), temporary storage (crane and two reach stackers) and empty containers (two reach stackers). The temporary storage serves as an intermediate station for empty containers, target or source location of which is at the storage for full containers. Two examples of plans are listed in Tab. 4 (train translation – empty containers) and Tab. 5 (direct translation of any container type between train and trucks). Because of simplicity of the model, all needed plans can be build automatically by Villon during simulation.

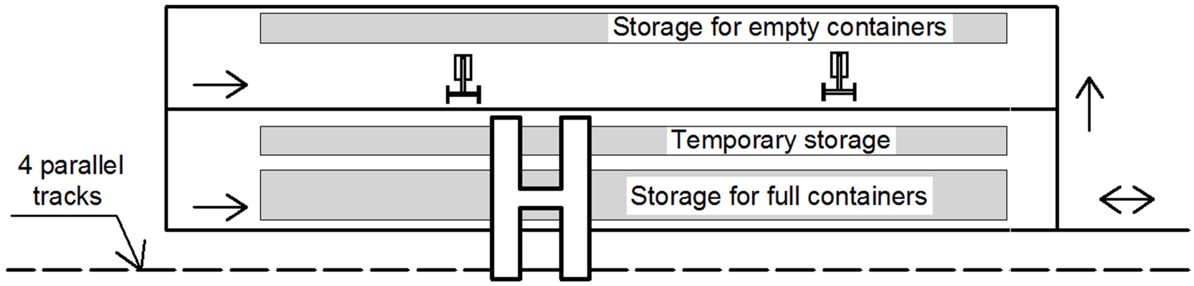


Fig. 3. Test model of container terminal.

Start: Set of Terminal tracks	*Target:* Storage for empty containers
Usable for Both Directions: yes	*Chains Count:* 1
Chain 1 Steps: 1. Translation with Crane – set: Crane 2. Stowage into Temporary Storage – set: Temporary storage 3. Translation with Reach Stacker – set: Reach stacker 1, Reach stacker 2	

Tab. 4. Translation plan for unloading and loading of train (empty containers).

Start: Set of Terminal tracks	*Target:* Set of Roads under crane
Usable for Both Directions: yes	*Chains Count:* 1
Chain 1 Steps: 1. Translation with Crane – set: Crane	

Tab. 5. Translation plan for direct translation (from train to track and vice versa).

4. Conclusion

In this paper we have introduced a concept of translation plans. These plans are used to describe the process of containers' translation in a container terminal simulation. We consider them to be an alternative (or at least support) tool for flowcharts. If the translation operations are quite simple, then the flowcharts are the right choice for translation process definition. However, the presented plans can replace them completely and the full power of plans reveals in more complicated and sophisticated simulations of container terminals. The suggested approach has been tested in simulation tool Villon and a test model has been presented. The future research will aim to the extension of the plans for simulations of warehouses and manufacturing systems. Furthermore, the automatic plans generation will be refined.

References

[1] STEENKEN, D., VOSS, S., STAHLBOCK, R. Container terminal operation and operations research – a classification and literature review, in *Container Terminals and Automated Transport Systems*, GÜNTHER, H. O., KIM, K. H. (Eds.), Berlin, Heidelberg: Springer-Verlag, 2005, pp. 3-50.

[2] KELTON, W. D., SADOWSKI, R. P., SWEETS, N. B. *Simulation with Arena*, 5th ed. Boston: McGraw-Hill, 2010.

[3] ADAMKO, N., KLIMA, V. Definition of operating procedures in generic agent based simulation model of transportation logistic terminals, in *Proceedings of the 2008 European Simulation and Modelling Conference*, October 27-29, Le Havre, France, 2008, pp. 312-316.

[4] KLIMA, V., KAVIČKA, A. Agent-based simulation model design, in *Proceedings of European simulation multiconference*, SCS, 1996, pp. 254-258.

[5] ADAMKO, N., KLIMA V., KAVIČKA, A. Villon – agent based generic simulation model of transportation logistic terminals, in *Proceedings of ESM 2007*, Eurosis, St. Julian's, Malta, 2007, pp. 364-368.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Body Area Network for Monitoring Human Vital Signs Using Smartphone

[*]Michal Kochláň, [*]Michal Hodoň, [*]Jana Púchyová
[*]University of Žilina, Faculty of Management Science and Informatics, Department of Technical Cybernetics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic,
{michal.kochlan, michal.hodon, jana.puchyova}@fri.uniza.sk

Abstract. Healthcare is changing. Modern ways of patient monitoring represent key issue for disease management. One of the interesting modern ways is wireless body area network technology. Wide range of sensors can give a complex picture of health that a smart application or a doctor can use for diagnose specification. This document presents a body area network for human vital signs monitoring with smartphone network coordinator. The growing popularity of smartphones opens the space for new nontraditional solutions based on them. The capabilities of relatively high computational power, long battery life and mobile operator subscription (making calls, sending short messages, etc.) make the smartphones an ideal choice for the network coordinator. The proposed sensor solutions and timing issue is designed with the respect to low power consumption and long network life.

Keywords: body area network, node timing, smartphone, health monitoring, vital signs monitoring.

1. Introduction

Healthcare is in need of change and in fact, the changes have already started. The population ages and the number of chronic and heart diseases increases [1]. Early diagnosis, in most cases, leads to successful treatment. Therefore, new ways of continuous and everyday patient monitoring can contribute to the early diagnosis designation. Wearable components around a human body that scan and sample various human vital signs, e.g. electric characteristics of the heart, blood pressure, brain activity, body temperature, etc. are *body sensor nodes*. These sensor nodes can be wired to the coordinator node as well as be wireless. Wireless technology is more convenient but pays for that with higher power requirements. Though, low power sensor node consumption with efficient communication mechanism extends the wireless network life [2].

1.1. Body Sensor Network

The body sensor nodes around a human body that communicate in a coordinated fashion create a Body Area Network (BAN). In most cases, a central and most powerful node acts as the coordinator of the network. In terms of health monitoring, there are body surface and pervasive body sensor nodes [3]. This paper assumes human body surface mount sensor components. These components can easily monitor heart rate or body temperature. Very important network component is a network coordinator. A smartphone has been chosen to be a network coordinator in the proposed application and is responsible for data collection, evaluation and further action decision (to call an emergency or to send a short message).

1.2. Network coordinator – A smartphone

Network coordinator node performs time consuming tasks. The reason for that is that this node is usually the most powerful and at the same time this node processes data sampled from peripheral sensor nodes that are being controlled by the coordinator node. At the same time, the network coordinator manages and initiates the inter-node communication. Usually powerful hardware components such as FPGAs are used. FPGA feature are described in [4] and [5]. But the proposed solution in this paper uses a smartphone. The benefit of this approach is that the smartphone already contains powerful microcontroller and these multi-core processing units in smartphones does not surprise. Besides the smartphone's application framework already handles the basic networking tasks, mobile operator tasks and commands touch screen for application interaction.

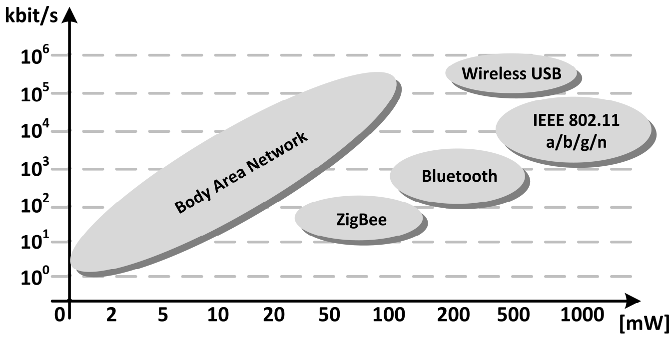


Fig. 1. Transfer rate comparison of different wireless technologies in power scale.

Due to health reasons the lowest communication power devices should be used. Fig. 1 illustrates the data transfer speed comparison of different wireless technologies in power scale. Except for the BAN, all the wireless technologies power starts at 50mW [6]. This fact disqualifies mentioned technologies from being effectively used as wireless technologies for health monitoring.

The research still faces key issues that must be solved – battery life improvement, power usability, size of the sensors, human body energy harvesting (mainly from body movements) [7].

2. Case Study of the Timing Scheme

The proposed solution for human vital functions monitoring consists of a central unit *C*, nodes N_i and a smartphone. Connection schematic is depicted in Fig. 2.

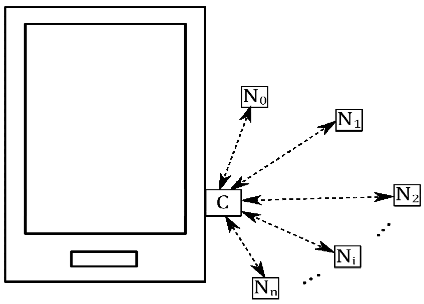


Fig. 2. Connection schematic of system for vital function monitoring.

The two-way wireless communication occurs only between *C* and N_i. *C* communicates only with one node at a time and N_i*s* don't communicate among each other. The main principle is that the central unit is awake all the time and listens for incoming data from N_i*s*. Each node is given own communication slot – time division multiple access (TDMA) is used. By default, the nodes are asleep. They are awake only in case of (i) Measurement procedure, (ii) Measurement values processing and (iii) During communication slot period (Fig. 3). Communication in $N_i \rightarrow C$ direction is established every time the new time slot starts. Communication in the opposite direction, $C \rightarrow N_i$ is established only in case a problem occurs or a synchronization request appears. For this purpose, a simple algorithm called *Cristian's algorithm* [8] has been chosen:

$$T_{syn} = T_s + \frac{T_c}{2}, \qquad (1)$$

where T_{syn} is (new) synchronized time, T_s is synchronizing time sent from central unit and T_c is communication time. The time synchronization is performed by each node separately after few time slots. Node transmits synchronization requests during its own time slot with the data measurements. In the same time slot the node receives the synchronizing time (T_s) and computes the synchronized time (T_{syn}) according to formula (2). Obviously it is expected, that both ways of communication last the same time. Each node is able to compute the start of its time slot from synchronized time:

$$S_i(k+1) = S_i(k) + (T_{syn} \bmod t_a) \quad ; \quad S_i(0) = si + T_{start} \tag{2}$$

where $S_i(0)$ is the start time of the first time slot of each node N_i, $i = \{1,2,3,\ldots,n\}$ (n is the number of all nodes), s is time slot size, T_{start} is the time of starting communication process, k is the slot sequence, T_{syn} is synchronized time and t_a is defined as:

$$t_a = sn + t_{cp} \quad ; \quad s = t_s + t_g \tag{3}$$

where t_{cp} is data block size being processed by central unit which implies t_a is the block size of the repeated transmission of the measured data of the same node. Size of time slot s is consists of communication slot time t_s and the gap time t_g. The gap time compensates the communication slot interlacing. The measurement itself and measurement handling is executed right before the communication slot in order to keep the transmitted data up-to-date and new. The node and central unit time management is presented by Fig. 3.

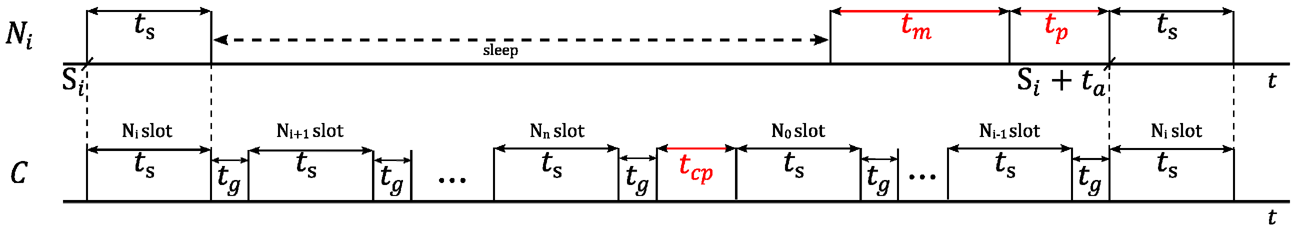


Fig. 3. Time diagram of node and central unit behavior.

Measured values are further processed in central unit; central unit processing slot is $\leq t_a$. The communication time slot t_s is wide enough to transmit data between node and central unit three times. The worst case happens when the node is unsynchronized and did/could not communicate in the previous time slot (Fig. 4, c). In the best case one-way communication is performed (Fig. 4, a). When the synchronization is needed, the two-way communication scheme is used (Fig. 4, b).

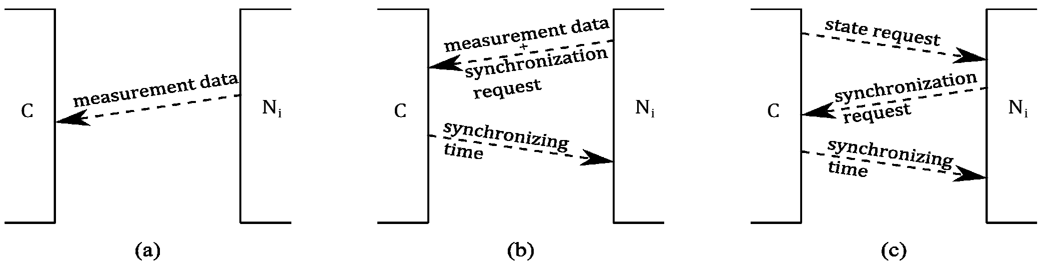


Fig. 4. Communication process, a) only measurement data transmission, b) measurement data transmission and request for synchronization, d) state request from non-synchronized node.

If the response does not come after the status request, the node N_i is considered to be damaged. In general, it is assumed that the nodes are homogeneous.

3. Communication and Sensor Survey

The IEEE 802.15 Task Group 6 defines a Medium Access Control (MAC) layer and several supporting physical (PHY) layers to enable Body Area Networks (BAN) used in, on, or around a body. IEEE 802.15.6 specifies a total of three PHYs, namely *Narrow Band* (NB), *Ultra Wide Band* (UWB) and *Human Body Communication* (HBC). All three PHYs are specified to address different system demands and target applications (Fig. 5).

Fig. 5. IEEE 802.15.6 base architecture [9].

The standard for IEEE 802.15.6 specifies the following operation frequency bands: 402-405 MHz (MICS), 420-450 MHz (Japan WMTS), 863-868 MHz (EU ISM), 902-928 MHz (US ISM), 950-958 MHz (Japan ISM), 2360-2400 MHz (MBANS), and 2400-2483.5 MHz (Worldwide ISM).

The mentioned specification is optimized to serve wireless communications needs for ultra-low power devices operating in or around the human body. Such communication system has to respect some specific restrictions related to the target application area (i) *Specific Absorption Rate* (SAR) – max. 1.6 W/kg per one gram of the tissue and 0.08 W/kg average values for the whole body; (ii) *Broadcast bands restrictions* – max. 25 μW radiated power in the frequency band; (iii) *Energy consumption restrictions* – tiny batteries with the maximum operating time or energy harvesting systems; (iv) *Transfer rate requirements* – 10 Kb/s ~ 10 Mb/s depending on the application.

4. Conclusion

The typical health-care BAN applications include patient’s physiological and vital functions permanent monitoring allowing detection of various diseases (high blood pressure, tachycardia, depressions, etc.); disabled patients’ monitoring and guidance indoors as well as outdoors; remote first aid through WSN actuators (deep brain stimulator, muscular stimulator, insulin pump, etc.). The proposed BAN survey in timing issues, sensor and coordinator selection serves as the basis for the upcoming work on the project. Recent project development involves the coordinator setup and sensor node manufacturing. Further development steps comprise suitable control unit development and tuning of the timing issue and communication protocol with respect to low power consumption.

Acknowledgement

This contribution/publication is the result of the project implementation “*Vývoj špeciálnej BAN pre monitoring pacienta v domácich podmienkach za využitia smartfónu ako riadiacej jednotky i koordinátora siete*” with grant no. *050/12* in grant program “*Vysokoškoláci a technika*” supported by Foundation Volkswagen Slovakia funded by Volkswagen Slovakia, a.s.

Nadácia Volkswagen Slovakia

References

[1] EUROSTAT, “*Eurostat regional yearbook 2012*”, Chapter Health, ISBN 978-92-79-24940-2, 2012.

[2] IINO, Y., HATANAKA, T., FUJITA, M., “*Event-predictive Control for Energy Saving of Wireless Networked Control System*”, American Control Conference, June 10-12, 2009.

[3] BRANDÃO, P., “*Abstracting information on body area networks*”, University of Cambridge Technical Report No. 812, ISSN 1476-2986, 2012.

[4] ŠEVČÍK, P., KOVÁŘ, O., “*Very efficient exploitation of FPGA block RAM memories in the complex digital system design*”, Journal of information, control and management systems, ISSN 1336-1716, Vol. 8, 2010, pp. 403-414.

[5] ŠEVČÍK, P., “*Implementation of the data preparing unit for the convolution computing in the FPGA device*” TRANSCOM 2007, Žilina, June 25-27, 2007, Slovakia, ISBN 978-80-8070-693-7, pp. 225-228.

[6] CHEN, M. *et al.*, “*Body Area Networks (BAN): A Survey*”, Mobile Networks and Applications, 2011 February 16, pp. 171-193.

[7] MONTÓN, E. *et al.*, “*Body area network for wireless patient monitoring*”, Telemedicine and E-health Communication Systems, Vol. 2, February 2008, pp. 215-222.

[8] CRISTIAN, F. *Probabilistic clock synchronization*. Distributed Computing. Springer. vol. 3, 1989, pp. 146-158.

[9] Testing BAN (WBAN) Networks and IEEE 802.15.6, [Online] Available: http://www.ni.com/white-paper/14285/en

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Building Decision Diagram from Espresso benchmarks

*Jozef Kostolny
*University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitna 8215/1, 010 26 Žilina, Slovak Republic, jozef.kostolny@fri.uniza.sk

Abstract. It is necessary to have sample inputs to verify the correct functioning of new algorithms. This paper presents a possible solution for creation of a Binary Decision Diagram (BDD) from Espresso benchmarks. Building a BDD benchmarks is possible through the synthesis of PLA benchmarks into simplified form of BDD, and the following unpacking of BDD.

Keywords: Binary decision diagrams, dot file format, parse, conversion, pla file format.

1. Introduction

Reliability is an essential characteristic of every technical system. Reliability analysis is used to estimate the characteristics of different types of systems. Two principal mathematical models are used in Reliability engineering - Binary-state systems (BSS) and Multi-state systems (MSS). Some aspects of these mathematical models have been considered in the paper [1]. According to [1] every of these mathematical models (BSS and MSS) has different types of mathematical forms for representation. Description of the differences between the BSS and the MSS representation is in the article [1].

A Multiple-Valued decision diagram (MDD) can be used for presentation of a MSS. An efficient form for definition of a BSS is a Binary Decision Diagram (BDD). BDDs are very often used to manipulate with Boolean expressions. This is caused by the fact that a BDD takes less memory for a large Boolean expression than other forms of a BSS. A generalization of a BDD is a MDD. A MDD is a directed acyclic graph that consists of sink nodes, which represent system states, and non-sink nodes that correspondent to system components. Each non-sink node (Fig.1) has outgoing edges connected to other nodes. A path from the top node to a sink node defines system state. A value of a sink node defines a system state – if the value is zero then system is failure, and if there are other values then system is working.

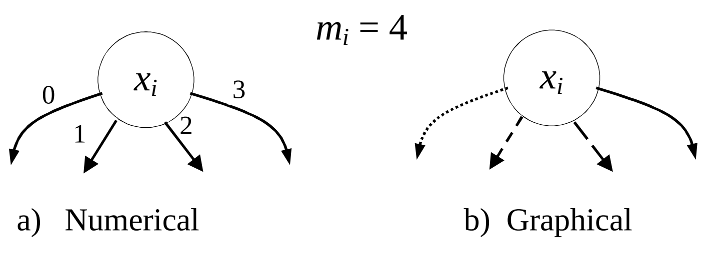


Fig. 1. Example of MDD node

In the article [3], there has been considered an algorithm for estimation of a system state by Direct Partial Logic Derivatives (DPLD), which are calculated using MDD. In this paper, a solution for building a BDD is presented. This solution can also be used for building a MDD and for testing the algorithm.

2. Benchmarks

Reliability analysis of MSSs presented as MDDs is considered in article [2] and [3]. In these articles, new algorithms for calculation measures of MSS have been proposed. The verification and testing of new algorithms are principal parts of algorithms implementation. One of methods of such verification is the algorithm testing based on the benchmarks. The benchmarks for the algorithm considered in [3] are not available in the required form and therefore they should be modified. In this article, the benchmarks LGSynth91 [4] are presented and used. These benchmarks are in a binary PLA BENCH (Espresso) format. This format is used to define the two-level logic circuits. The PLA file consists of cubes with two parts. A cube is considered a representation of a row in a PLA file. The first part of a cube is an assignment of inputs and the second part is an output (outputs). The PLA file defines a structure and values of switching function expressly [5].

a) PLA file

```
.i 5
.o 1
.ilb d c b a e
.ob xor5
.p 16
11111 1
01110 1
10110 1
00111 1
11010 1
01011 1
10011 1
00010 1
11100 1
01101 1
10101 1
00100 1
11001 1
01000 1
10000 1
00001 1
.e
```

b) PLA schematic

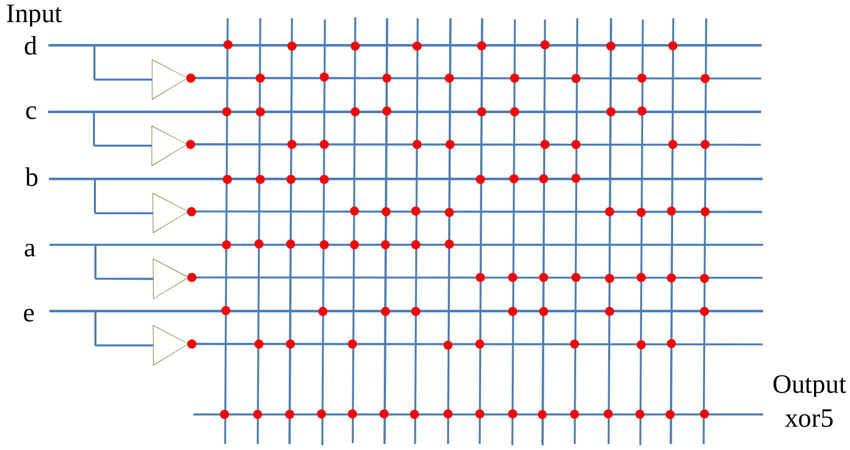


Fig. 2. Benchmarks xor5 specified in PLA format (a) and the PLA schematic (b)

2.1. The Modification of the PLA File Format

The PLA file format is used in tools working with logical circuits. One of these tools is tool ABC (A System for Sequential Synthesis and Verification [6]), which is developed by Berkeley Verification and Synthesis Research Center. This tool can be used for the modification of PLA benchmarks files. The ABC tool can load a PLA file and make an output in a form of a BDD. This tool has built-in support for the construction of a BDD by CUDD package [7], which is already part of this tool. Therefore, this tool has been used for the synthesis of PLA benchmark into a BDD. The output of this tool is a file in a DOT format, which contains the resulting BDD benchmark.

2.2. The DOT file format

The DOT (graph description language) file format is used to write graphs and digraphs in simple text. This format can be displayed by a graphics library and tool Graphviz [8] that has defined this format. From the last update, this tool works with files that have ".gv" suffix, which is actually same as a ".dot". Changing of the file suffix was performed for one reason: Microsoft has registered this suffix for MS Word. DOT file notation is simple. It consists of a header "graph" or "digraph" according to the type of an output. Curly brackets "{}" separate all major parts. The next part describes levels of the diagram, which consist of nodes. Every node has name and/or value.

The last part of the DOT file contains a description of edges between defined nodes. The description of edge contains name and type of line. Example file entry is shown on Fig. 3 (a).

a) DOT file

```
digraph "DD" {
        size = "7.5,10"
        centr = true;
        edge [dir = none];
        {  node [shape = plaintext];
           edge [style = invis];
       "CONST NODES" [style = invis];
       " d " -> " c " -> " b " -> " a " -> " e " -> "CONST NODES"; }
        {  rank = same; node [shape = box]; edge [style = invis];
           "  xor5  "; }
        { rank = same; " d ";
          "370"; }
        { rank = same; " c ";
          "36b"; }
        { rank = same; " b ";
          "35e"; }
        { rank = same; " a ";
          "341"; }
        { rank = same; " e ";
          "2f6";}
        { rank = same; "CONST NODES";
          { node [shape = box]; "2ee";}
        }
        "  xor5  " -> "370" [style = solid];
        "370" -> "36b";
        "370" -> "36b" [style = dotted];
        "36b" -> "35e";
        "36b" -> "35e" [style = dotted];
        "35e" -> "341";
        "35e" -> "341" [style = dotted];
        "341" -> "2f6";
        "341" -> "2f6" [style = dotted];
        "2f6" -> "2ee";
        "2f6" -> "2ee" [style = dotted];
        "2ee" [label = "1"];
}
```

b) DOT visualization

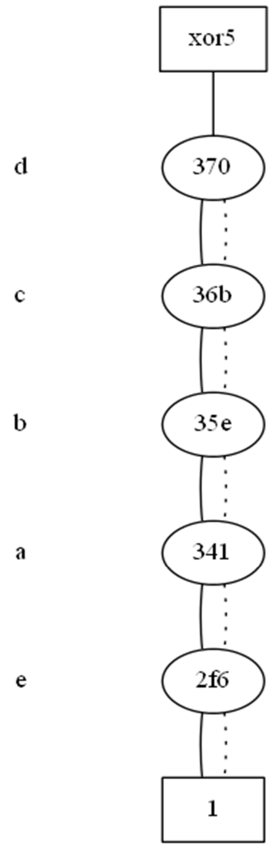


Fig. 3. Example from DOT file for xor5 benchmark (a) and Graphviz visualization xor5 benchmarks as BDD (b)

3. Benchmark processing

Back to benchmark processing. ABC tool uses CUDD library for processing of BDD. CUDD library works with a simplified form of BDD. Example is on Fig. 3 (b). This BDD notation is used to reduce the size of BDD in memory. Therefore, the output node is one and it has a value 1. Edges “then” (or 0) are indicated by solid lines, and the “else” (or 1) have a dual notation – dotted or dashed lines. An edge with the dotted line is also called “complemented edge”. Important of complemented “else” is that, if a path from input node to output node contains an odd number of dotted edges, then the result value of output node is 0. If there is an even number of visited dotted edges then the value is 1 [9].

3.1. Conversion algorithm

When we want to use the simplified form of BDD in our BDD tool for finding DPLD, we have to convert this BDD to full size. Conversion algorithm to full size BDD (Fig. 4) is following.

The conversion consists of traversing all possible ways in BDD. With each transition of edge, new node is added and connected on this edge. Final node (leaf) value is determined by the number of transitions of complemented edges, as has been already described above. If we traverse every path, then we have constructed a full size BDD.

This format of BDD benchmark can be used for testing of algorithms for calculation the DPLD of BDD. The other option is the possibility of using this BDD for generating the MDD benchmarks that will be used for verification algorithms on Multi-state systems.

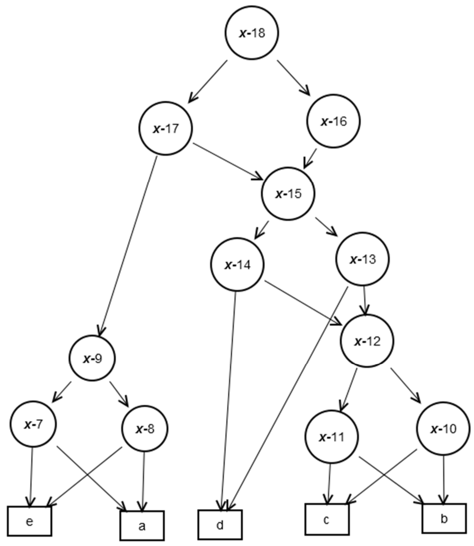


Fig. 3. UML .FRI visualization of xor5 benchmark in full size BDD

4. Conclusion

In this paper, a solution for creation of a Binary Decision Diagram for testing of algorithms in reliability analysis is presented. Espresso benchmarks are used for testing. These benchmarks are in PLA file format. We can build BDD in simplified form from benchmarks. As we present in this paper, the simplified BDD can be transformed into the full form BDD quite simple. In the further process of research, it is possible to use the BDD benchmarks for conversion into Multi-valued Decision diagram.

References

[1] E. Zaitseva, J. Kostolny, “Multi-state system for reliability analysis in application problem,” in *Digital technologies 2011 - 8th international workshop*, 2011, pp. 8 – 15.

[2] E. Zaitseva, V. Levashenko, and J. Kostolny, “Multi-State System Importance Analysis Based on Direct Partial Logic Derivative,” in *International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2012*, 2012, pp. 1514–1519.

[3] E. Zaitseva, V. Levashenko, J. Kostolny, and M. Kvassay, “A Multi-Valued Decision Diagram for Estimation of Multi-State System,” in *EUROCON 2013 - International conference on computer as a tool*, 2013 - in press.

[4] The Benchmark Archives at CBL online: http://www.cbl.ncsu.edu:16080/benchmarks/LGSynth91 /twolexamples. – 13.03.2013.

[5] Input file format for Espresso PLA file format at The University of Iowa online: http://www.engineering.uiowa.edu /~switchin/OldSwitching/espresso.5.html - 13.03.2013.

[6] ABC: A System for Sequential Synthesis and Verification at Berkley University online: http://www.eecs.berkeley.edu/~alanmi/abc/ - 13.03.2013.

[7] CUDD package at University Colorado online: http://vlsi.colorado.edu/~fabio/CUDD/cuddIntro.html - 13.03.2013.

[8] Graph Visualization Software at Graphviz online: http://www.graphviz.org/content/dot-language - 13.03.2013

[9] E. L. Schreiber, “A CUDD Tutorial,” 2008.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

RPL Routing Protocol for 6LoWPAN Wireless Sensor Networks

*Jozef Kozak, Martin Vaculik
*University of Zilina, Electrotechnical, Department of Telecommunications and Multimedia, Univerzitna 1, 01026 Žilina, Slovakia, {jozef.kozak, martin.vaculik}@fel.uniza.sk

Abstract. The Internet of Things (IoT) foresees billions of objects to connect the physical world to the digital world. Network stacks and routing protocols that can scale with network size and density are needed while remaining energy-efficient and lightweight. The Routing Protocol for Low-Power and Lossy Networks (RPL) is a routing protocol for constrained nodes (with limited processing power, memory and energy) and networks (links with high loss rates, low data rates and instability). It was developed by IETF ROLL WG for 6LoWPAN (IPv6 over Low power Wireless Personal Area Networks) to support various application areas of the Internet of Things. This paper makes a survey of RPL protocol fundamental features. We intend to use the COOJA simulator to evaluate performance of implemented Contiki RPL according to set of parameters of the stack and optimize system performance for specific application.

Keywords: RPL, WSN, IPv6 routing, Internet of Things

1. Introduction

IoT could be explained as a new application option of the ordinary Internet, where billions of tiny devices interacting with physical environment are connected. In the last few years, there has been an increasing trend towards enabling technologies of this concept [3]. Integral parts of IoT are wide range of IP-enabled smart objects, sensors and actuator networks.

Wireless Sensor Networks (WSNs) used to communicate via their own tools and technologies, often with proprietary protocols for isolated applications. With idea of IoT there are efforts to reuse Internet technologies to integrate WSNs such as IEEE 802.15.4 into existing Internet infrastructure. The essential component in the area of WSNs is the 6LoWPAN, adaptation layer that allows efficient use of IP protocol in constrained networks.

Design of routing protocols in LoWPANs with several constraints constitutes great challenges. Successful solutions should take into account the specific application requirements, along with IPv6 behaviour and 6LoWPAN mechanisms. While WSNs and ad-hoc networks are both wireless multi-hop networks, they are different in three important factors: (a) primary goal of WSN is energy efficiency, (b) the amount of data transported by a WSN is typically low in most proposed applications and (c) substantial portion of information flows are destined towards limited number of nodes in WSN. Routing protocols designed for ad-hoc networks are hence inadequate in large and dense WSNs [1].

There have been many efforts for designing suitable routing protocols for LoWPANs, but the most consolidated and widely recognized proposal is RPL, a new IPv6 routing protocol standardized by the IETF ROLL working group. It is a gradient based routing algorithm that organizes the network topology as a Destination Oriented Directed Acyclic Graph (DODAG), where each node selects the next hop as preferred parent according to largest gradient to the destination (based on some metrics). This routing structure is optimized for traffic to or from one or more root nodes [5]. RPL operates at the IP layer and thus allows for routing across multiple link layers, in contrast with other form of “routing” operating at lower layers.

The rest of this paper is organized as follows. In section 2, a brief overview of RPL routing principles is sketched. Section 3 describes our intended simulation for RPL performance evaluation. Section 4 concludes the paper.

2. RPL Routing Overview

RPL is a distance vector IPv6 routing protocol for LoWPANs that specifies how to build a DODAG using Objective Function (OF) and set of metrics/constraints. The OF defines how to compute the best path. There could be several OFs in operation on the same node and mesh network to carry traffic with different requirements for path quality. A node can join one or more graphs (RPL instances) according to graph characteristic to support QoS aware and constraint based routing [4].

2.1. Routing Topology Construction

The DODAG construction is based on the Neighbour Discovery (ND) process. To establish a DODAG, a root node informs about its presence by periodically sending messages called DODAG Information Object (DIO). They are sent as options of IPv6 Router Advertisements transmitted using multicast frames. DIO messages contain common configuration attributes in RPL instance (rank of the sending node, DODAG-ID, OF, etc.) [5].

Once a node receives a DIO, it calculates its rank using OF from the rank of sending node and cost of reaching node form itself. When a node want to join the DODAG, adds the DIO sender to its parent list and forwards the DIO message with the updated rank information. This way DIOs are propagated to distant nodes from the root to create a DODAG. The rank represents the location of a node within the DODAG and it strictly decreases in the upstream direction. Any node that has lower rank than the node itself is considered as a candidate parent. When a node joins the network, it may wait to receive a DIO or it may multicast a DODAG Information Solicitation message (DIO) to probe neighbour nodes. When a node already associated with DODAG receives another DIO message, it can discard it according to some criteria, process the message to either maintain its location in the DODAG or improve its location by getting lower rank. When a node changes its rank, it must discard all nodes with the higher rank in the parent list to avoid loops. After the construction of the DODAG, each client node would have a default upstream route to the DODAG root. Construction of the DODAG is on the Fig. 1.

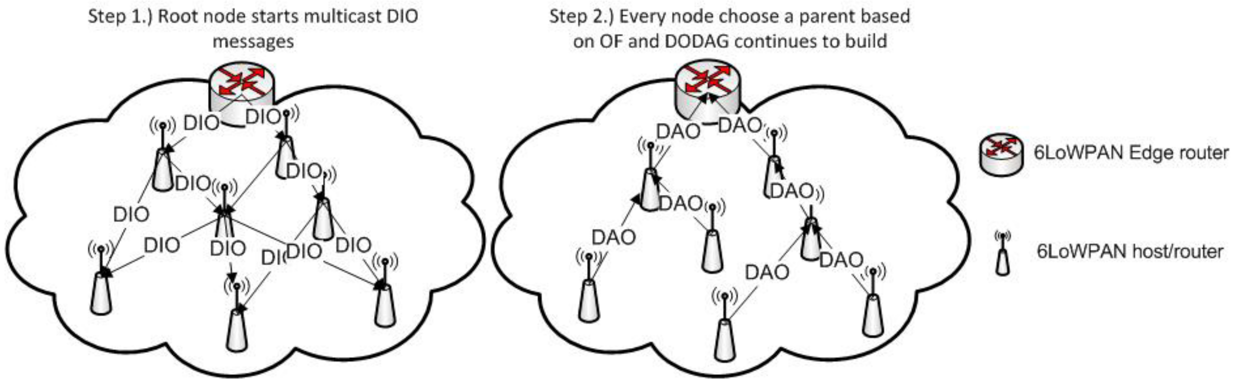


Fig. 1 DODAG construction process

To propagate reverse route information Destination Advertisement Object messages (DAO) are used. They are sent by each node to populate routing tables with prefixes of their children and to advertise their addresses to their parents. As a response by a DAO recipient DAO Acknowledgement (DAO-ACK) is sent [3]. The way DAO messages are transmitted and processed by intermediate routers depends on the mode of operation. In the first mode, called non-storing mode (intermediate nodes do not store any routes), unicast DAOs with included parent set are sent to the root. DODAG root then can construct a source route by recursively looking at the DAO parent information. In the second mode called storing mode (intermediate nodes have available memory to store routing tables), each RPL node sends unicast DAOs to its parents. Each node then store information about reverse paths to downward subDODAG and forwards it to its parents.

2.2. Objective Function

The OF defines how RPL nodes select and optimize routes within a RPL Instance. The OF is defined within the DIO message by an Objective Code Point (OCP). The OF defines how nodes translate one or more metrics and constraints into value called rank, which approximate the node's distance from a root. The OF also defines how nodes selects parents. Variety of LoWPAN applications with different requirements can be enabled using specific OFs. The design of efficient OFs is still an open research issue.

IETF proposed two objective functions:

Objective function 0 (OF0): uses only information from DIO header. A node rank is computed by adding a normalized scalar (*RankIncrease*) to the *Rank* of a preferred parent and therefore minimizes hop count. OF0 is designed as a default OF that will allow interoperation between different implementations. It ignores metric containers and it leaves to implementation the responsibility to compute how link properties are transformed into a *RankIncrease.*

Minimum Rank Objective Function with Hysteresis (MHROF): is designed to find paths with minimal cost while preventing excessive traffic in the network. A node switches to the *NewPathCost* only if it is smaller than *CurrentPathCost*- γ (where γ is parent switch threshold implementing hysteresis). This objective function may be used with any additive metric, typically sum of ETX (Expected Number of Transmission) or delay to the DODAG root. It employs a DODAG parent set with only one node, which is automatically chosen as the preferred parent [6].

2.3. Metrics and Constraints

A metric is a scalar quantity used as input for best path selection. A constraint, on the other hand, is used as an additional criterion to prune links or nodes that do not meet the set of requirements. Metrics and constraints can be node (node state, energy, etc.) or link (latency, reliability, link colour, etc.) based [4].

2.4. Trickle Algorithm

To control rate of RPL messages trickle algorithm is used. It implements a consistency check model to verify if RPL nodes have out-of-date information. If the routing information is up-to-date, trickle algorithm will exponentially reduce the rate of DIO messages. Otherwise it immediately schedules the transmission of a DIO message to update it. During the trickle time interval RPL node uniformly schedules DIO messages at a given time t. During this period, the RPL node counts all DIO messages that content routing information compliant with its own state of network. If the number of messages is lower than redundancy threshold (k), the scheduled message is sent. Otherwise, transmission is cancelled and the trickle timer is doubled (up to Imax) and a new DIO transmission is scheduled. RPL resets the trickle timer to a minimum value (Imin) if there is an inconsistency (such as out-of-date DIO message).

3. Contiki/COOJA Simulator

Widely used network simulation tools may lack credibility because of oversimplified models. On the other hand, real testbed scenarios are more precise, but more expensive and more difficult to deploy. In our future work we intend to use cross-level emulation tool known as COOJA. It is a Java based tool for simulation of sensor networks running Contiki OS (designed for constrained embedded systems) typically used in applications for the Internet of Things. Contiki code is open source and thus it is easy to control the network parameters in simulations. Contiki natively supports IP communication implementing 6LoWPAN adaptation layer and RPL routing. Its uIPv6 stack became the first low-power wireless operating system to provide a certified IPv6 Ready stack.

ContikiRPL includes all basic mechanisms specified by RPL standard, except non-storing mode (source routing) and security mechanisms. It adopts both OFs: OF1 based on ETX and the default OF0, which minimizes the average number of hops from the root. It relies on the IPv6 ND for address resolution and verifying neighbour reachability. There is set of radio duty cycling mechanisms such as ContikiMAC, X-MAX and LLP available.

We would like to evaluate performance of Contiki RPL implementation based on specific requirements of selected applications. To better characterize 6LoWPAN stack and to optimize its performance, experiments will be drawn and discussed in future works. We will investigate impact of routing parameters (like OF, selected metrics, timers, etc.), network size, link quality, etc. on overall network quality parameters.

Acknowledgement

This work has been supported by

Centre of excellence for systems and services of intelligent transport II., ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

"Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ"

4. Conclusion

For communication between objects and services in the IoT many standard Internet protocols are used. In our view, IoT must rely on standard solutions for extending the current Internet paradigms into the world of embedded sensing and actuating objects to maintain interoperability between different applications.

In this paper we discussed RPL routing protocol for 6LoWPAN and its characteristics. It represents suitable routing protocol for low-power IP enabled objects. Our study shows possibility for RPL routing protocol simulation in Contiki/COOJA simulator.

References

[1] WINTER, T., THUBERT, P., *RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks.* IETF: RFC 6550, 2012

[2] TRIPATHI, J., OLIVEIRA, J.C., VASSEUR, P., *A performance evaluation study of RPL: Routing protocol for low power and lossy networks*. CISS 44th, Princeton 2010

[3] GADDOUR, O., KOUBAA, A., CHAUDHRY, S., *Simulation and performance evaluation of DAG construction with RPL*. ComNet 3rd, Hammamet 2012

[4] VASSEUR, J.P., AGARVAL, N., *RPL: The IP routing protocol designed for low power and lossy networks*. IPSO Alliance whitepaper, 2011

[5] BRUNO, R., CONTI, M., *RPL routing protocol in advanced metering infrastructures: An analysis of the unreliability problems*. SustainIT conference, Pisa, 2012

[6] PALATTELA, M. R., ACCETTURA, N., *Standardized protocol stack for the Internet of (important) Things*. Communications Surveys & Tutorials, Issue: 99, IEEE, 2012

[7] BORMANN, C., SHELBY, Z., *6LoWPAN: The wireless embedded Internet*. ISBN: 978-0-470-74799-5 ,WILEY, 2009

[8] ACCENTURA, N., GRIECO, L.A., *Performance analysis of the RPL routing protocol*. International Conference on Mechatronics, Istanbul, 2011

[9] ACCENTURA, N., CAMACHO, C.A., *Performance assessment and tuning rules for low-power and lossy stacks*. WFCS 9th, Lemgo, 2012

[10] WATTEYNE, T., MOLINARO, A., *From MANET To IETF ROLL Standardization: A Paradigm Shift in WSN Routing Protocols*. Communications Surveys & Tutorials, IEEE, Vol. 13, Berkeley, 2011

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Identification of the Language in the Task of the Automatic Recognition of Reproduced Fragments of the Text Documents

*Yury Krapivin
*Brest State Technical University, Faculty of Electronic Information Systems, Department of Intelligent Information Technologies, Moskovskaja 267, 224017 Brest, Belarus, {ybox}@list.ru

Abstract. The analysis of the main methods to solve the problem of the automatic identification of the language of the text document has been done. The effective algorithm ensuring the solution of the stated problem has been defined. It unites possibilities of minimal statistical and linguistic analysis of lingual data.

Keywords: Text processing, Automatic language identification, Linguistic analysis.

1. Introduction

The information systems that work with huge volumes of text documents of unconditioned knowledge domain and successfully solve different applied tasks are in demand by businesses, organizations and individual users. The growing amount of information in different languages that is stored both in full-text databases and in the Internet requires the quality of its operational processing. It implies, inter alia, the solving of the fundamental task of automatic text processing – automatic language identification of the text document. The specified problem arises during the creation of the systems that operate in a multilingual information environment: the machine translation, the optical text recognition, the information retrieval and the automatic recognition of reproduced fragments of the text documents. The base functionality of the system of the automatic recognition of reproduced fragments of the text documents for the Russian and Belarusian languages, that firstly needs solving the problem of the identification of the language has been described in the work [1].

2. Methods of Identification of the Language of the Text Document

There are following basic As shown by our study, the most known methods used to solve this problem at present are based on the same principle, that can be described using the terminology of information retrieval in the following way. At first, a language search profile (LSP) is created for each natural language (NL) of a finite set of languages involved in the identification process, secondly, a document search profile (DSP) is created for each analysed (input) document according to the rules of the LSP creation, and, thirdly, the strategy of the comparison of the DSP and the LSP is chosen, used as the basis for making decisions of relating the input document to the specific languagemethods of the identification of the language of the text document, depending on the rules used to create search patterns and the strategies of their comparison: the short words [2, 3], the frequent words [4, 5], the N-grams [2, 6-8], the statistical [9-11], the string kernels [12], the alphabetical [13, 14], the function words [13], the non-grammatical words [13]. According to the results of our study, these methods are sufficiently effective only under the certain restrictions on the size of the input text, and comparative characteristics of the natural languages, on the set of which the problem is solved, and, thus, can't provide the creation of the effective algorithm separately, which is represented in this work.

3. The Algorithm of the Identification of the Language of the Text Document

Any text T of the language L can be considered as a finite chain $T = a_1 a_2 ... a_n$, $a_i \in A$, $i = \overline{1, n}$, A – alphabet of the language L, formed in accordance with the set of its lexico-grammatical, syntactical and semantic rules. It's obvious, that all methods, listed above, refer to the use of the knowledge of a natural language from the alphabetic level up to the grammatical one, which is quite acceptable from the standpoint of complexity of their use in the industrial systems of the automatic text processing. However, it's necessary to achieve greater efficiency (accuracy) of these methods, remaining at these levels in solving the problem, in our case, on the set of Belarusian, Russian and some European languages (English, French, German).

The trigram method, the alphabetic method, the function word method and the alphabetic-trigram method have been included in our experiments. The well-known linguistic text corpus Leipzig Corpora Collection (LCC) [15] has been used to investigate the first method as the training set for European languages. It is based on randomly selected sentences from newspaper articles and web documents. In regard to the Russian and Belarusian languages, the proper text corpus has been constructed according the principle of the LCC. The text sets of 100 thousand sentences have been used for each of the five languages. The list of function words (LSP) has been determined for the second method adequately and strictly linguistically. It includes, for example, for English: determinatives (few, last, many, only), prepositions (about, above), conjunctions (and, or, but), articles (a, an, the), modal and auxiliary verbs (can, could, has), pronouns (I, we, it, him, mine), determinants (as, more, very, enough), adverbs (below, near, now, there), particles (to, not) etc. It has been obtained via the automatic processing of the sizeable, for example, about 200 thousand word-forms of the English language and about 1 million 200 thousand word-forms of the Belarusian language, tagged with lexico-grammatical classes of words reference dictionaries: 1532 function words of the Belarusian language (BE), 1435 – of the Russian language (RU), 311 – of the English language (EN), 559 – of the French language (FR), 968 – of the German language (DE). The idea of the latter method is to use the lists of the most frequent trigrams of the natural language as the search profile of the language, instead of classical trigram method, in this case – 400 trigrams for each language. The nature of the algorithm remains the same (the comparison of the two lists). For all these methods the language of the input text is identified by the presence of the alphabet items unique to the language, plus function words and trigrams, respectively.

There have been selected 100 texts of 14 words each for all five languages as well as 100 texts of 7 words each to test these methods. These lengths have been taken on the assumption of the fact that, if the constructing algorithm is sufficiently accurate to solve the considering problem for the texts of 14 words long, then it will be even more effective to solve the problem for the sentence of the average length (e.g. this value found to be about 19 words for the well-known text corpus LOB Corpus [16]). If it is working in this way for the texts two times shorter (7 words), it can be used successfully to detect insertions of sentences and short phrases in other languages in the input text. The maximum score of $|A(T) \cap A_i|$, where $A(T)$ – the alphabet of the input text T and A_i – the alphabet of the language L_i, the set of their pairwise different words and the set of unique function words, respectively, the set of pairwise different trigrams and the set of the most frequent unique trigrams for the alphabetic method, the function word method and the alphabetic-trigram one, respectively, has been the criteria of the acknowledgement of the language L_i as the problem solution. If the maximum score equal for several languages simultaneously, it at least made possible to identify the language group of the text T.

The experiments have demonstrated the following results. The alphabetic method, using minimal resources, makes it possible to rate accurately the input text to the languages of the same or similar alphabets, in this case – to the language group BE-RU or EN-FR-DE. Thus, it can be used, for example, as one of the means to optimize the constructing algorithm. The function words method remains the most effective one, but taking into consideration the situation when none of the short texts contains the function words, it, obviously, allows the strengthening by resources of the unique trigrams.

Thus, the effective algorithm of solving the problem, that unites the possibilities of minimal statistical and linguistic analysis of lingual data, can be based on the combination of the alphabetic method, the function word method and the alphabetic-trigram method. So, due to the facts stated above, there is the schematic diagram of the algorithm:

1. Begin.
2. Process the text *T* with the alphabetic method for all NL of the given set. If the result is just one language L_i, then designate it as the problem solution and move to step 5.
3. Set the language group of the text *T* and process it with the function word method for those NL that belong to this group. If the result is just one language L_i, then designate it as the problem solution and move to step 5.
4. Set the language group of the text *T* and process it with the alphabetic-trigram method for those NL that belong to this group. Designate L_i as the solution of the problem.
5. End.

It's obvious, if the L_i at the step 4 will not be the only one, the supplementary resource to remove the ambiguity is needed, at least, for example, the interaction with the user.

4. Experimental Results

The defined algorithm has been implemented on the set of the above mentioned natural languages and tested on 500 texts, 100 texts for each language, including 25 texts for each group (texts of 4 Kb long, 5 sentences long, 14 words long, 7 words long). The test data included texts from the above mentioned text corpora used to investigate the trigram method of solving the problem.

The use of the alphabetic method in the first stage of the algorithm has allowed to determine the group of the languages of the same alphabet (BE-RU and DE-EN-FR), as well as accurately identify the language in the group BE-RU: BE – for texts of 4 Kb and 5 sentences long, RU – for texts of 4 Kb long. In this case the accuracy of the method has been 100%. There have been 24 (96%) texts correctly classified in the group BE-RU for RU from 25 texts of 5 sentences long. The method has successfully identified the language in case of the 14 words long texts: BE – 96 %, RU – 92 %, and the 7 words long texts: BE – 84 %, RU – 92 %.

The use of the function words method in the second stage of the algorithm has allowed to accurately (100%) redefine the languages in the group BE-RU: BE – for texts of 14 words long, RU – for texts of 5 sentences, 14 words and 7 words long. The language has accurately been identified in the group DE-EN-FR for texts of 4 Kb, 5 sentences, 14 words long in German, English and French languages. The method has successfully redefined the language of the 7 words long texts of BE – 75%, as well as correctly identified the languages DE, EN, FR in 96% of cases.

The use of the alphabetic-trigram method in the third stage of the algorithm has allowed to accurately redefine the languages in the group BE-RU: BE – for texts of 7 words long – 100%. There have been accurately identified the languages FR and DE for texts of 7 words long in the group DE-EN-FR.

The language of one 7 words long text in English hasn't been identified from 500 input texts. Thus, the accuracy of the algorithm was 99,8%. The total time needed to solve the problem for all 500 texts, via AMD Athlon 1800, RAM 1 Gb, has been 920,526 seconds.

5. Conclusion

The most famous existing methods, used to solve the problem of automatic identification of the text document, have been surveyed in this work. Those methods refer to the knowledge of a natural language from the alphabetic level up to the grammatical one and are sufficiently effective only under the certain restrictions on the size of the processing text and comparative characteristics of the natural languages. The effective algorithm of the combination of methods – the alphabetic, the function word and the alphabetic-trigram – has been presented. In general, the obtained results permit to consider it as an efficient tool to identify the text language, which can be used particularly in the systems of the automatic recognition of reproduced fragments of the text documents.

Acknowledgement

The author thanks scientific adviser professor I. Sovpel for participation and useful discussions.

References

[1] KRAPIVIN, Y. Plagiarism identification in multilingual information environment. *OWD: Proc. of the XIII International PhD Workshop.* – Gliwice. 2011. – P. 465-469.

[2] GREFENSTETTE, G. Comparing two language identification schemes. *The Third Intern. Conf. on Statistical Analysis of Textual Data.* – Rome, 1995.

[3] SIBUN, P., SPITZ, A.L. Language Determination: Natural Language Processing from Scanned Document Images. *Proc. of the 4th ACL Conf. on Applied Natural Language Proceeding (ANLP).* – Stuttgart, Germany, 1994.

[4] COWIE, J., LUDOVIC, Y., ZACHARSKI, R. Language recognition for mono- and multilingual documents. *Proc. of the Vextal Conference*. – Venice, 1999.

[5] SOUTER, C. Natural Language Identification using Corpus-based Models. *Hermes Journal of Linguistics*. – 1994. – № 13. – P. 183–203.

[6] CAVNAR W.B., TRENKLE J.M. N-Gram-Based Text Categorization. *Proc. of the 3rd Annual Symposium on Document Analysis and Information Retrieval (SDAIR).* – Las Vegas, 1994. – P. 161–175.

[7] PRAGER, J.M. Linguini: Language identification for multilingual documents. *Proc. of the 32nd Hawaii Intern. Conf. on System Sciences.* – Maui, Hawaii, USA, 1999.

[8] DUNNING, T. Statistical Identification of Language. *Computing Research Laboratory. Technical report MCCS.* – New Mexico State University, 1994. – P. 94–273.

[9] SIBUN, P., REYNAR, J.C. Language identification: Examing the issues. *Proc. of the 5th Symposium on Document Analysis and Information Retrieval (SDAIR).* – Las Vegas, 1996. – P. 125–135.

[10] POUTSMA, A. Applying MonteCarlo Techniques to Language Identification. *Proc. of Computational Linguistics in the Netherlands.* – Amsterdam, Netherlands, 2001.

[11] BIEMANN, C., TERESNIAK S. Disentangling from Babylonian Confusion – Unsupervised Language Identification. *Proc. of the CICLing-2005.* – Mexico City, 2005.

[12] KRUENGKRAI, C. Language Identification Based on String Kernels. *Proc. of the 5th Intern. Symposium on Communications and Information Technologies (ISCIT-2005).* – Beijing, China, 2005.

[13] GIGUET, E. Categorization according to Language: A step toward combining Linguistic Knowledge and Statistic Learning. *4th Intern. Workshop of Parsing Technologies.* – Prague, Karlovy Vary, Czech Republic, 1995.

[14] NEWMAN, P. Foreign language identification: First step in the translation process. *Proc. of the 28th Annual Conf. of the American Translators Association.* – Albuquerque NM, USA, 1987. – P. 509–516.

[15] QUASTHOFF, U,. RICHTER, M., BIEMANN, C. Corpus Portal for Search in Monolingual Corpora. *Proc. of the Fifth Intern. Conf. on Language Resources and Evaluation, LREC 2006.* – Genoa, 2006. – P. 1799–1802.

[16] JOHANSSON, S., LEECH, G., GOODLUCK, H. Manual of information to accompany the Lancaster-Oslo-Bergen Corpus of British English, for use with digital computers [http://khnt.hit.uib.no/icame/manuals/lob/index.htm]

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

An Algorithm for Finding All Minimal Cut Vectors in Reliability Analysis

[*]Miroslav Kvaššay
[*] University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, Miroslav.Kvassay@fri.uniza.sk

Abstract. One of the main methods that are used in reliability analysis is a method of minimal cut sets. A cut set is a set of components, whose simultaneous failure leads into the failure of the whole system (if a system has been operational). A cut set is minimal, if no component can be removed from it. Another term, which is very close to a cut set, is a cut vector. In this paper, a relation between a cut set and a cut vector is presented and a basic algorithm for computation of minimal cut vectors is derived and analyzed.

Keywords: Structure function, fault tree, minimal cut set, minimal cut vector.

1. Introduction

The analysis of minimal cut sets (MCS) is one of the basic methods for estimation of measures in reliability engineering [1], [2], [3]. For mathematical analysis, a system under consideration is represented by two mathematical models: binary-state and multi-state. In what follows, we only deal with binary-state systems that allow defining two states in system reliability/availability: failure/unavailable (it is 0) and functioning/available (it is 1). Every studied system consists of n components (elements), which can also be in one of two states (failure and functioning). The correlation between the system performance level and the system components states is defined by the structure function [4], [5], [6]:

$$\phi(\mathbf{x}) = \phi(x_1, x_2, \ldots, x_n) : \{0,1\}^n \to \{0,1\}, \tag{1}$$

where x_i is the i-th component state and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a vector of components states. A system is coherent if each system component is relevant, and the structure function (1) is non-decreasing [4], [5].

Every component is characterized by functioning (failure) state probability:

$$p_i = \Pr(x_i = 1),\; q_i = \Pr(x_i = 0),\;\; p_i + q_i = 1,\; i = 1,2,\ldots,n. \tag{2}$$

The availability and unavailability of the whole system are defined as follows:

$$A = \Pr(\phi(\mathbf{x}) = 1),\; Q = \Pr(\phi(\mathbf{x}) = 0),\;\; A + Q = 1. \tag{3}$$

2. Mathematical Background

2.1. A Minimal Cut Set

One of the basic definitions of a MCS is based on the fault tree analysis, which is one of the most commonly used techniques for risk and reliability studies. A fault tree is a logic diagram that displays the interrelationships between a potential critical event (accident) in a system and the causes for this event. In the fault tree terminology, a cut set is defined as follows:

A cut set in a fault tree is a set of basic events whose occurrence (at the same time) ensures that the TOP event (the system failure) occurs. A cut set is said to be minimal if the set cannot be reduced without losing its status as a cut set [7].

This definition can be used for coherent and non-coherent systems. However, this paper only deals with a coherent systems and so from the fact that every coherent system can be described by a fault tree that consists of OR, AND gates and basic events, which only represent components failures [8], we can reformulate the previous definition of the MCS to the following one:

A cut set is a set of components of a system whose simultaneous failure leads into the failure of the system (if the system has been operational). A cut set is minimal, if no component can be removed from it without losing its status as a cut set.

2.2. A Minimal Cut Vector

In the rest of the paper, using the convention that $\mathbf{x} > \mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ are two states vector, means that $x_i \geq y_i$ (for $i = 1,2,\ldots,n$) and there exists at least one i such that $x_i > y_i$.

The last definition of a MCS, introduced in the previous section, is not bounded to the fault tree terminology but only to a system structure. The first part of the definition implies that if all components in a cut set are failed then the system is failed, regardless of states of other components. It means that if all components, which are not in the cut set, are functioning (they are in 1-state) and the cut set components are failed (they are in 0-state), then the system is failed (it is in 0-state). This statement leads into the following definition of a cut vector [4]:

A state vector $\mathbf{x}$ is a cut vector if $\phi(\mathbf{x}) = 0$.

The coherency of a system implies that there exists at least one state vector, for which a system is working. It means that if all components of a cut set are failed and other components are functioning (the system is failed), then there exists at least one subset of the cut set, such that if all components of the subset are working then the system is also functioning. When there is no subset that is different from the cut set, then the second part of the second definition of a MCS implies that the cut set is minimal. According to [4] the definition of a minimal cut vector (MCV) is as follows:

A state vector $\mathbf{x}$ is a cut vector if $\phi(\mathbf{x}) = 0$. A cut vector $\mathbf{x}$ is minimal if $\phi(\mathbf{y}) = 1$ for any $\mathbf{y} > \mathbf{x}$.

The results of the previous paragraphs are:

- if a state vector $\mathbf{x}$ is a (minimal) cut vector, then the corresponding (minimal) cut set consists of components, which are in 0-state,
- if there is a (minimal) cut set, then the corresponding (minimal) cut vector has 0-value in the positions of all components from the (minimal) cut set and 1-value in other.

3. An Algorithm for Finding MCVs

In the following section, we present a simple algorithm for computation of MCVs. This algorithm is based on the definition of a MCV and from the relation between a MCV and a MCS it is clear, that it can be used to find all MCSs of a system that is defined by structure function.

Firstly, we have a state vector $\mathbf{s} = (s_1, s_2, \ldots, s_n)$, where $0 \leq s_i < 2$, for $i = 1,2,\ldots, n$. Using the following notation:

$$\mathbf{s}^{(i)} = \left(s_1, s_2, \ldots, s_{i-1}, 1, s_{i+1}, \ldots, s_n\right), \tag{4}$$

it is clear that:

$$\forall \mathbf{y}; \mathbf{y} > \mathbf{s} : \left(\exists i \in \{1,2,\ldots,n\} : \mathbf{s}^{(i)} \leq \mathbf{y}\right). \tag{5}$$

This statement implies the following definition of a MCV:

$$\left(\mathbf{s} \in \text{MCVS}\right) \Leftrightarrow \left(\forall i \in \{1,2,\ldots,n\}; \mathbf{s}^{(i)} \neq \mathbf{s} : \phi\left(\mathbf{s}^{(i)}\right) = 1\right), \tag{6}$$

where MCVS is a set of all MCVs of a system. This statement leads into the following algorithm for computation of MCVs:

```
1. s = (0,0,…,0)
2. for i = 1,2,…,n:
      if s^(i) ≠ s AND φ(s^(i)) = φ(s) then
         go to step 4
3. MCVS = MCVS ∪ {s}
4. s++
5. if s < (1,1,…,1) then
      go to step 2
   else
      stop
```

The operation $\mathbf{s}++$ is implemented as follows:

```
i = n
s_i = s_i + 1
while (i > 0 AND s_i = 2):
   s_i = 0
   i = i - 1
   s_i = s_i + 1
```

The logical test $\mathbf{s} < (1,1,\ldots,1)$ can be implemented with respect to the implementation of the operation $\mathbf{s}++$ as follows:

```
s_1 < 2
```

Now, we investigate the computational complexity of this algorithm. The first step has complexity $O(n)$. Other steps represent a loop that is repeated 2^n times. So, we can analyze an average complexity of individual steps and then multiply it by numbers of repetitions.

In the second step, a structure function value of a state vector **s** is compared with structure function values of vectors $\mathbf{s}^{(i)}$. The average number of comparisons can be derived as follows:

In a system, there exists only one vector, which must be compared up to n vectors (a state vector (0,0,…,0)). Then, there are n vectors that are being compared to mostly (n -1) vectors (state vectors (1,0,…,0), (0,1,…,0),…, (0,0,…,1)). Obviously, if a state vector has number 1 at k positions then it must be compared up to (n - k) vectors, and there also exist $\binom{n}{k}$ different state vectors that have number 1 at k positions. So, a maximal number of comparisons is:

$$\mathrm{MNC} = \sum_{k=0}^{n}(n-k)\binom{n}{k} = \sum_{k=0}^{\frac{n}{2}}(n-k+k)\binom{n}{k} = n\sum_{k=0}^{\frac{n}{2}}\binom{n}{k} = \frac{n}{2}2^n . \tag{7}$$

Since the algorithm has to investigate all state vectors (there exist 2^n different state vectors), then we can compute the average number of comparisons:

$$\mathrm{ANC} = \frac{\frac{n}{2}2^n}{2^n} = \frac{n}{2}. \tag{8}$$

When the average number of operations for calculation of a structural function value of a state vector is $f(n)$, then the average complexity of the second step is $O(nf(n))$.

The third step is a special case, because it is executed only if a state vector **s** is a MCV. From this, it is clear that the overall complexity of this step is $O(|\mathrm{MCVS}|)$.

In the fourth step the operation $\mathbf{s}{+}{+}$ is repeated 2^n times. The first two (or three) steps of this operation are repeated every time. The other steps, which are within the while-loop, are executed only if the while-condition is true:

During the execution of the algorithm, the while-condition is satisfied 2^{n-1} times at least once (the three steps within the while-loop are executed at least once), the condition is true 2^{n-2} times at least twice (the three steps within the while-loop are executed at least twice),…, the condition is true 2 times at least (n -1) times (the three steps within the while-loop are executed at least (exactly) (n -1) times). So, we have for the number of repetitions of steps in $\mathbf{s}{+}{+}$ operation:

$$\sum_{k=1}^{n} 2^k = 2\frac{2^n - 1}{2-1} = 2^{n+1} - 2 . \qquad (9)$$

We get for average number of repetitions of steps in the operation $\mathbf{s}{+}{+}$:

$$\frac{2^{n+1} - 2}{2^n} \approx 2 . \qquad (10)$$

So, the average complexity of the fourth step is $O(1)$.

Finally, the average complexity of the fifth step is $O(1)$. As we can see from the analysis of individual steps of the algorithm for finding MCVs, the average computational complexity of all steps in the main loop (except the third step) is $O(nf(n))$ and it depends only on the third step. Then, it is clear that the overall complexity of the whole algorithm is:

$$O\left(n + \left|\mathrm{MCVS}\right| + nf(n)2^n\right) = O\left(nf(n)2^n\right). \qquad (11)$$

From the analysis of the algorithm, it is clear that the efficiency of the algorithm for finding all MCVs of a coherent system can be increased by reducing a number of investigated state vectors or by reducing an average number of comparisons of state vectors.

4. Conclusion

MCSs are widely used in reliability analysis. Another term that is very close to a MCS is a MCV. In this paper, the relation between these two terms is presented and an exact algorithm for computation of MCVs is derived. The analysis of the algorithm has shown that there are two possible ways how to improve its effectiveness. These ways can be used in further research that can be oriented to design more effective algorithms.

References

[1] FUSSELL, J. B., *How to hand-calculate system reliability and safety characteristics*. *IEEE Transactions on Reliability*, vol. R-24, no. 3, 1975, pp. 169 -174.

[2] VESELY, W. E., *A time-dependent methodology for fault tree evaluation. Nuclear Engineering and Design,* vol. 13, no. 2, 1970, pp. 337 - 360.

[3] RAUZY, A., *Mathematical foundations of minimal cutsets. IEEE Transactions on Reliability*, vol. 50, no. 4, 2001, pp. 389 -396.

[4] LISNIANSKI, A. and LEVITIN, G., *Multi-State System Reliability*. World-Scientific, 2003.

[5] ZAITSEVA, E., *Importance analysis of a multi-state system based on multiple-valued logic methods*. *Recent Advances in System Reliability,* Springer Series in Reliability Engineering Lisnianski, A. and Frenkel, I., Edt., Springer London, 2012, pp. 113-134.

[6] ZAITSEVA, E. and LEVASHENKO, V., *Dynamic reliability indices for parallel series and k-out-of-n multi-state systems. Reliability and Maintainability Symposium*, *2006. RAMS '06. Annual,* 2006, pp. 253 -259.

[7] RAUSAND, M. and HØYLAND, A., *System reliability theory,* Wiley-Interscience, 2004.

[8] ANDREWS, J. D. and BEESON, S., *Birnbaum's measure of component importance for noncoherent systems*. *IEEE Transactions on Reliability,* vol. 52, no. 2, 2003, pp. 213 - 219.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Relational Databases Time Management

*Michal Kvet
*University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, Michal.Kvet@fri.uniza.sk

Abstract. Data processing requires keeping actual, historical and future valid data in the database. Conventional tables store only actual valid objects. Uni-temporal table offers storing whole states of the object during its existence, even after logical delete. This paper describes the problems, defines and compares solutions – special methods, structures, operations to provide easy access to these data. Thus, each database object is not represented by only one row, but by the set of the rows with limited validity.

Keywords: Temporal table, conventional table, valid time, snapshot of the database, log files, backups.

1. Introduction

Data processing requires easy access to extensive data to provide data in proper form and time using procedures, functions and other methods of the database technology.

Most of the data kept in the database describe current states, not what they used to be like. If the object state is going to be changed, data in the database are updated and the database will still contain only currently applicable object states. But all things have its own history, many of them also future as well (customer changes its address, status, products are modified and updated). Management often requires providing historical data for decision making. It is usually possible to restore historical data using log files and backups, together with associated delays and costs. Temporal models can be used in many fields - industry, medicine, transportation – mostly in systems which unauthorized modification would cause great problems or systems are processing previous states of the objects.

2. Indexes

The aim of creating indexes is to improve performance of the database system by reducing execution time of the query. However, the issue is more complicated and requires complex solutions. The problem is that the traditional indexes contain variables (pointers) to the rows based on discrete values. However, time interval is not discrete, although it contains discrete time points based on clock-tick of the database - the smallest distinguishable time.

Intervals are thus defined as a continuous bounded sequence of discrete values (times). Begin and end time limit the effective time (validity) or transaction time. In fact, the intervals of the validity (uni-temporal databases) and transaction intervals (bi-temporal databases) cannot be joined together because they characterize the time in two separate time axes. Therefore, it is necessary to describe these periods separately. Traditional indexes constructed on the basis of weighted tree work effectively with discrete values. Their efficiency rapidly decreases with processing time intervals.

Nevertheless, new indexing methods have been developed. They can process intervals, but not the time. Specifically, bi-temporal index can model validity interval and also transaction time interval in the graph as a Cartesian rectangle - the validity dimension of time is represented on the X axis, transaction period (or other periods) on the Y-axis.

Another approach models these intervals separately because of the different meaning and different behavior. While the transaction time is always moving forward, the validity representation can define not only the future, but also the past. This means that the row in bi-temporal model can be inserted pro-actively (defining the future) based on validity, but not the transaction time. [1] [2] [3]

3. Conventional, uni-temporal and bi-temporal table

The row in a relational database table can be defined in three different ways concerning the time (fig. 1). *ID* is a unique identifier; *PK* refers to a primary key. *BD1* and *ED1* are pairs of columns defining the beginning and end value of the period, *BD2* and *ED2* define the second time interval. The primary key can be also composite without changing the functionality of used methods. The first model does not use time for definition. This is a standard model, called the conventional model. [3] [4]

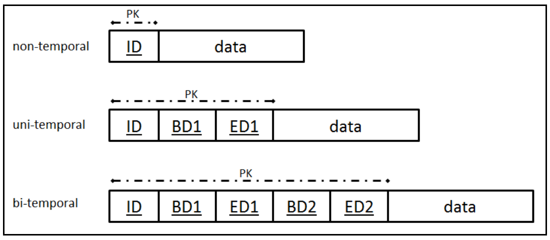


Fig. 1. Conventional and temporal table

The following table shows the uni-temporal table - the object identified by an attribute value *ID=1* in the time period <September 2012 - December 2013>. We can see that the attribute *data* has the value *123* at the time <September 2012-June 2013>, in the second part of the interval, attribute *data* has the value of the *234*. The representation of the time interval is the closed-closed type.

ID	BD1	ED1	data
1	**September 2012**	**June 2013**	**123**
2	January 2013	November 2014	555
1	**July 2013**	**December 2013**	**234**

Tab. 1. Uni-temporal table [4]

The problem occurs, if the end date of the time period is not defined. There are two possibilities, in general. The first is to deal with undefined values - NULL. In this case, if there is a request to obtain current state of the object, respectively overall database, standard methods to manipulate time cannot be used. The second solution is to replace an undefined time with the maximal value that can be used - *December 9999*, or other variants depending on time granularity. If we need to update the state of the object, the attribute value defining the end of the validity (attribute *ED*) is replaced by the actual time value.

The last - third basic model – bi-temporal model - is based on the concept of uni-temporal tables, but uses two time intervals. Thus, it allows not only defining several rows for one object, but also multiple rows for a particular object's state at the time. The reason for this model is the need to record an updated status.

Primary key of the bi-temporal model consists of three logical units [4] [5]:

- The object identifier (*ID*).
- Interval (*BD1*, *ED1*) - the time during which the object has been describing the characteristics of the row (e.g. the period during which a customer has the characteristics - name, address, status, etc.).
- The last component of the logical primary key is a pair of dates (or timestamps according to the representation of a time granularity of data). These dates limit the period during which we believe the value of the row is correct. This component limits the time interval defined by the second component (*BD1*, *ED1*).

4. Uni-temporal table modelling

Transformation of the conventional table to temporal model is not trivial problem and needs special structures and resources. In addition, the requirement of the users is to provide compatibility, easy manipulation and proper time consumption. Therefore, triggers, procedures and functions must be declared, original tables can be transformed to views (if necessary). Conventional database tables transformation recommends single attribute primary key in every table to create a common temporal table for all tables containing temporal attributes. Thus, ID is suitable; each record can be clearly defined and referenced. Moreover, the ID does not have special denotation and the need for its change is irrelevant (e.g. personal identification number contains the birthdate and if the mistake in time of insert occurs, the record must be updated).

Information of any change of temporal column is recorded in the table managing changes - temporal_table. However, it contains information only about table containing temporal column. This table consists of these attributes (see also fig. 2): [4]

- *ID change*
- *ID previous change* **–** references the last change of an object identified by ID. This attribute can also have NULL value that means, the data have not been updated yet, so the data were inserted for the first time in past and are still actual.
- *ID_tab* – references the table, record of which has been processed by DML statement (INSERT, DELETE, UPDATE).
- *ID_orig* - carries the information about the identifier of the row that has been changed.
- *ID_column, ID_row* – hold the referential information to the old value of attribute (if the DML statement was UPDATE). Only update statement of temporal column sets not null value.
- *BD* – the begin date of the new state validity of an object.

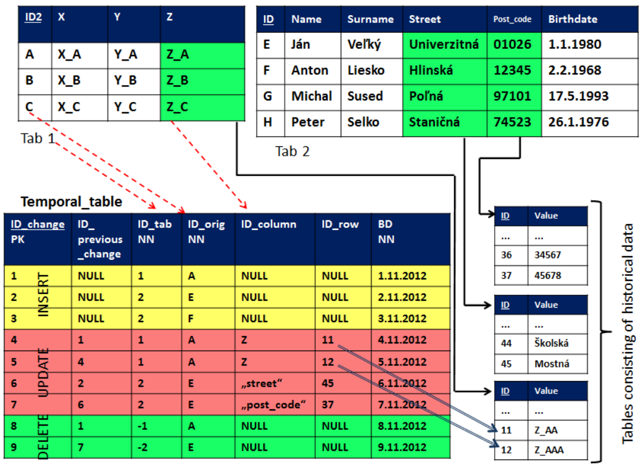


Fig. 2. Temporal_table model [4]

5. Experiments

Figure 3 shows the total execution time of DML operations (insert, delete, update) for conventional and uni-temporal tables. The total number of operations is 10 000.

Model 1 does not use the table with maximal values of changes, data are inserted into historical tables only if they have not been there yet. Model 2 does not use table with maximal values of changes, data in historical tables are always saved. Model 3 improves the storage problem of model 2 by inserting data into historical tables only if they have not been there yet. Model 4 manages maximal changes, historical data are always saved. The last model saves the maximal change of the object in the main table. The reason is, that a lot of data during the time are deleted, so the number of real used data is lower. We must also store info about deleted data in change id table, too.

The overall slowdown of this model (5) is:

- insert operations – 3%,
- delete operations – 34%,
- update operations – 34%.

As it was already mentioned, the results of the experiments can be considered satisfactory, because of the storing the entire history of temporal columns.

		model 1		%		model 2		%		model 3		%	
INSERT	temp	0:24:775	0:27:272	101%	104%	0:25:578	0:25:781	101%	101%	0:26:28	0:26:4493	105%	104%
	conv	0:24:474	0:26:265	100%	100%	0:25:256	0:25:513	100%	100%	0:25:00	0:25:475	100%	100%
UPDATE (10)	temp	13:22:678	16:8:667	384%	468%	12:17:029	11:16:427	276%	270%	7:36:267	6:53:765	178%	196%
	conv	3:44:534	3:59:003	100%	100%	4:41:323	4:14:05	100%	100%	4:12:192	3:33:79	100%	100%
DELETE	temp	0:52:638	0:54:042	2605%	2470%	0:58:896	53:989	3166%	2885%	0:4:099	0:4:045	173%	170%
	conv	0:2:021	0:2:188	100%	100%	0:1:86	1:871	100%	100%	0:2:364	0:2:38	100%	100%

		model 4		%		model 5		%	
INSERT	temp	0:26:345	0:28:538	104%	101%	0:25:123	0:25:994	104	101%
	conv	0:25:112	0:28:173	100%	100%	0:25:236	0:25:712	100%	100%
UPDATE (10)	temp	7:01:754	7:36:109	158%	162%	6:32:273	6:12:109	138%	131%
	conv	4:44:265	4:55:91	100%	100%	4:45:185	4:43:572	100%	100%
DELETE	temp	0:3:825	0:4:226	140%	138%	0:3:212	0:4:412	131%	137%
	conv	0:2:732	0:3:063	100%	100%	0:2:452	0:3:263	100%	100%

Fig 3. Research results

6. Conclusion

Each instance in a conventional database is represented by one row. Temporal database concept offers new opportunities by adding additional time interval attributes limiting the validity of the object or attribute. Thus, object can be represented by more rows. They characterize the properties and description of the object during given time point or time interval. Database has information about the whole life-cycle of the objects, even after logical delete. It brings the opportunity to create prognoses and analyses.

Acknowledgement

This contribution is the result of the project implementation:
Centre of excellence for systems and services of intelligent transport II.,
ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

"Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ"
This work was partially supported by the project **"Creating a new diagnostic algorithm for selected cancers,"** ITMS project code: 26220220022 co-financed by the EU and the European Regional Development Fund.
The work is also supported by the project VEGA 1/1116/11, Adaptive data distribution.

References

[1] DATE, C. J., DARWEN, H., LORENTZOS, N.A. *Temporal data and the relational model*, Morgan Kaufmann, 2003. ISBN: 1558608559

[2] JENSEN, Ch. S., SNODGRASS, R. T. *Temporally Enhanced Database Design*

[3] JOHNSON, T., WEIS, R. *Managing Time in Relational Databases*, Morgan Kaufmann, 2010. ISBN: 9780123750419

[4] KVET M., LIESKOVSKÝ, A., MATIAŠKO, K. *Temporal data modelling*, 2013.(IEEE conference ICCSE 2013, 4.26. – 4.28.2013) in press

[5] KVET M., MATIAŠKO, K. *Conventional and temporal table*, EDIS, 2012. (virtual conference ARSA, 12.3. – 12.7.2012). ISBN: 978-80-554-0606-0

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Measuring of Jitter Noise in Symbol Map for M-QAM Digitally Modulation

*Radek Martinek, *Mohamed Al-Wohaishi
*Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, 17. Listopadu 15, 708 33 Ostrava, Czech Republic
{radek.martinek.st1, jan.zidek}@vsb.cz

Abstract. Usually for RF applications such as system transmission and reception of digital signals (including satellite systems) and DVB-T, it is necessary to understand Jitter noise by analysis the effect of it on the symbol map for modulation schemes during the transceiver process.

Authors of this paper present real results which analysis the effect of Jitter noise in the demodulated IQ data for a number of symbols for a modulation schemes, which illustrate the link between number of symbols in M-state QAM modulation, SNR, BER, Constellation diagrams and Spectral Analysis diagram of transmitted and received symbols that used in a widespread use in current and emerging technologies. The obtained results are important to make a communication system more immune to channel noise.

Keywords: Jitter Noise, SNR, Spectral Analysis, M-QAM, LabVIEW.

1. Introduction

In digital communication system, there are two specific factors affect proper demodulation of the carrier signal. These factors include the signal-to-noise ratio and order (M) of the modulation scheme [1].

In this paper, we will evaluate the relationship between these two factors. By adjusting the simulated noise level of the physical channel for different order (M) of QAM modulation to observing the Jitter noise accuracy that happened in demodulated IQ data for order 4, 8, 16, 32, 64, 128, and 256 QAM [5].

Due diversity of communication protocols implement in QAM modulation. Current protocols include such as Digital Video Broadcast (DVB) and 802.11b wireless Ethernet (Wi-Fi) both implement 64-QAM modulation. It can significantly use QAM modulation with a large number of symbols to get faster data rates than other modulation schemes such as PSK, ASK.

2. System Model

This modulation is used primarily in systems where high spectral efficiency is required. Quadrature amplitude modulation (QAM) requires changing the phase and amplitude of a carrier sine wave.

This modulation scheme has M symbols (states), each of them represents n bits, where $M=2^n$. These symbols are transmitted with duration of symbol period Ts:

$$s_1(t), s_2(t), \ldots, s_M(t). \tag{1}$$

The symbols are expressed by general formula [2]:

$$s_{MQAM\,(t)} = \sqrt{\frac{2E_{\min}}{T_s}} a_i \cos(2\pi f_c t) + \ldots$$
$$\ldots + \sqrt{\frac{2E_{\min}}{T_s}} b_i \sin(2\pi f_c t). \tag{2}$$

For $0 \le t \le T_s, i = 1,2,\ldots, M$, where a_i, b_i are the pairs of independent integers corresponding to the position of symbols in constellation diagram. One of the easiest ways to implement QAM with hardware is to generate and mix two sine waves that are 90 degrees out of phase with one another. Adjusting only the amplitude of either signal can affect the phase and amplitude of the resulting mixed signal.

Assuming that the modulating pulses are rectangular, it is possible to express the signal $S_{MQAM}(t)$ using basis functions, which are defined by following equations [2]:

$$\Phi_1(t) = \sqrt{\frac{2}{T_s}} \cos(2\pi f_c t);$$
$$\Phi_2(t) = \sqrt{\frac{2}{T_s}} \sin(2\pi f_c t), \tag{3}$$

for $0 \le t \le T_s$. The coordinates of i-the point (symbol) are [2]:

$$a_i \sqrt{E_{\min}}, b_i \sqrt{E_{\min}}. \tag{4}$$

Our Experiment Model represents the use of LabVIEW Modulation Toolkit. This Experiment demonstrates continuous acquisition and demodulation of a QAM signal see figure 1 below.

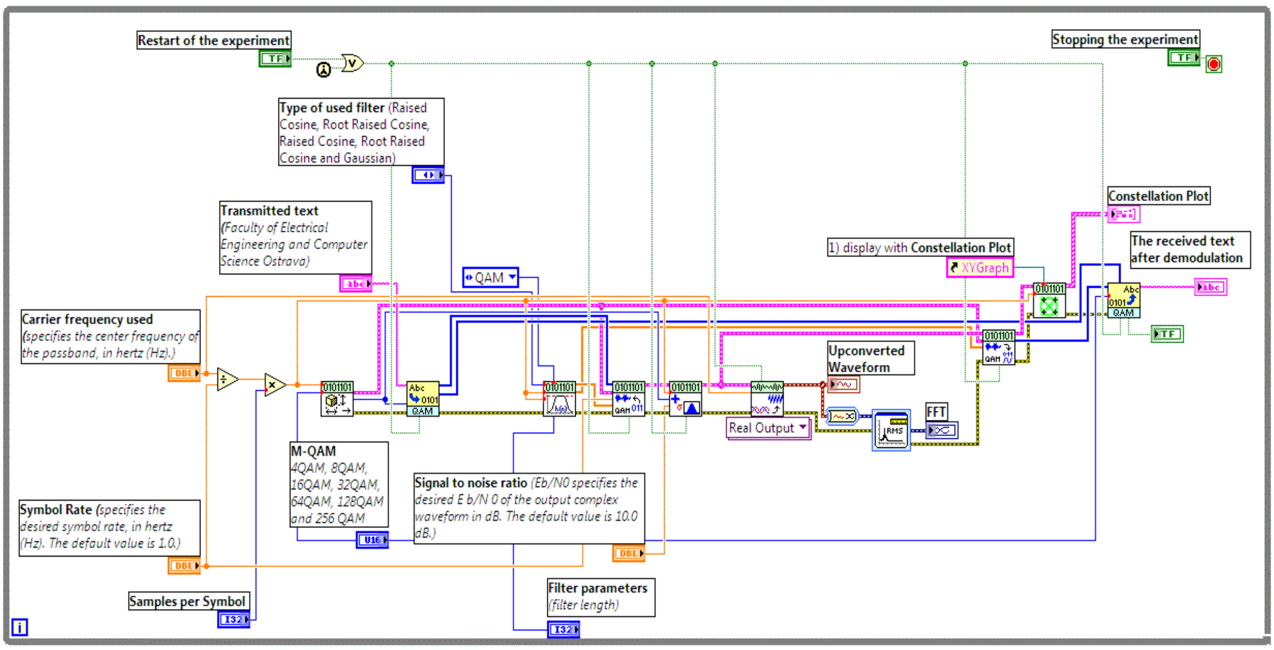


Fig. 1. Model Experiment.

3. Analysis of Jitter

Jitter is fundamentally an expression of phase noise. Mathematically, jitter is the undesired variation in the phase of a signal given by the term, ϕ (t), in the expression [3]:

$$S(t) = P(2\pi f_d t + \varphi(t)). \tag{5}$$

Where S is the received signal, P represents the sequence of signal pulses as a function of time, and f_d is the data rate. Jitter isn't measured simply to create statistics; it is measured because jitter can cause transmission errors. For example if jitter results in a signal being on the "wrong side" of the transition threshold at the sampling point, the receiving circuit will interpret that bit differently than the transmitter intended, causing a bit error [4]. See figure 2.

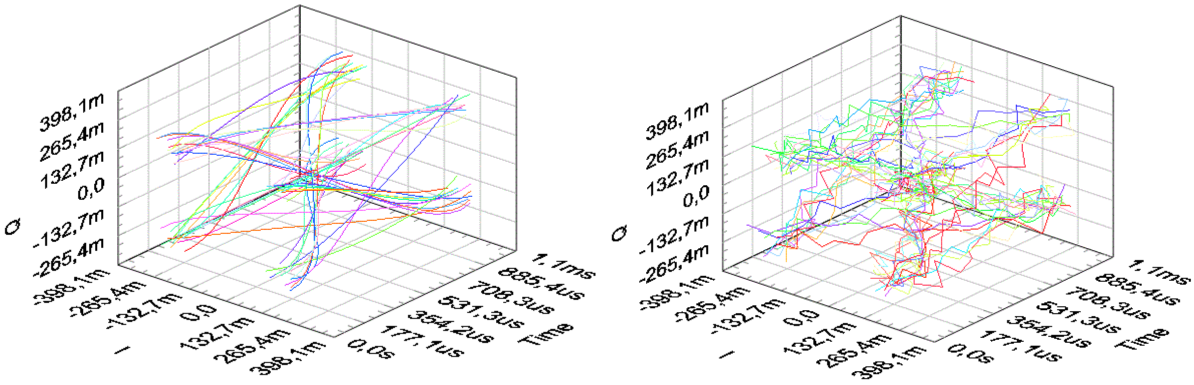


Fig. 2. Jitter noise in 3D Eye diagram for 4-QAM respectively with Eb/N0 (100dB) & (30dB).

4. M-QAM Modulation

Quadrature Amplitude Modulation, QAM, has fast become the dominant modulation mechanism for high speed digital signals. From the wireless 802.11 protocols to ADSL modems to personal communicators for the military, QAM has become a necessary part of our daily lives. With increases in processing power, QAM as a part of software defined radio (SDR) schema is now easily achievable.

QAM is a modulation scheme which is carried out by changing (modulating) the amplitude of two carrier waves. The carrier waves are out of phase by 90 degrees, and are called quadrature carriers - hence the name of the scheme. QAM can be viewed as a combination of ASK and PSK. That means the digital information is carried in both the phase and the amplitude of the carrier signal [5].

Quadrature amplitude modulation, or QAM, extends the idea of PSK to modulating the pulse amplitudes as well as the phases. As a generic communication term, QAM implies linear I and Q modulation and carrier, in contrast to PAM, which implies single-channel linear baseband modulation. A general from of a QAM signal is once again Eq, with the data mapped to M two-tuples (a_n^I, a_n^Q), but this time the resulting constellation is more than just the PSK circle. Some QAM signal space constellations are shown in figure 2, see [1].

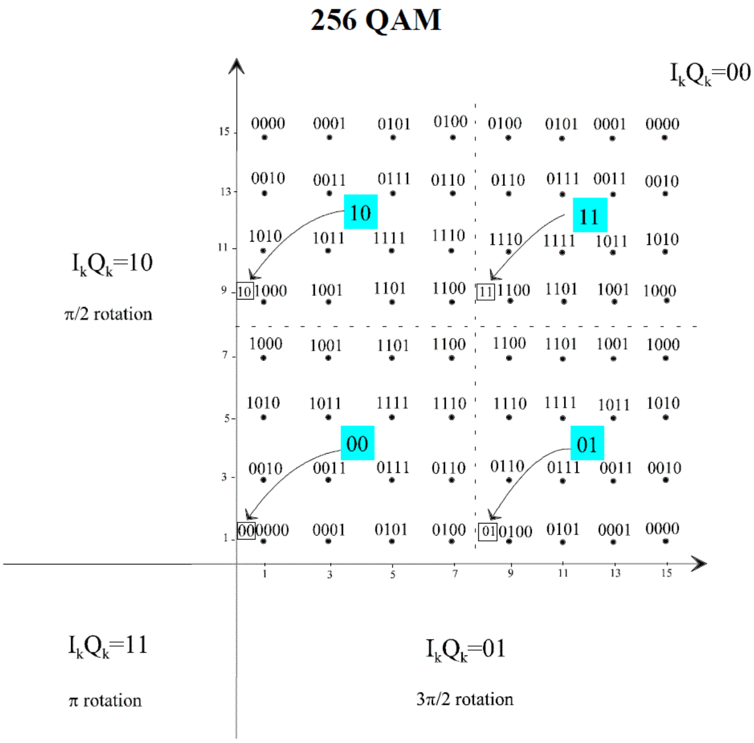


Fig. 3. Constellation diagram respectively for 256 QAM.

5. Results and Discussion

From this experiment we illustrated the effect of channel noise on an M-ary QAM signal. By selecting some order (M) of QAM modulation and adjust the simulated noise level of the physical channel, as shown in figure 4.

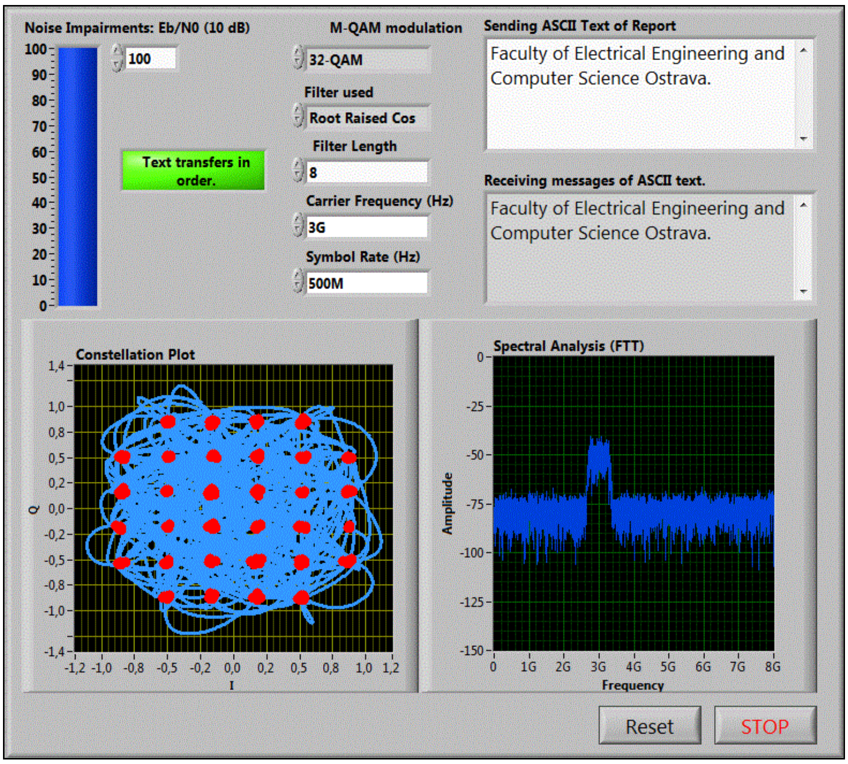


Fig. 4. Front panel parameters for measuring Jitter noise.

We will examine the behavior can occur when Bit error ratio is significant enough to prevent or occurrence of Jitter noise to the received IQ symbol for a specific order (M) of QAM modulation in constellation and spectral diagram as shown in figure 5.

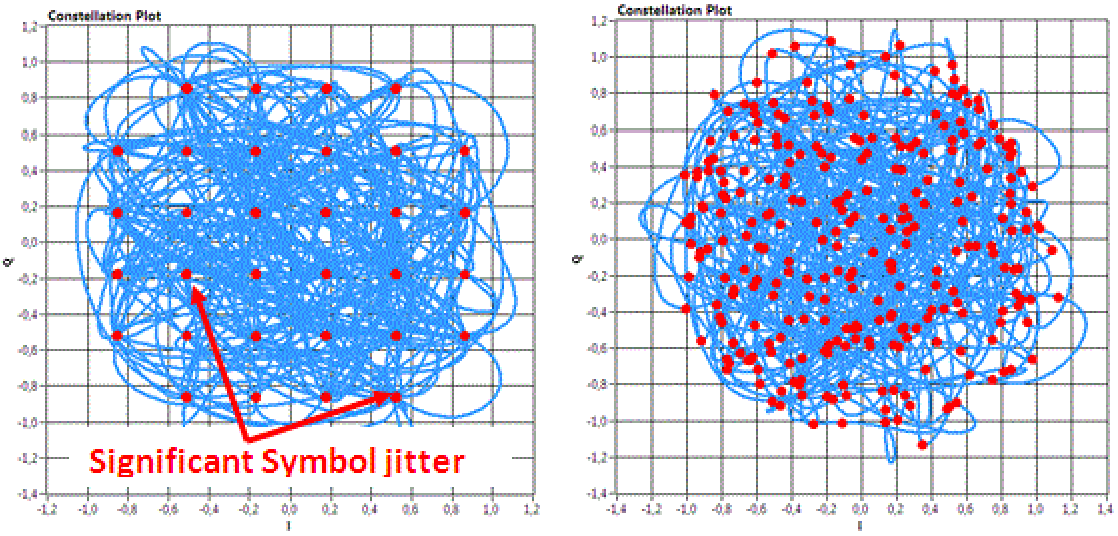


Fig. 5. Jitter noise represents in the physical channel.

The constellation plot or every symbol in the QAM symbol map is distinguished by a unique phase and amplitude. That's mean; the order of the modulation scheme defines how many symbols are in the specific symbol map. In figure 6 respectively we observe that when we set Eb/N0 (Signal to noise ratio) set to 100 dB (minimal channel noise) the constellation plot shows each received symbol mapped accurately to an ideal symbol location and the result that the string match (green), means the (message to send). In the other image, increasing the (Signal to noise ratio) to 30 dB the demodulated bitstream cannot be matched with the original bitstream.

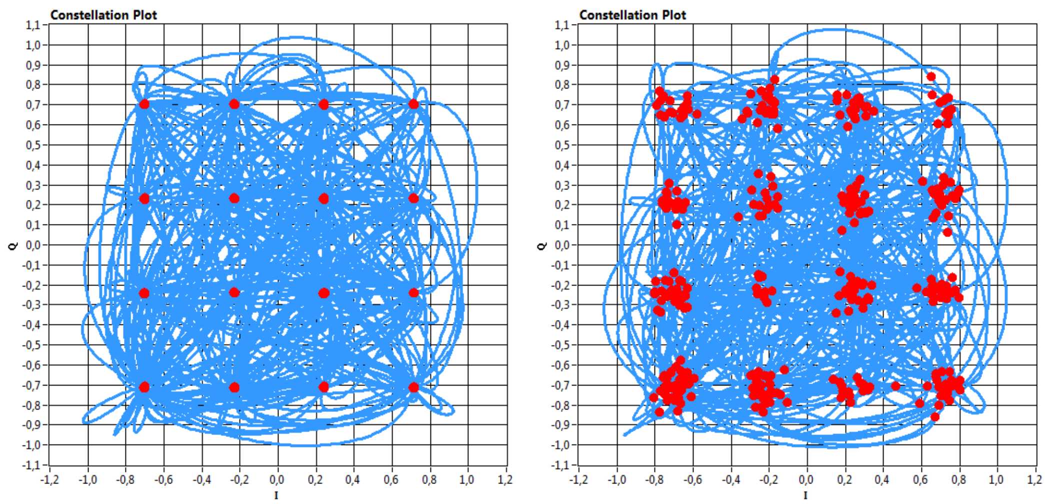


Fig. 6. Constellation plot16 QAM respectively with 100(dB) & 30 (dB).

In the other order of 64 QAM with the same parameters to get the result and this occurs because the phase of the carrier cannot be properly recognized we can observe it below as shown in figure 7.

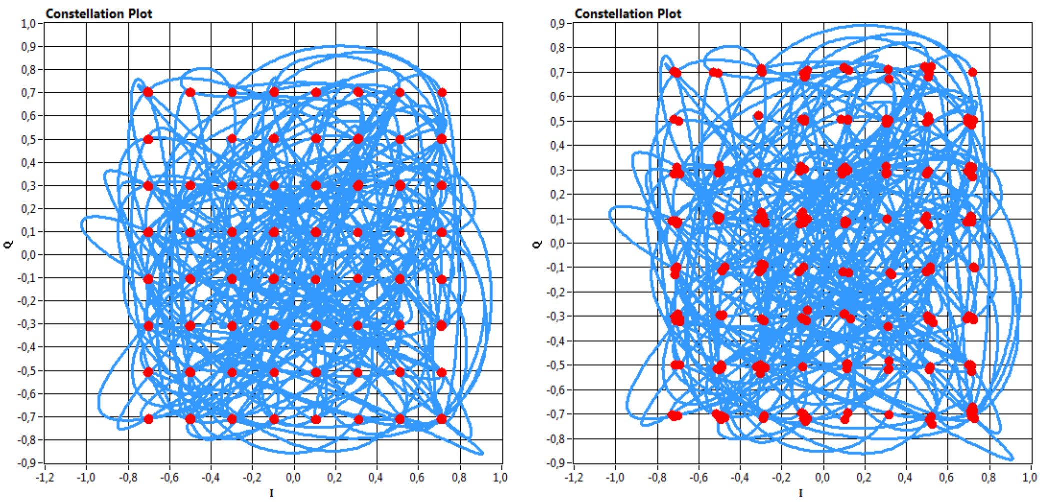


Fig. 7. Constellation plot 64 QAM respectively with 100(dB) & 30 (dB).

For the same order 16 QAM and 64 QAM with the same parameters we determined the Jitter noise by using the spectral analysis by determining the frequency domain representation of a time domain signal as shown in figure 8 and 9 below.

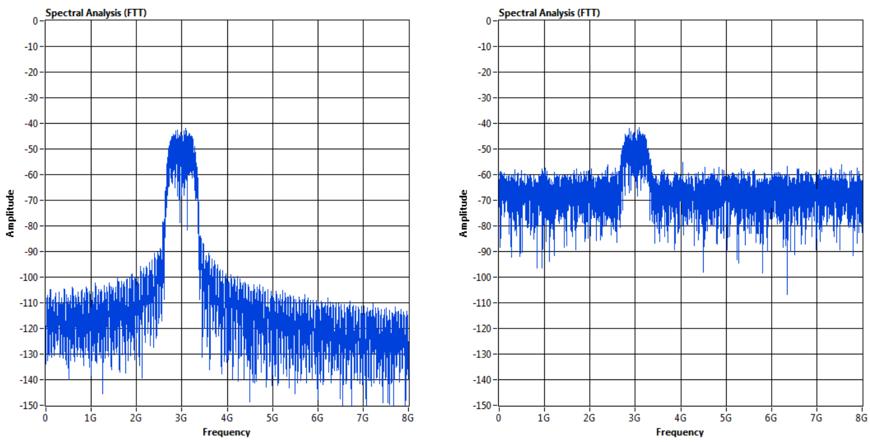


Fig. 8. Spectral analysis 16 QAM respectively with 100(dB) & 30 (dB).

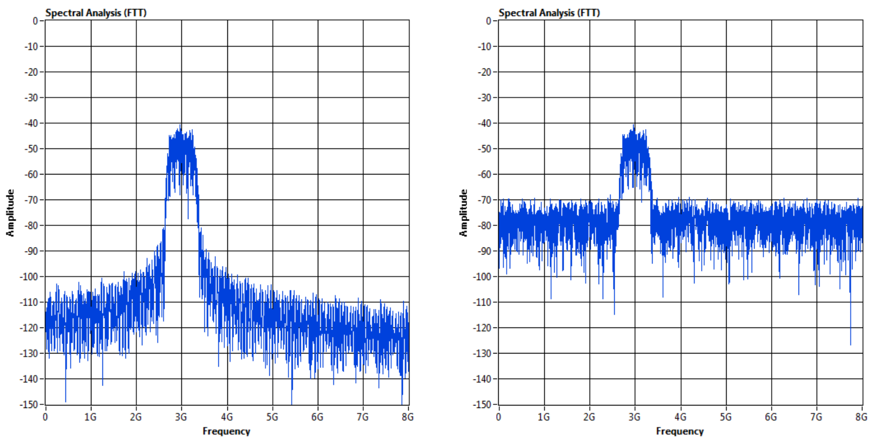


Fig. 9. Spectral analysis 64 QAM respectively with 100(dB) & 30 (dB).

Fig. 10 shows a diagram BER (Bit Error Rate [11]) vs. Eb/No (the energy per bit to noise power spectral density ratio).

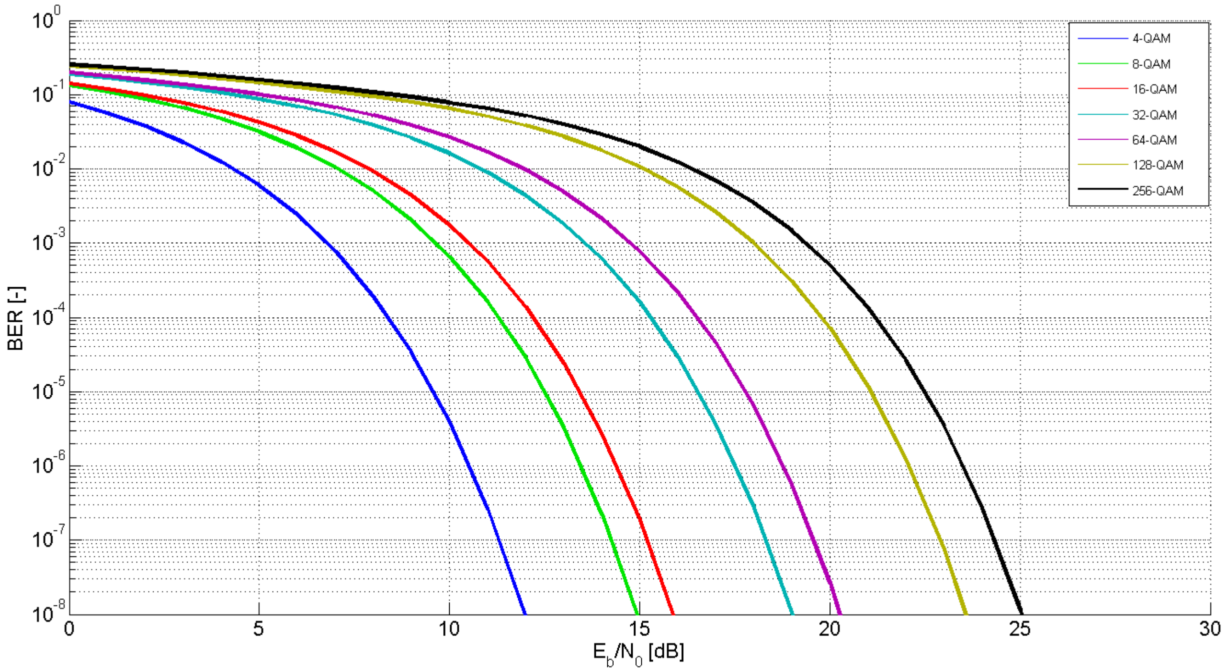


Fig. 10. BER vs. Eb/No for M-QAm.

6. Conclusion

It was clearly observed that noise in the physical channel can have a significant effect on the ability of the transceiver to recover the original bit stream. From the figures result we can recognize that the received IQ symbol are still susceptible to Jitter noise for the order (M), we can see in the bigger number of symbols for a symbol map will increase the bitrates and channel throughput, it makes a communication system more susceptible to channel noise. In contrast, by using a smaller number of symbols for the same received IQ symbol it will be easy to distinguish between them even when there is significant Jitter noise.

Acknowledgement

This research was supported in part by The Ministry of Education, Youth and Sports of Czech Republic under the project KONTAKT II registration number LH12183.

References

[1] M. Al-Wohaishi; R. Martinek, J. Židek; “Analysis of M State Digitally Modulated Signals in Communication Systems Based on SDR Concept”. In Proceedings of The 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2011); first issue. Praha, Czech Republic; IEEE Ukraine & ČVUT Praha, 2011; 5 pages.

[2] Ke-Lin Du, M. N. S- Swamy, (2010) “Wireless Communication Systems’’ Cambridge University Press, ISBN-13: 9780521114035.

[3] Agilent (no author listed), Measuring Jitter in Digital Systems, Application Note 1448-1, http://cp.literature.agilent.com/litweb/pdf/5988-9109EN.pdf, June 1, 2003.

[4] S. Saad, The effect of accumulated time jitter on some sine-wave measurements,” IEEE Trans. Instrum. Meas., 44, pp. 945–951, 1995.

[5] W. Rao, K. Yuan, Y. Guo and Ch. Yang, "A simple constant modulus algorithm for blind equalization suitable for 16-QAM signal" , the 9th international conference on signal processing, vol, 2, pp. 1963-1966, 2008, Print ISBN: 978-1-4244-2178-7, INSPEC Accession Number: 10411271.

[6] R. Martinek, M. Al-Wohaishi, J. Zidek, "Software based flexible measuring systems for analysis of digitally modulated systems". RoEduNet 2010 9th, Sibiu, Romania 2010, Page(s): 397 – 402, ISSN 2068-1038, ISBN 978-1-4244-7335-9.

[7] MARTINEK, Radek; ZIDEK, Jan. Application of synthetic instrumentation that applies the trend of software-based approach for measuring on the field of modern wireless transfer systems. In International Journal of Digital Information and Wireless Communications (IJDIWC), Vol. 2, No. 3. Published by The Society of Digital Information and Wireless Communications, Kowloon, Hong Kong, Dec. 2012. pp. 208-221, ISSN: 2225-658X.

[8] MARTINEK, Radek; ZIDEK, Jan; TOMALA, Karel. BER Measurement in Software Defined Radio Systems In Journal: Przeglad Elektrotchniczny (Electrical Review), R. 89 NR 2b/2013, Warszawa, Poland, February 2013, pp. 205-210, ISSN 0033-2097.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Performance of Visual Descriptors in Object Recognition

*Slavomir Matuska, *Robert Hudec, *Miroslav Benco
*University of Transport, Faculty of Computer Engineering, Department of Multimedia, Univerzitna 2, 01026 Žilina, Slovak Republic, {slavomir.matuska, robert.hudec,miroslav.benco}@fel.uniza.sk

Abstract. In general, object recognition consists of three steps – feature extraction from training database, training classifier and evaluate test image. Visual descriptors are used to describe local features and capture image characteristics. This presented paper proposes about local visual descriptors and their performance in object recognition.

Keywords: SURF, SIFT, FAST, STAR, MSER, GFTT, object recognition.

1. Introduction

In computer vision, based idea of object recognition is creation a representation of particular classes, which characterized appearance of object from given class and in this way determine unknown objects class. Success rate of object recognition depends especially on good object representation. Moreover, object representation depends on good object characterization, which can be achieved by visual descriptors. Therefore, this paper proposes some different visual descriptors and their performance in object recognition [1, 2].

2. Object Recognition Process

Object recognition process is shown in Fig. 1a and can be divided into two parts: Training and Testing. Task of training part is to create a classification model from the training data. Training data contain a collection of images of each class. The extraction of primary images features are extracted at their low-level by different methods. Most common used methods are SIFT (Scale Invariant Feature Transform), SURF (Speeded Up Robust Features), MSER (Maximally Stable Extermal Regions), GFTT (Good Features to Track), FAST (Features from Accelerated Segment Test) etc. These methods will be detailed described in section 3. Moreover, a low-level features extracted from images are used to creation a classification model.

To the input of the testing part enters an images and their still picture objects designated to the classification. Moreover, these objects have the same metadata description like data in training part. Based on these data, the classifier is able to regarding to classification model successfully evaluate an unknown objects to the appropriate class [1, 2].

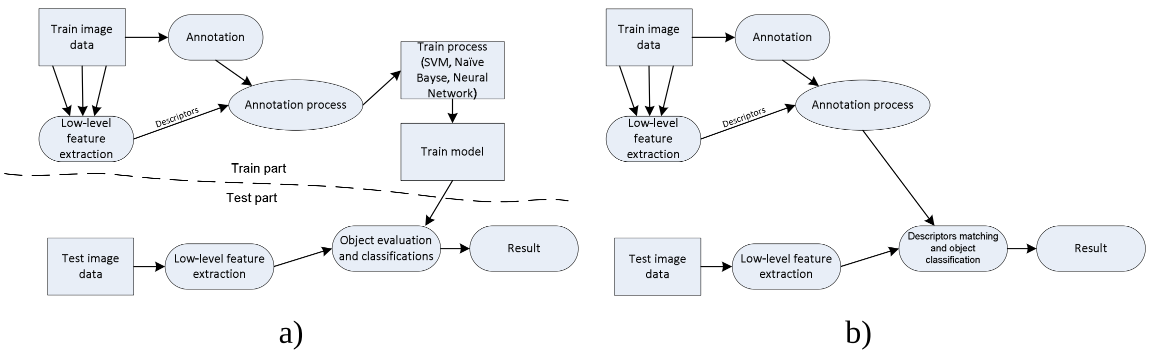


Fig.1 Principle of: a) object recognition process, b) test process.

3. Visual Descriptors

Visual descriptors are used to capture the local appearance of objects. They are calculated from the neighbor pixels. Visual descriptors need to be discriminative enough to distinguish a large number of object classes. Some of them are visually similar and they need to have also invariance to noise, changes of illumination and viewpoints [3].

3.1. SIFT

SIFT is the most widely used local visual descriptors. It has reasonable invariance to changes in illumination, rotation, scaling, and small changes in viewpoints. SIFT involves two stages: features extraction and description. First, the difference of Gaussians operator is applied to an image to identify features of potential interest – key point. Then a SIFT descriptor of key point is obtained by first computing the gradient magnitudes and orientations of pixels in the neighborhood region of the key point, using the scale of the key point to select proper Gaussian kernel to blur the image. The orientation of histograms within the sub-regions around the key point are computed and combined into SIFT feature vector. Produced vector is normalized to improve the invariance to changes of illumination. More detailed information about SIFT can be found in [3, 4, 10].

3.2. SURF

The SURF detector is based on the determinant of the Hessian matrix $H(\chi,\sigma)$, shown in (1).

$$H(\chi,\sigma)=\begin{bmatrix} \dfrac{L_{xx}(\chi,\sigma)}{L_{xy}(\chi,\sigma)} & \dfrac{L_{xy}(\chi,\sigma)}{L_{yy}(\chi,\sigma)} \end{bmatrix}, \tag{1}$$

where $L_{xx}(\chi,\sigma)$ is the convolution of the Gaussian second order derivation of image at point $\chi(x,y)$ in scale σ and similarly for $L_{xy}(\chi,\sigma)$ and $L_{yy}(\chi,\sigma)$.

The discriminant value is used to classify the maximum and minimum of the function by second order derivative test. The first step in the SURF descriptor extracting is construction a square window around the interest point. This window contains the pixels which will form entries in the descriptor vector. Furthermore, the window is oriented along the reproducible orientation. The descriptor window is divided into 4 x 4 regular sub-regions. Within each of these sub-regions Haar wavelets of size 2σ are calculated for 25 regularly distributed sample points. Therefore each sub-region contributes four values to the descriptor vector leading to an overall vector of length 4 x 4 x 4 = 64. Results of SURF descriptor are invariant to rotation, scale, brightness and after reduction to unit length or contrast. More detailed information about SURF can be found in [3, 5, 6, 10].

3.3. Other Keypoints Detectors

MSER – instead of detecting key points, MSER detect a regions which are darker or brighter than backgrounds. It is affinely-invariant and robust to change of illuminations [3, 11].

FAST – is high-speed corner detector, which use corner to capture object appearance. This algorithm was evolved for real-time frame-rated application [7].

STAR – keypoints detector is derived from CenSurE (Center Surrounded Extrema) detector for real time feature detection and matching [8].

GFTT – finds the most prominent corners in the image or in the specified image region [9].

4. Experimental Results

Training database consist of 5 classes: dinosaur, dolphin, butterfly, ant and elephant. The examples of images from training database are shown in Fig 2. Ten images per class were randomly chosen from training database and were used as test database. Tested algorithm is shown in Fig 1b.

The first step of the test process, the low-level features from training images was extracted. In next step, the extracted descriptors together with annotation record in order to create a representation of particular class were used. Thus, created data representation consists of extracted descriptor data just test database only.

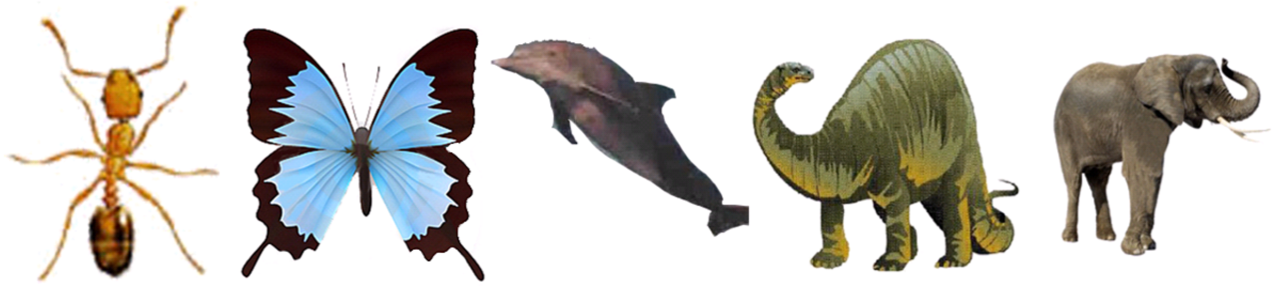


Fig. 2. The images from training database.

As it can be seen in Fig. 1, the absence of learning algorithm is main difference between object recognition and test processes. Therefore, two algorithms for matching descriptors to determine result class for unknown objects were used. For each feature data extracted from test image by selected descriptor, BruteForce matcher finds a closest feature vector extracted from images located in the training database. To the designation of two feature vectors distance, the Euclidean distance was used. Similar approach how to find out the minimum distance of two feature vectors is called FlannBased matcher. It trains a flann index on training feature vectors collection and calls its nearest search methods to find best matches. Good matches were counting per each class and class with maximal good matches was identified as appropriate class. Match is labeled as good match, when (2) is valid:

$$match\ .\ distance < 2 * \min\_dist\ , \qquad (2)$$

where *min_dist* is minimal distance from distance collection calculated by matcher between test feature vectors and training feature vectors. *Match.distance* is distance between actual feature vector and the closest training feature vector.

In experiments a total 6 keypoints detectors, namely, SURF, SIFT, STAR, FAST, MSER and GFTT were used. Moreover, to describing a keypoint by two descriptors: SIFT or SURF and two metrics/matchers: Brute Force or FlannBased were used too. All combinations of detectors, descriptors and matchers were combined into 24 standalone runs and they were programmed in C++ language with support of OpenCV (Open source Computer Vision) library. In Table 1, a success rates for particular classes for SIFT descriptor in combination with SIFT, SURF, FAST or MSER detectors and FlannBased matcher are shown. There is also shown the average scores and computation times, achieved for both, training and evaluation process part related to set of 50 database images per class. From the experiment is evident that combination of SURF detector with SIFT descriptor and FlannBased metric offer almost highest score for acceptable computation time.

Class	Parameters: Detector, Descriptor, Matcher			
	SIFT, SIFT, FlannBased Success rate [%]	SURF, SIFT, FlannBased Success rate [%]	FAST, SIFT, FlannBased Success rate [%]	MSER, SIFT, FlannBased Success rate [%]
Dinosaur	20	**90**	80	50
Dolphin	60	**80**	60	70
Butterfly	70	80	80	**90**
Ant	**90**	70	**90**	20
Elephant	90	80	**100**	**100**
Average score	66	80	**82**	66
Computation time (s)	56.6	245.39	882.08	52.34

Tab. 1. Results for SIFT descriptor, FlannBase matcher and four different detectors.

In Table 2, a success rates with spending computation times for combinations of various detectors a descriptors with BruteForce and FlannBased matchers are presented. Regarding to realized experiments, FAST descriptor with SIFT detector and both matchers produce a best results although for a longest computation time.

Matcher	Detector	Descriptor					
		SIFT	SURF	FAST	STAR	MSER	GFTT
BruteForce	SIFT	64%, 27 s	78%, 280 s	**82%, 473 s**	46%, 10 s	66 %, 28 s	52%, 119 s
	SURF	16%, 23 s	10%, 53 s	10%, 280 s	24%, 12 s	16%, 46 s	16%, 80 s
FlannBased	SIFT	66%, 57 s	80%, 245 s	**82%, 882 s**	46%, 31 s	66 %, 52 s	52%, 338 s
	SURF	16%, 56 s	8%, 226 s	10%, 913 s	24%, 30 s	13%, 43 s	16%, 347 s

Tab. 2. Average score for combination detector, descriptor and BruteForce or FlannBased matchers

5. Conclusion

In this paper, the performance of keypoint detectors and visual descriptors were tested. From the realized experiment is evident that highest classification success rate was achieved by algorithm in combination FAST detector/SIFT descriptor with BruteForce or FlannBased matchers. Moreover, success rate higher than 80 % was achieved by combination SURF detector, SIFT and FlannBased too. All descriptors with SURF detector achieved poor results with success rate of classification under 25 %.

Acknowledgement

The work presented in the paper has been supported by the Slovak Science project Grant Agency, Project No. 1/0705/13 "Image elements classification for semantic image description" and EUREKA project no. E! 6752 – DETECTGAME: R&D for Integrated Artificial Intelligent System for Detecting the Wildlife Migration.

References

[1] SONKA, M, HLAVAC, V., BOYLE, R., *Image Processing. Analysis. and Machine Vision*. Thomson Learning, part of the Thompson Corporation, ISBN: 10: 0-495-24438-4, 2008.

[2] CAO, T.P., *Object Recognition*. ISBN 978-953-307-222-7 2011.

[3] NGI, NGAN K. HONGLIANG, L., *Video Segmentation and Its Applications,* Springer Science+Business Media 2011, ISBN 978-1441994813.

[4] NIXON, M. & AGUANDO, A., *Feature extraction & image processing, second edition*, ISBN 978-0-1237-2538-7, 2008.

[5] BAY, H., TUYTELAARS, T., AND LUC, V. G.: SURF, *Speeded Up Robust Features,* ETH Zurich, Preprint submitted to Elsevier, 10 September 2008.

[6] EVANS, CH. *Notes on the OpenSURF Library*, January 18, 2009.

[7] ROSTEN E., DRUMMOND T., *Machine learning for high-speed corner detection,* Computer Vision – ECCV 2006, Lecture Notes in Computer Science Volume 3951, 2006, pp 430-443.

[8] AGRAWAL, M. and KONOLIGE, K. and BLAS, M.R., “*CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching*”, ECCV08, 2008.

[9] SHI, TOMASI C., *Good Features to Track*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 593-600, June 1994.

[10] ZACHARIASOVA, M., HUDEC, R., BENCO, M., KAMENCAY, P., "*Automatic extraction of non-textual information in web document and their classification,*" Telecommunications and Signal Processing (TSP), 2012 35th International Conference on , vol., no., pp.753,757, 3-4 July 2012.

[11] MATAS, J., CHUM, O., URBAN, M., PAJDLA, T., Robust wide baseline stereo from maximally stable extremal regions. In Proc. British Machine Vision Conference, 2002.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Ubiquitous Localization via Android Smartphone

*Michal Mlynka, *Peter Brida, *Jozef Benikovsky
* University of Zilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitna 8215/1, 010 26 Zilina, Slovakia,
{michal.mlynka, peter.brida, jozef.benikovsky}@fel.uniza.sk

Abstract. Development of ubiquitous positioning has become interesting topic in last years. In this paper, implementation of location based application for Android smartphone is proposed. This application is based on a system which is able to provide ubiquitous positioning. This system is designed as modular localization system. Android smartphones have been chosen because they are the world most popular pocket devices. Most of them are equipped with all necessary chipsets (GSM, Wi-Fi, and Global Positioning System (GPS)). Positioning information based on radio information from these chipsets is used by the system to estimate the device position. Ways how to obtain the radio information from smartphone to modular system, its architecture and final application solution are presented in the paper. Designed application allows estimating the position of the mobile device by the system. It can also be used to scan and analyze GSM and Wi-Fi radio networks and create radio maps for fingerprinting localization methods.

Keywords: Android, fingerprinting, GSM, localization, modular, Wi-Fi.

1. Introduction

Possibility to estimate the position of a device is very important. There are many reasons why people need this information. The most common reason is associated with navigation of people or things in unfamiliar surroundings. There are many applications which help to find a way to unknown destination or to the nearest PoI (Point of Interest) e.g. cash machine, parking, pump or pub [1] - [3]. Aim of these applications is to provide Location Based Services (LBSs) with added value. It is important that applications have been targeted into small pocket-sized device which offer portability. The most widely used mobile devices in the world are smartphones with Android platform. These devices can usually determine their position using GPS or based on information of radio communication networks (e.g. GSM, Wi-Fi etc.).

Unfortunately, particular solutions are optimal for use in different conditions, i.e. one positioning solution is not able to localize a mobile device everywhere and all the time. These reasons were motivation to implement a positioning system which is able to localize a mobile device in all kinds of environments without necessary change of a system or an application by the user. For instance, the user uses only one device and one positioning application for the localization. The system can be implemented as modular localization system which consists of mentioned positioning solutions. According to actual conditions, the system is deciding for selection of optimal positioning module. We present the location based application as an important part of modular localization system. The application is designed for Android smartphones. The application was created on the basis of the previous research described in [4].

The paper is structured as follows. Section 2 describes the Android Operating System (OS). Section 3 contains information about the implemented modular localization system. Software solution for a mobile device is described in the Section 4. The last Section concludes the paper.

2. Mobile Platform Android

Android is an open-source operating system based on Linux kernel and optimized for mobile devices. This OS is currently the world's most popular mobile platform. Android has been created by the Open Handset Alliance [5]. Its applications have been usually developed in object-oriented language Java but there are also other possibilities e.g. Flash or C/C++ [6], [7]. The source code does not compile to native code, but in the intermediate stage called byte-code that is not dependent on a specific hardware. Linux Kernel is an abstraction layer between the hardware and the rest of the software. It provides services as security, memory management, process management, network stack, and driver model. Over Linux Kernel is a set of C/C++ libraries used by various components of the Android system. These capabilities are exposed to developers through the Android Application Framework. Application Framework allows free access to features and hardware options for a developer (e.g. Java class as WifiManager, LocalizationManager and TelephonyManager). Android Runtime layer includes Dalvik Virtual machine code and core Java Libraries. This virtual machine executes Dalvik byte-code, which is compiled from programs written in the Java. Dalvik Virtual Machine is not a Java Virtual Machine. Application block consists of users' applications [8] - [10].

3. Modular Localization System

Created modular localization system is designed to maximize the coverage and increase the reliability of a localization based services [4]. It can be achieved by utilization of the different radio networks which work with the different radio frequency. Almost all Android smartphones are equipped by GSM (900 MHz), GPS and Wi-Fi (2.4 GHz) chipsets. Information from these networks is used by modular system.

Logical model of the system is depicted in Fig.1. The lowermost layer consists of the individual localization modules. Each module estimates mobile device position independently from the particular data. Wi-Fi and GSM modules estimate the position by fingerprinting method utilizing RSS (Received Signal Strength) information [11], [12]. Google Module determines location based on availability of cell base stations and Wi-Fi AP (Access Points) [9], [10].

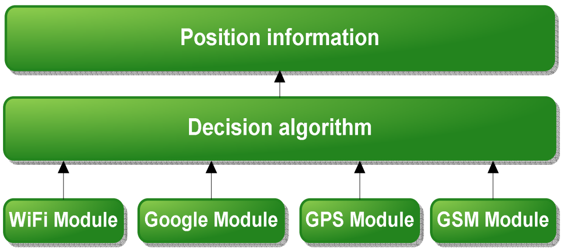


Fig.1. Logical model of Modular location system.

The second layer includes the decision algorithm. The task of this layer is to determine which position estimate will be provided to the user. This decision is based on the predetermined criteria. The top layer ensures view of the estimated position of user on a map. As maps can be used OpenStreetMap [2], Google Maps or custom maps for indoor environment [10]. There can be also shown radio signals parameters and position information obtained by Decision algorithm layer. This information can be used for environment analysis (e.g. how many useable APs or BTSs (Base Transceiver Stations) is there or the GPS accuracy) [4], [8] – [12].

4. Location Based Application

The application is an interface between a user and the modular system. Decision algorithm is also implemented in the application. The application was developed through the Android software development kit (SDK). Important information for positioning is got from Application Framework. Positions obtained by GPS and Google modules are given by LocationManager.

```
LocationManager locationManager = (LocationManager)
                          this.getSystemService(Context.LOCATION_SERVICE);

locationManager.requestLocationUpdates(String provider, long minTime,
                               float minDistance, LocationListener listener)
```

Parameters used by the requestLocationUpdates method are shown in Tab.1

provider	the name of the provider with which to register (GPS or Network)
minTime	minimum time interval between location updates, in milliseconds
minDistance	minimum distance between location updates, in meters
listener	a `LocationListener` whose `onLocationChanged(Location)` method will be called for each location update

Tab. 1. Parameters of requestLocationUpdates method.

Position obtained by WiFi module depends on IEEE 802.11 b/g radio signal status. Information about the signal status are obtained as

```
List<ScanResult> results;
WifiManager wifi;
wifi = (WifiManager) getSystemService(Context.WIFI_SERVICE);
wifi.startScan();
results = wifi.getScanResults();
for (int j = 0; j < results.size(); j++) {
results.get(j).BSSID;
results.get(j).SSID;
results.get(j).level;
}
```

where BSSID is physical address of the access point, SSID is the network name and level is the signal level in dBm. Position obtained by GSM module depends on GSM radio signal status. Information about this signal status id obtained as

```
TelephonyManager tm = (TelephonyManager)
                            getSystemService(Context.TELEPHONY_SERVICE);
     for (int i = 0; i < tm.getNeighboringCellInfo().size(); i++){
tm.getNeighboringCellInfo().get(i).getCid();
tm.getNeighboringCellInfo().get(i).getRssi();
tm.getNeighboringCellInfo().get(i).getLac()));
   }
```

where getCid() gives information about cell id in GSM, getRssi gives received signal strength from this cell and getLac() gives information about location area code [8] – [10].

Obtained information is sent to the Localization server. The server determines device positions from the received Wi-Fi and GSM data. The positions are estimated by fingerprinting localization method [11]. Intensity of radio scans can be setup by a user.

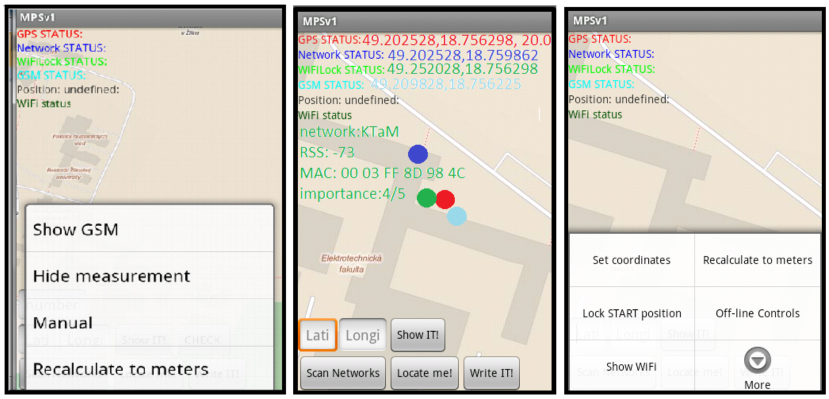


Fig.2 The screen of developed application.

Information about obtained positions is shown in top left corner as is showed in Fig.2. It is also possible to monitor information about the current GSM and Wi-Fi radio signal. The application can offer multiple position estimates. This feature allows investigate the accuracy of the individual modules in different environments. It is mainly implemented for research reason. The red color shows the position estimate obtained by GPS, blue by Google network provider, green by Wi-Fi, cyan by GSM.

5. Conclusion and Future Work

In this paper, implementation of the location based application based on Modular Localization System to an Android smartphone was discussed. Implementation was realized by Java programming language in Eclipse development environment. Google and MapQuest libraries were used for our project. Smartphone features were used by the Application Framework. Information achieved by WifiManager and TelephonyManager are used to determine the position estimates by the modules. These modules extend the capabilities of positioning and offer greater coverage for LBS.

Acknowledgement

This work has been supported by the Slovak VEGA grant agency, Project No. 1/0394/13 and by

Centre of excellence for systems and services of intelligent transport II., ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

"Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ"

References

[1] Google, *Google Maps for mobile*, http://www.google.com/mobile/maps/, [Online] Google, February 2013.

[2] MAPQUEST, *Open Directions Service Developer's Guide*, http://open.mapquestapi.com/directions, [Online] MAPQUEST, February 2013.

[3] APPLE, *Apple Maps*, http://www.apple.com/ios/maps/, [Online] Apple, February 2013.

[4] BENIKOVSKY, J., BRIDA, P., MACHAJ, J., Proposal of User Adaptive *Modular Localization System for Ubiquitous Positioning*, ACIIDS 2012, 2012.

[5] Open Handset Alliance, http://www.openhandsetalliance.com/, [Online] Open Handset Alliance, February 2013.

[6] Adobe AIR, *Adobe AIR for Android*, http://www.adobe.com/inspire-archive/august2010/articles/article1/, [Online] Adobe, 2013.

[7] Google, *Android NDK*, http://developer.android.com/tools/sdk/ndk/index.html, [Online] Google, 20 February 13.

[8] MURPHY, L., *Beginning Android 2* 2nd ed. Computer Press, 2011.

[9] CONDER, S., DARCEY, L., *Android Wireless Application Development*, Addison-Wesley Professional, 2009.

[10] Google, *Android – Developers*, http://developer.android.com/index.html, [Online] Google, February 2013.

[11] BAHL, P., PADMANABHAN, V. N., *RADAR: An In–Building RF–based User Location and Tracking System*, Proceedings of the IEEE Infocom, 2000.

[12] REDDY, S., BAGARIA, A., AGGARWAL, D., ARORA, N., *Analysis of Fingerprint Data in Cellular Networks: An Android Application Case Study*, IPCS, 2012.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Novel Parameter-Based Model Estimating Quality of Synthesized Speech Transmitted over IP Network Based on Different RNN Architectures

*Miroslava Mrvová
*University of Žilina, Department of Telecommunications and Multimedia, FEE, Univerzitná 1, 01026 Žilina, Slovak Republic, {miroslava.mrvova}@fel.uniza.sk

Abstract. This paper deals with a design of parameter-based speech quality estimation model based on Random Neural Network (RNN). In particular, the model has been designed to estimate a quality of synthesized speech transmitted through IP channel, considering an appropriate set of quality-affecting parameters. In this study, four different RNN architectures have been investigated. The performance results obtained by the designed estimation model have proven several facts. Firstly, the mathematical advantage of RNN model (good generalization of new inputs not considered in training phase) has been proven in this study. Secondly, more complex structures of RNN model provide more accurate results, as clearly expected.

Keywords: Random neural network, speech quality estimation, synthesized speech, packet loss.

1. Introduction

Up to the present, many methods for assessing a quality of speech samples transmitted through the telecommunication channel have been developed. The first way, how to assess speech quality is to have wide group of human subjects, which participate in a quality assessment process following common rules recommended in Rec. ITU-T P.800 [1]. This subjective testing is impractical in real conditions, because of reserved opinion of every subject on quality of samples, high costs, time-consumption, etc. In order to avoid/overcome the drawbacks of the subjective testing, objective testing has been developed. Objective methods are able to automate estimation process by means of computer programs modeling average user behavior [2]. Finally, the other way of assessing the speech quality (presented in this paper) represent models based on machine learning techniques (e.g. neural networks, genetic programming). This approach is able to offer real-time and continuous evaluation of speech quality in comparison with objective testing. In regard of the fact that terminal point is satisfaction of the end user, a common goal of last two methods is to provide results highly correlated with subjects opinions obtained from subjective testing [2].

The interest in synthesized speech has grown extremely in recent years and synthesized speech has become a part of modern daily life. In addition, some voice-based telecommunication services have been replaced by services based on synthesized speech. To date, no one model designed for estimating a speech quality of the services based on synthesized speech has been evolved and standardized. Such model could be helpful for network operators and service providers in planning phase or early-development stage of telecommunication services based on synthesized speech. In this paper, I present the novel parameter-based model estimating a quality of synthesized speech transmitted through IP channel based on Random Neural Network (RNN). The main advantage of using RNN for this purpose is according to the literature (for instance [3]) that, it provides accurate estimations, good computational efficiency and has also good ability to generalize for new inputs.

2. Experiment Description

Generally, the RNN is a learning tool [4] serving to map the relationship between input and output parameters and generalize it for new inputs. In other words, the idea is to apply an appropriate set of quality-affecting parameters as an input together with their corresponding MOS (Mean Opinion Score) values (values expressing speech quality perceived by the user ranging from 1 "bad quality" to 5 "excellent quality") in order to define a relationship between them. One of the quality-affecting parameters used as an input of the designed estimation model was packet loss, because of its significant impact on speech quality proven by several studies [5-7]. In principle, following parameters have been used to properly capture packet loss impact in the designed model:

- conditional loss probability (as indicator of packet loss burstiness, denoted as clp) with values: 0% (theoretically) - Bernoulli loss model (modeling independent/random losses), 70% and 80% - Gilbert loss model (modeling dependent losses)
- unconditional loss probability (as indicator of packet loss rate, denoted as ulp) with values: 0%, 1,5%, 3%, 5%, 10% and 15%.

The simulated clp and ulp values reflect packet loss conditions, which can be commonly obtained in real telecommunication networks. The third considered input parameter was related to a type of synthesized speech signal. There were used two kinds of Slovak synthesizers, namely Diphone synthesizer (Kempelen 1.6 [8]) and Unit-selection synthesizer (Kempelen 2.1 [8]). Diphone and Unit-selection synthesis are corpus-oriented types of synthesis, which means they use finite number of speech (acoustic) units (diphones, syllables, words, sentences, etc.) stored in a database in order to generate a synthesized version of a desired phrase or sentence.

For the purpose of this research 40 measurements for each value of ulp considering all values of clp and two types of synthesized speech signal were carried out in order to get quality scores for designing the estimation model. These speech quality scores were predicted by PESQ (Perceptual Evaluation of Speech Quality - objective model designed for naturally-produced speech). The decision to use PESQ as a predictor of synthesized speech quality was based on experiments carried out in [6], [9], [10] which have proven that PESQ is able to provide accurate predictions of quality of synthesized speech impaired by the impairments used in this study (with correlation coefficient above 0.9). In total, 1440 synthesized speech samples with length 12 seconds were recorded, hence 36 average speech quality scores (6 ulp x 3 clp x 2 types of synthesized speech signals) called configurations were chosen. These 36 average configurations served to create two training databases. The first one (denoted as AVG1) had knowledge of overall range of configurations during the training process, whereas the second one (denoted as AVG2) did not cover any boundary condition during training process. The synthesized speech samples, used for designing the estimation model were coded by encoding scheme ITU-T G.729AB [11].

2.1. Training Phase

During the training phase, the task of RNN is to capture a numerical or logical relationship between quality-affecting parameters as an input (overall three input parameters used: clp, ulp, type of synthesized speech signal) and corresponding MOS values (predicted by PESQ) as an output and generalize it for new inputs. Gradient descent method was used as a learning algorithm, because of its simplicity and strong generalization capability even for small training data sets. Four different architectures were used in a design process of the estimation model, namely 2-layer RNN, 3-layer RNN (with 7 and 10 hidden neurons) and 4-layer RNN model (with two hidden layers).

As mentioned above, totally 36 configurations were used; 30 of them were employed for training process (to acquire the relation between input and output) and the rest, namely 6 configurations (common ratio of 80:20) for testing process to verify a performance of the designed estimation model. RNNSIM v.2 developed by Abdelbaki [12] was used as RNN environment for this research.

2.2. Testing Phase

As already mentioned above, 6 configurations were used to evaluate a performance of the designed model. Those 6 testing configurations were randomly chosen from the 36 configurations of quality-affecting parameters and corresponding MOS values (predicted by PESQ) and naturally were not used in training process as required by common rules in machine learning community. The MOS values provided by the designed estimation model (ranging from 1 “bad quality” to 5 “excellent quality”) were compared with the MOS values (of synthesized speech signals) predicted by PESQ for corresponding configurations. The performance of the designed estimation model was quantified in terms of the Pearson correlation coefficient R, the respective root mean square error (rmse) and epsilon-insensitive root mean square error (rmse*) described in [13].

3. Results

Five simulations were conducted for each network architecture and training database in order to obtain statistically stable and reliable results, i.e. in total 40 simulations (2 training databases * 4 different RNN architectures * 5 simulations per architecture and training database). In Table 1, the best simulations selected from all realized simulations are presented. They were selected according to the root mean square error (rmse) and epsilon-insensitive root mean square error (rmse*). A comparison of real (predicted by PESQ) and estimated MOS values by the designed model for both databases is depicted in Figure 1.

RNN architecture	AVG1			AVG2		
	R	rmse	rmse*	R	rmse	rmse*
3_1	0.0979	0.4567	0.3813	-0.8189	0.9005	0.8313
3_7_1	0.7811	0.2588	0.2091	0.7210	0.3844	0.3261
3_10_1	0.7837	0.2591	0.2107	0.7218	0.3843	0.3254
3_5_4_1	0.7759	0.2594	0.2084	0.8020	0.3500	0.3070

Tab. 1. Chosen best simulations in regard of rmse and rmse*.

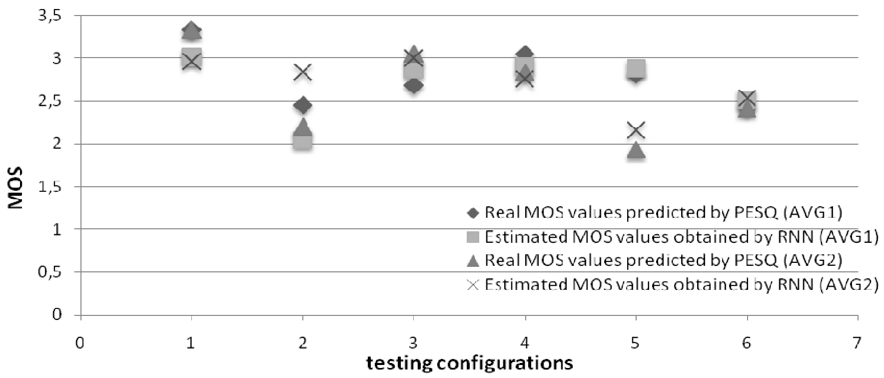


Fig. 1. The real and estimated values of speech quality (4-layer architectures (3_5_4_1) for both databases).

From the Table 1, we can clearly see that more complex RNN architectures (with one or more hidden layers) yield to provide more reliable and accurate predictions than simple RNN architectures (without hidden layer). This fact has also been proven in [7]. It is worth noting that AVG1 database had knowledge of overall range of configurations during the training process, whereas AVG2 database had not covered any boundary condition during training process. On the other hand, a performance values obtained for database AVG2 show that the values of Pearson correlation coefficients (R) are slightly lower (apart from 4-layer RNN model denoted as 3_5_4_1 which is surprisingly higher) and the respective root mean square error (rmse) and epsilon-insensitive root mean square error (rmse*) values are substantially higher than those obtained for database AVG1. Anyhow, the reported correlation R, rmse and rmse* values are still satisfactory for this kind of quality estimation model. On the other hand, taking into account the fact that AVG2 database had not covered any boundary condition during the training process, we can pronounce on the basis of relatively small differences obtained between the investigated databases that RNN has good ability to generalize for new inputs. In other words, the curve characterizing relationship between the perceived quality of synthesized speech and quality-affecting parameters involved in the databases is relatively well-defined even beyond the specified interval; therefore designed model is able to provide good results even outside the area defined in the training process.

4. Conclusion

The main task of this paper was to investigate the performance of the designed speech quality estimation model based on different RNN architectures. In particular, the designed RNN model has estimated the quality of synthesized speech (as perceived by the end user) transmitted through IP channel, taking into account quality-affecting parameters, like packet loss parameters and type of synthesized speech signal. The designed estimation model has achieved high correlation and small rmse and rmse* values in comparison with the MOS scores predicted by PESQ model, especially for more complex architectures. Moreover, the designed model has confirmed good estimation accuracy for very limited training database (30 configurations involved in training database) and good ability of generalization for new inputs not involved in training database. On the basis of the obtained performance values, we can conclude that the designed estimation model is suitable to accurately estimate the quality of synthesized speech transmitted over the IP network. Furthermore, the designed estimation model is able to provide estimations of quality of synthesized speech transmitted over IP network automatically and in a real-time. For this reason, the model can be useful for network operators and service providers.

References

[1] ITU-T P.800: *Methods for subjective determination of transmission quality*. International Telecommunication Union, Geneva,1996.

[2] De RANGO, F., TROPEA, M., FAZIO, P., MARANO, S. *Overview on VoIP: Subjective and Objective Measurement Methods*. IJCSNS International Journal of Computer Science and Network Security, Vol.6, No.1B, January 2006.

[3] MOHAMED, S., RUBINO, G. *A Study of Real –Time Packet Video Quality Using Random Neural Networks*. IEEE Transactions on Circuits and Systems for Video Technology, Vol.12, No.12, December 2002, pp. 1071-1083.

[4] GELENBE, E. *Random Neural Networks with Negative and Positive Signals and Product Form Solution.* Neural Computation 1 (4), 1989, pp. 502-510.

[5] MOHAMED, S., RUBINO, G, VARELA, M. *Performance Evaluation of Real-time Speech through a Packet Network: a Random Neural Networks based Approach*. Performance Evaluation 57 (2), 2004, pp. 141-162.

[6] RADHAKRISHNAN, K., LARIJANI, H. *Evaluating Perceived Voice Quality on Packet Networks using Different Random Neural Network Architectures*. Performance Evaluation 68, 2011, pp. 347-360.

[7] DA SILVA, A. P., VARELA, M., DE SOUZA E SILVA, E., LEAO, R., RUBINO, G. *Quality Assessment of Interactive Voice Applications*. Computer Networks 52 (6), 2008, pp. 1179-1192.

[8] DARJAA, S., RUSKO, M., TRNKA, M. *Three generations of speech synthesis systems in Slovakia*. In Proceedings of XI International Conference Speech and Computer (SPECOM 2006), Sankt Petersburg (Russia), pp. 297-302, 2006, ISBN 5-7452-0074-X.

[9] POČTA, P., HOLUB, J. *Predicting the Quality of Synthesized and Natural Speech Impaired by Packet Loss and Coding Using PESQ and P.563 Models*. In Acta Acustica united with Acustica, vol. 97, No. 5, pp. 852-868, 2011, ISSN 1610-1928.

[10] BASTERRECH, S., RUBINO, G., VARELA, M. *Single-sided Real-time PESQ Score Estimation*. In Proceedings of Measurement of Speech, Audio, and Video Quality in Networks (MESAQIN'09), Prague, Czech Republic, 2009.

[11] ITU-T Rec. G.729: *Coding of speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Exited Linear prediction (CS-ACELP)*. International Telecommunication Union, Geneva (Switzerland), 2007.

[12] ABDELBAKI, H. E. *Random Neural Network Simulator (RNNSIM v.2).* For Use with MATLAB, September 1999, online: http://www.cs.ucf.edu/~ahossam/rnnsimv2/rnnsimv2.pdf

[13] ITU-T Rec. P.1401: *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*, International Telecommunication Union, Geneva (Switzerland), 2012.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Risk assessment of safety signalling systems based on individual risk

*Ľubomír Pekár
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 2, 010 26 Žilina, Slovak Republic, lubomir.pekar@fel.uniza.sk

Abstract. Currently an increased attention to safety is required in the design of new safety signalling systems and safety controlled processes as well. The risk is closely related to the safety and an objective human safety can be only achieved by the objective risk analysis and the objective safety evaluation of safety signalling systems. This article focuses on an overview of frequently used ways and methods of risk assessment and risk calculation (estimation) in praxis. In the article the analysis of the individual in the process control and the impact of controlled process on an individual are discussed.

Keywords: individual risk, risk calculation (estimation) and risk assessment, THR, SIL, safety signalling system.

1. Introduction

Currently, in the discipline of process control the different control systems that perform different functions are used. The control systems related to the safety belong to one of the subsets of control systems. They are called the safety related control systems (SRCS) and are used in various fields of life. One of these fields is the field of railway traffic, where the SRCS are called the safety signalling systems. They are designed to perform the safety functions necessary to achieve or to maintain a safe state (a state in which safety is achieved) of controlled process. Such control systems have a so-called fail-safe feature.

The risk arising from the undesired events is closely related to the safety of railway traffic. In our case, the undesired events can be the formation of hazard or an accident, etc. The formation of such events and their consequences can endanger a human as an individual, or as a group of human respectively, material properties, environment and so on. In the paper, the object of our interest is the impact of the controlled process on human as an individual.

2. Human in Process Control

In general, in the safety control discipline, the relationship between the human, safety signalling system (SRCS) and controlled process can be considered. Since the safety signalling system is always required to be reliable and safe. Reliability and safety of the system can very often depend on the human factor and also on the many other factors. Therefore it is quite important and necessary to pay an increased attention to the relationship between the human and the controlled process.

If we consider the relationship between the human and controlled process (Fig. 1.), it is also necessary to deal with the following issues:

1. How can the human impact the controlled process, respectively how can he intervene in the process control?
2. How can the controlled process impact the human in the case of formation of the hazard and what consequences will it have?

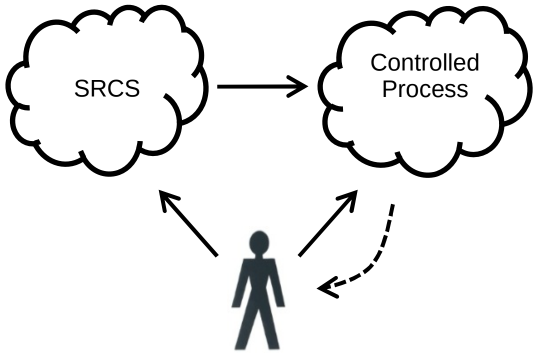


Fig. 1. Relationship Human – controlled process

2.1. The impact of controlled process on human after putting SRCS into operation

If we consider a direct relation from the controlled process to human, it is necessary to consider whether the SRCS is put into operation at all, and how can the controlled process endangers the human in the case of formation of the hazard. During these considerations, it is important to ensure that the elements of subjectivity will be eliminated by the process of the risk analysis and safety analysis of SRCS (there are different ways of analyses). Therefore, if we want to objectify human safety, we must make an objective risk analysis. If every risk assessor used another method of the risk analysis (subjective), in the result of different cases there would be different safety features of the SRCS or the design of other safety functions of SRCS could be altered. These safety functions could then have different tolerable hazard rates (THR) and different safety integrity levels (SIL) (in the terms of the SIL defined on the safety function of the SRCS, not on the entire system).

2.2. The position of an individual in regard of controlled process

An individual can come into contact with the process control at different circumstances. In general, the following cases (Fig. 2.) can be considered:

- an individual comes into contact with SRCS in the position of operator, where his orders impact the controlled process (i.e. case a);
- an individual comes into contact with the controlled process in the position of maintenance worker, where he involves in the control during scheduled maintenance (case b);
- an individual comes into contact with the controlled process in the position of passenger, where his life is directly endangered in the case of formation of an undesired event (case c);
- an individual comes into contact with the controlled process in the position of third person, i.e. as an indirect participant of the control. There can be directly endangered his life. For example, it can be the case of participant of the road traffic, respectively the pedestrian, whose life can be endangered by the controlled process (a train) (case d).

The object of our interest is particularly the case of passenger as an individual within the train (case c) and in the case of formation of undesired events his life can be endangered, therefore we consider his individual risk. An undesired event can occur at the railway station or open line section. In our case, we focus on the case of open line section. From the statistics of Slovak Railways (ZSR) it is obvious that the most undesired events occur at the place of the points. There are places, where the railway lines are crossed in the open line section or where the train crosses a station head, i.e. places, where the open line section starts, respectively ends.

If we consider the open line section, besides the crossed points there are also some other parts of controlled process at risk. It can be the cases of riding the train through the level crossing system, railway tunnel, railway bridge, etc., which can endanger the lives of passengers due to an undesired event occurrence. The selected cases are shown in Fig. 2. A similar analysis is suggested in the publication [6], focusing on specific acts of human.

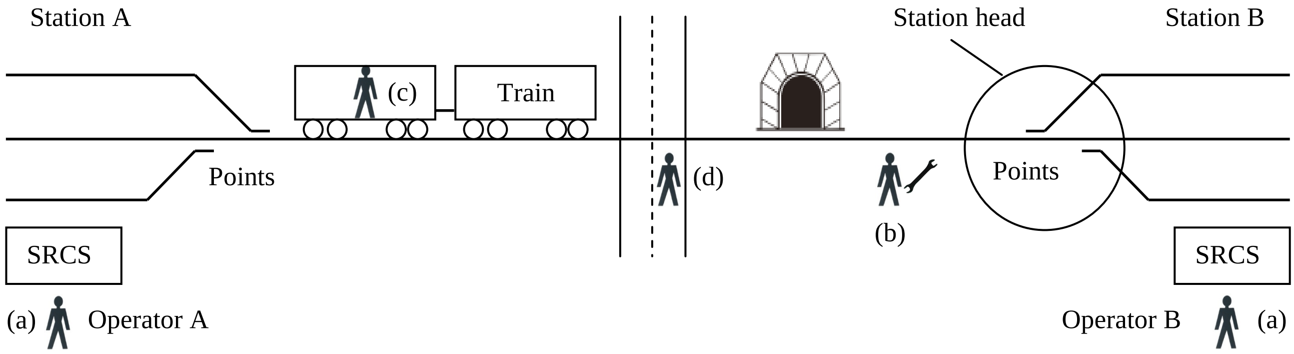


Fig. 2. Human and his risk in process control

2.3. Specification of safety functions of the safety signalling system

The safety function of safety signalling systems is not clearly defined by the standard. In the terms of objectivisation of the safety of safety signalling systems, it is necessary to define THR for safety function of these systems. Currently, in some cases, the determination of THR for safety function can be characterized by the fact that for safety signalling systems that perform a large number of safety functions, the calculated value of THR is high. Conversely, for safety signalling systems that perform a small number of safety functions, the calculated value of THR is low. The safety function of these systems isn't clearly defined in the standards, so it is important to unify this concept within the scope of safety objectivisation of safety signalling systems.

It is therefore particularly important to determine risk for each undesired event (hazard) and to determine the THR and SIL for each defined safety function of safety signalling system through the risk assessment. It is necessary to keep in mind that the value of tolerable risk cannot be exceeded and the process of risk analysis has to be executed objectively through risk determination. In the case of the open line section it is different to determine the risk to all open line sections within the system as a whole (the system has only one safety function) from the determination of the risk to the individual line sections of open line section (the system has more safety functions).

3. Risk calculation (estimation)

Currently we can find various definitions of the risk. According to the standard [1] the risk is defined as a combination of the frequency (intensity) or probability and consequence of a specified hazardous event. In general the risk can be calculated as the combination (product) of the intensity of origin of hazards h and their consequences S.

3.1. Individual risk

Standard [1] defines the individual risk as the risk that is associated with only one individual. It is assumed that the individual isn't protected and he is exposed to adverse circumstances (hazard) at all the time of operation of the safety signalling system.

The most important parameter is the intensity of origin of undesired events (hazards), i.e. the number of hazards for a unit of time (in literature the year is quite often used as the unit of time).

3.2. Tolerable risk

At the risk management of SRCS, it is needed either to tolerate, reduce or delete the risk. There are many measures which can be used to reduce the risk, whether it is an application of organizational or technical measures. It is a process whose result is a risk reduction under the limit of tolerable risk.

The tolerable risk presents a risk level which the individual or company are disposed to tolerate. If the tolerable risk is greater than the total risk, it does not need to perform any interventions within the risk management. In the case the tolerable risk is less than the total risk, the difference of these risks is possible to identify as risk to reduction. After reducing the risk to the desired level, there remains the risk, which is called the residual risk. The relationship between the tolerable risk and risk to reduction is shown in Fig. 3. The residual risk describes the extent of the hazard and it must be less than or equal to the tolerable risk. If the residual risk is equal to the tolerable risk, the residual risk is also called the limit risk.

Generally, it is possible to state that SRCS is considered as safe, when the individual risk is less than or equal to the defined limit value for each individual, who is coming into contact with the system (consciously or also unconsciously) [2].

<table>
<tr><td colspan="3">Risk</td></tr>
<tr><td colspan="3">Risk without safety functions</td></tr>
<tr><td colspan="2">Tolerable risk</td><td>Risk to reduction</td></tr>
<tr><td>Residual risk</td><td colspan="2">reduced risk</td></tr>
<tr><td></td><td>Reduced the risk of technical measure</td><td>Reduced organizational risk measure</td></tr>
</table>

Fig. 3. Illustration of risk reduction

It is necessary to define the safety functions based on the risk analysis, determine the value of SIL, respectively THR for each safety functions of SRCS and design the control system for the analysed controlled process in accordance with these requirements.

3.3. Calculation ways of individual risk

According to [3] the connections that lead to undesired events at the formation of some hazard are taken into account at the determination of the individual risk. These considerations can be summarized as follows:

1. An individual uses the safety signalling system (e.g. a station interlocking system, etc.). The profile of the operations of the safety signalling system is described by the number of the system uses N_i by the individual (per year or per hour).
2. At the using of the safety signalling system, the individual is exposed to hazards due to failure of the system or its subsystem, etc. This is described in the list of hazards with the corresponding intensities of formation of the hazards h_j. The probability that the individual is exposed to the hazard H, is also dependent on the time of hazard duration D_j (the latent time) and also on risk time E_{ij} (the exposure time of the risk), during which the individual is exposed to the consequences of the hazard. This probability is based on the assumption that the hazard already exists, when an individual comes into contact with the safety signalling system (h_j. D_j) and the assumption that the hazard persists during the system use of the individual (h_j. E_{ij}).
3. One or more types of undesired events (accidents) can occur from each hazard. This event is described by the probability of consequences s_j^m that the m-tuple accident of the d_m type occurred for each hazard. This probability is valid for the external factors of the risk reduction that can be derived from the consequence analysis. For each type of accident d_m an appropriate degree of severity f_i^m is derived from the analysis of the damages. It can be assigned to i-tuple individual as the probability of fatality of the individual.

This causality corresponds to the formula of individual risk in case of fatality of the individual (IRF_i), which is defined by the formula [3]:

$$IRF_i = N_i \sum_{hazards\_H_j} \left[\left(h_j \cdot D_j + h_j \cdot E_{ij} \right) \sum_{accidents\_d_m} \left(s_j^m \cdot f_i^m \right) \right]. \qquad (1)$$

The publication [4] provides another approach for calculation of the individual risk. The basis for this calculation IR is the collective risk R_c per unit of time (for example, the number of fatalities per year) which has been calculated as part of a detailed analysis. The traditional approach (based on frequency of use) takes into account the number of people involved every year N_P and the maximum number of occasions at which the system is used per unit of time W_{max} by one of these persons [4]:

$$IR = \frac{R_C}{N_P} \cdot W_{\max} . \tag{2}$$

In order to take into account the different system exposure times, the inclusion of the exposure time is recommended as specified in [4]:

$$IR_{E1} = \frac{R_C}{E_P} \cdot E_1 , \tag{3}$$

where E_1 corresponds to the time during which the individual person is exposed to the system and E_P to the time for which all the persons involved are exposed (cumulative period of exposure). IR_{E1} is referred as the individual risk based on the period of exposure. It's a virtual risk to which a person is exposed in case that the individual uses permanently the system per unit of time.

4. Risk assessment in the process of risk analysis

4.1. The process of risk analysis

According to the standard [5] the definition of the safety requirements for the safety signalling system, but also the subsequent evaluation of the system safety, must be based on the risk analysis. The process of risk analysis consists of the following steps:

- ***Defining the system*** - the goal is to define the system independent on its future technical execution. It is necessary to define the integration of safety signalling system in the operating environment, the system role in the process of the operating control, the interface between the safety signalling system and controlled process, and so on;
- ***Identification of hazards*** - the occurring hazards are identified during the life cycle of the safety signalling system. Hazards identification must be made together with the causal analysis (the analysis of causes, how and why can a particular hazard occur) and estimate intensities of formation of the hazards;
- ***Analysis of the consequences of hazards*** - will contain an estimation of the probable consequences, if the hazard occurs. Identification of the consequences must be carried out for each hazard depending on the specific operational situation;
- ***Calculation (estimate) of total risk*** - the total risk associated with the control of the analysed controlled process is determined based on the information about intensities of the hazards and consequences of the hazards. Risk assessment can be based on the use of the qualitative and quantitative methods.
- ***Determination of tolerable hazard rate (THR)*** - the resulted tolerable hazard rate must be derived in regard to the tolerability criteria (acceptation) of the risks. These criteria are determined by the requirements of the legislative at the national or European level [2].

4.2. Quantitative assessment of individual risk

Individual risk can be assessed through the various qualitative, semi-quantitative or quantitative methods. In terms of human losses THR can be determined based on the following principles using the quantitative methods:

- ***principle ALARP (As Low As Reasonably Practicable)*** - this principle is based on the idea that the control system must provide such safety level that verifies generally acceptable rate of risk, so called acceptable field (as low as reasonably achievable risk). If this rate of risk is not exceeded, we consider the system as safe. There is a certain limit above which the risk is intolerable, so called unacceptable field. In case, if it is impossible to reduce the intolerable risk below the upper limit, the control system will be unusable in operation. The ALARP field is between these two limits. In this field the risk will be accepted only if the costs of the risk reduction are disproportionately high in regard to the achieved improvement, or the risk reduction is impossible. It isn't enough only to show that the risk is in the ALARP field. The risk must be reduced in the way that it is reasonably possible [2];

- ***principle GAMAB (Globalement au moins aussi bon) –*** it tells that all the new control systems of the railway traffic process must provide the overall level of risk at least so good as any of existing equivalent (reference) systems. This principle is based on the statistics of safety of existing reference systems and it allows dividing the total risk associated with the control of railway traffic into the partial risks of subsystems. The subsystems are involved in the controlled process (for example, equipment on the track and in the train), so that the level of overall risk is better (or at least so good) as of the existing reference systems [2];
- ***principle MEM (Minimum Endogenous Mortality) –*** it is based on general statistic of a minimum endogenous mortality of population in the state and it is focused on the age group from 5 to 15 years, for which is the lowest. The minimum endogenous mortality R_m was defined as $R_m = 2.10^{-4}$ [fatalities / person × year]. It is required that the hazard caused by the new system would not produce a significant increase of R_m. According to the principle MEM it is possible to use the value of tolerable risk of fatality of an individual $R_{TO} \leq 1.10^{-5}$ [fatalities × year] for determining the tolerable hazard rate [2].

5. Conclusion

Based on the available used methods I would like to focus on the analysis and evaluation of the individual risk of the passenger by train in the open line section in my future work. I would like to focus on the analysis of origin of the various hazards that can occur in open line section, whether at the crossing of the points, in the railway tunnels, or at level crossings. In future work, I would like to focus mainly on the critical evaluation of existing methods and their modification, respectively creating custom quantitative methods of analysis of individual risk.

Acknowledgement

This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0388/12 "Quantitative safety integrity level evaluation of control systems in railway application".

References

[1] STN EN 50 129 *Dráhové aplikácie – Komunikačné a signalizačné systémy a systémy na spracovanie údajov. Elektronické signalizačné systémy súvisiace s bezpečnosťou*, 2004.

[2] Záhradník, J., Rástočný, K., Kunhart, M. *Bezpečnosť železničných zabezpečovacích systémov*, EDIS Žilina, ISBN 80-8070-296-9, 2004.

[3] Braband, J. *Risikoanalysen in der Eisenbahn-Automatisierung,* Siemens AG, Signal – Draht Edition, Hamburg, ISBN 3-7771-0335-7, 2005.

[4] Slovák, R., Meuli, H., Schlatter, H. *Assessing the individual risk of rail transport for passengers and staff,* in the 9th FORMS/FORMAT 2012 Symposium on Formal Methods for Automation and Safety in Railway and Automotive Systems, Braunschweig, Germany, 2012.

[5] STN EN 50 126 *Dráhové aplikácie – Stanovenie a preukázanie bezporuchovosti, pohotovosti, udržiavateľnosti a bezpečnosti (RAMS)*, 2002.

[6] Rástočný, K., Ždánsky, J., Nagy, P. *Some specific activities at the railway signalling system development*. in Communications in Computer and Information Science, Volume 329, Telematics in the Transport Environment, Springer-Verlag, Berlin, pp. 372-381, ISSN 1865-0929, ISBN 978-3-642-34049-9, 2012.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

An Application of Selected Theoretical Tools from the Artificial Intelligence Area to the Breast Cancer Recognition Problem

*Jakub Piekoszewski
*Kielce University of Technology, Department of Electrical and Computer Engineering,
Al. 1000-lecia P.P. 7, 25-314 Kielce, Poland, j.piekoszewski@student.tu.kielce.pl

Abstract. This article presents several selected theoretical tools from the field of artificial intelligence and theirs application to the breast cancer classification problem. The considered tools are: feed forward neural network (multilayer perceptron), neural network with the radial basis functions, decision tree, bayesian network, support vector machine, and *k*-nearest neighbour algorithm. A comparative analysis of the considered tools with other alternative techniques has also been presented. All numerical experiments presented in this paper are performed with the use of the *Wisconsin Breast Cancer* dataset, that is available on the WWW server of the University of California at Irvine.

Keywords: Artificial intelligence, Computational intelligence, Decision support systems, Data classifiers, Supervised learning, Breast cancer classification problem.

1. Introduction

Designing techniques of the classification systems are the subject of intensive scientific research since many years. In particular, several theoretical tools from the field of artificial intelligence play an important role in this task. The tools allow us to build such systems, in a certain sense, in an automatically way from data (i.e. from a representative set of examples of a given classification problem). There are two fundamental categories of the classifiers: the so called black-boxes and the grey-boxes - distinguished according to the way of representation of the knowledge about the considered problem. The black boxes contain the knowledge in the form of a set of real numbers. The meanings of particular numbers (and in consequence, the entire knowledge) are very difficult to understand and to interpret by a human (see, e.g. [1], [2], [12], [14]). Usually, the neural networks or decision trees are used in order to build this type of classifiers. The gray-boxes contain the knowledge in the form of decision rules – clear and relatively easy to interpret by a human (see, e.g. [3] – [8], [10], [11], [13]). This time, two main theoretical tools, based on the fuzzy set theory or the rough set theory are employed to create the classifiers. The clear representation of the knowledge is a huge advantage of these tools – unfortunately, it is achieved by a small reduction in their accuracy (in comparison to the black-box-type classifiers).

In this work, the breast cancer diagnosis problem is considered. In order to design the classifiers supporting the diagnosis, several selected classification methods (belonging to the above mentioned two categories) are used. One of the first approaches to this problem was presented in the paper [15]. W. H. Wolberg and O. L. Mangasarian applied the linear programming technique to determine the parameters of the decision hyperplanes, which separate the cases of healthy people from the cases of sick people. For the purpose of the numerical experiments, they developed a special dataset describing the disease of breast cancer, known as *Wisconsin Breast Cancer* dataset. It is available now at the WWW server of the University of California, at Irvine (http://archive.ics.uci.edu/ml). This dataset is commonly used to test of the modern design methods of the classifiers, including the above mentioned techniques from the artificial intelligence area. A great number of papers (about 350 in the Web of Knowledge database) suggest that the considered problem is still open. In this paper, the *Wisconsin Breast Cancer* dataset has also been used to test the proposed classifiers. The following theoretical tools: multilayer perceptron (MLP), neural network with the radial basis functions (RBF), decision tree, Bayesian network, support vector machine (SVM), and *k*-nearest neighbour algorithm (*k*-NN) are applied to design six classifiers. A theoretical overview of these tools can be found in [9]. All numerical experiments presented in this paper have been carried out with the use of WEKA (Waikato Environment for Knowledge Analysis) software, that is available at the WWW server of the University of Waikato in New Zealand (http://www.cs.waikato.ac.nz/ml/weka). Finally, a comparative analysis of the considered classifiers with other alternative approaches has been conducted.

2. The Classifiers Design – an Outline

A classification system with *n* inputs x_1, x_2, …, x_n and *m* outputs y_1, y_2, …, y_m is considered. The number of classes recognized by the system is equal to the number of its outputs. The response of the system on the *j*-th output $y_j \in [0,1]$, $j = 1,2,...,m$ (calculated when the input data are presented to its inputs) represents the degree of membership of input vector for the *j*-th class. The system is designed from data using the supervised learning scenario and the learning dataset $L_1^{(lrn)}$ in the form of p pairs of the input-output data samples:

$$L_1^{(lrn)} = \{\boldsymbol{x}_s'^{(lrn)}, c_s'^{(lrn)}\}_{s=1}^{p}, \tag{1}$$

where $\boldsymbol{x}_s'^{(lrn)} = [x_{s1}'^{(lrn)}, x_{s2}'^{(lrn)}, ..., x_{sn}'^{(lrn)}]$ ($x_{si}'^{(lrn)} \in \Re$, $i = 1,2,...,n$) is a vector of input data (presented to the classifier inputs), $c_s'^{(lrn)}$ is a class label to which the vector $\boldsymbol{x}_s'^{(lrn)}$ is assigned. In the frame of the so-called data preprocessing stage, the original learning dataset $L_1^{(lrn)}$ is transformed into a new numerical learning dataset:

$$L_2^{(lrn)} = \{\boldsymbol{x}_s'^{(lrn)}, \boldsymbol{d}_s'^{(lrn)}\}_{s=1}^{p}, \tag{2}$$

where $\boldsymbol{x}_s'^{(lrn)}$ is the same as in (1), while $\boldsymbol{d}_s'^{(lrn)} = [d_{s1}'^{(lrn)}, d_{s2}'^{(lrn)}, ..., d_{sm}'^{(lrn)}]$ ($d_{sj}^{(lrn)} \in \{0,1\}$, $j = 1,2,...,m$) is a desirable response of the system, when the input vector $\boldsymbol{x}_s'^{(lrn)}$ is presented to the inputs ($d_{sj}^{(lrn)} = 1$ if the vector $\boldsymbol{x}_s'^{(lrn)}$ is assigned to the *j*-th class and $d_{sj}^{(lrn)} = 0$, otherwise).

During the learning process, the response error $Q_{RMSE}^{(lrn)}$ of the system for the learning dataset $L_2^{(lrn)}$ is minimized:

$$Q_{RMSE}^{(lrn)} = \sqrt{\frac{1}{p}\frac{1}{m}\sum_{s=1}^{p}\sum_{j=1}^{m}(y_{sj}'^{(lrn)} - d_{sj}'^{(lrn)})^2}, \tag{3}$$

where $y_{sj}'^{(lrn)} \in [0,1]$ is a real response of the system on the *j*-th output y_j, calculated when the vector $\boldsymbol{x}_s'^{(lrn)}$ is presented to the inputs. In order to control the learning process, the response error $Q_{RMSE}^{(tst)}$ of the system for the test data $L_2^{(tst)}$ is also calculated in an analogous way as in (3).

3. Experiments and Results

In order to test the considered classification techniques we use the *Wisconsin Breast Cancer* dataset, which contains 699 cases of the breast cancer diagnosis. Each case is assigned to one of two possible classes: *benign* (458 instances) or *malignant* (241 instances) and it is described by nine numeric attributes: *clump thickness*, *uniformity of cell size*, *uniformity of cell shape*, *marginal* *adhesion*, *single epithelial cell size*, *bare nuclei*, *bland chromatin*, *normal nucleoli*, and *mitoses*. The attribute values are integer numbers from 1 to 10. The original dataset has been divided into two subsets: learning $L_1^{(lrn)}$ and test data sets $L_1^{(tst)}$. The first set contains 466 randomly selected data samples (66% of the total number of instances) and the second set contains the remaining 233 data samples (34% of the total number of instances). The appropriate numerical datasets: the learning set $L_2^{(lrn)}$ and test set $L_2^{(tst)}$ have been prepared as it is described in Section 2.

Six classifiers have been designed on the base of: a) two-layer perceptron (2 neurons in the output layer and 9 neurons in the hidden layer) using the backpropagation algorithm (with learning rate equals to 0.3 and momentum rate equals to 0.2; the learning process lasted 1000 epochs), b) RBF, using *k*-means learning algorithm, c) decision tree, using C4.5 learning algorithm (the minimum number of cases in the leaf equals to 2), d) SVM, e) Bayesian network and f) *k*-NN algorithm (for the tools d), e), and f) the default parameters provided by WEKA software have been used). The details concerning the learning methods of all the considered classifiers used in this paper can be found in [1]. The results of the experiments are presented in Tab. 1.

Classifier	Learning data set						Test data set					
	Confusion matrix			Number of correct decisions	Number of incorrect decisions	$Q_{RMSE}^{(lrn)}$	Confusion matrix			Number of correct decisions	Number of incorrect decisions	$Q_{RMSE}^{(tst)}$
		b[1)]	m[2)]					b[1)]	m[2)]			
MLP[3)]	b[1)]	**302**	**3**	**462**	**4**	**0.0928**	b[1)]	**143**	**10**	**214**	**19**	**0.2645**
	m[2)]	**1**	**160**	**(99.1%)**	**(0.9%)**		m[2)]	**9**	**71**	**(91.9%)**	**(8.1%)**	
RBF[4)]	b[1)]	299	6	453	13	0.167	b[1)]	147	6	219	14	0.2451
	m[2)]	7	154	(97.2%)	(2.8%)		m[2)]	8	72	(94%)	(6%)	
Decision tree	b[1)]	302	3	461	5	0.0938	b[1)]	146	7	217	16	0.2485
	m[2)]	2	159	(98.9%)	(1.1%)		m[2)]	9	71	(93.1%)	(6.9%)	
SVM[5)]	b[1)]	301	4	456	10	0.1466	b[1)]	146	7	215	18	0.2778
	m[2)]	6	155	(97.9%)	(2.1%)		m[2)]	11	69	(92.3%)	(7.7%)	
Bayesian network	b[1)]	297	8	455	11	0.1442	b[1)]	145	8	224	9	0.1865
	m[2)]	3	158	(97.6%)	(2.4%)		m[2)]	1	79	(96.1%)	(3.9%)	
k-NN [6)]	b[1)]	305	0	455	11	0.0989	b[1)]	146	7	209	24	0.2777
	m[2)]	11	150	(97.6%)	(2.4%)		m[2)]	17	63	(89.7%)	(10.3%)	

[1)] b = benign [2)] m = malignant, [3)] MLP = Multi Layer Perceptron, [4)] RBF = neural network with the Radial Basis Functions, [5)] SVM = Support Vector Machine, [6)] *k*-NN = *k*-Nearest Neighbour algorithm.

Tab. 1. The results of classification for *Wisconsin Breast Cancer* data set

The best system from Tab. 1 (MLP) has been compared with other, to some extent, alternative methods. The results of comparison are shown in Tab. 2. The first four systems: AIRS [2], LS-SVM [13], ME [14] (the abbreviations are shown below Tab. 2) and MLP represent the black-box-type classifiers. The remaining systems: RIAC [7], AR+NN [8], Fuzzy-GA2 [11], NeuroRule2a [13], NEFCLASS [10], GFRBS-1, and GFRBS-2 [4] represent the gray-box-type classifiers.

4. Conclusion

In this paper, six classification systems and their applications to breast cancer recognition problem have been presented. A comparative analysis with other, to some extent alternative approaches has also been conducted. It seems apparent, that the MLP is the best solution to the considered problem (it is more accurate than the other ones).

Acknowledgement

The numerical experiments reported in this paper have been performed using computational equipment purchased in the framework of the EU Operational Programme Innovative Economy

(POIG.02.02.00-26-023/09-00) and the EU Operational Programme Development of Eastern Poland (POPW.01.03.00-26-016/09-00).

Method	Number of correct decisions for learning data set	Number of correct decisions for test data set	$Q_{RMSE}^{(lrn)}$	$Q_{RMSE}^{(tst)}$	Additional information
MLP[1)]	**99.1%**	**91.9%**	**0.0928**	**0.2645**	
AIRS[2)] [2]	97,2%	n/a	n/a	n/a	
LS-SVM[3)] [11]	98,5%	n/a	n/a	n/a	
ME[4)] [14]	98,9%	n/a	n/a	n/a	
RIAC[5)] [7]	95%	n/a	n/a	n/a	
AR + NN[6)] [8]	97,4%	n/a	0,099	n/a	
Fuzzy-GA2[7)][9]	97,4%	n/a	n/a	n/a	3 fuzzy rules, 16 fuzzy sets
NeuroRule2a[8)] [13]	98,1%	n/a	n/a	n/a	3 fuzzy rules, 11 fuzzy sets
NEFCLASS[9)] [10]	95,1%	n/a	n/a	n/a	2 fuzzy rules, 5 input attributes, 10 fuzzy sets
GFRBS-1[10)] [4]	96%	95%	0.2207	0.2138	7 fuzzy rules, 6 input attributes, 12 fuzzy sets
GFRBS-2[10)] [4]	95%	95%	0.2262	0.2132	3 fuzzy rules, 2 input attributes, 4 fuzzy sets

[1)] MLP = Multi Layer Perceptron, [2)]AIRS = Artificial Immune Recognition System, [3)] LS-SVM = Least Square Suport Vector Machine, [4)] ME = Mixture of Experts network structure, [5)] RIAC = Rule Induction through Approximate Classification, [6)] AR+NN = Association Rules and Neural Network, [7)] Fuzzy-GA2 = fuzzy systems and evolutionary algorithms [8)]NeuroRule2a = Neural network Rule extraction) , [9)] NEFCLASS = NEuro-Fuzzy CLASSification, [10)]GFBRS = Genetic Fuzzy Rule-Based System (two structures: more accurate and less interpretable, as well as more interpretable and less accurate),

Tab. 2. The results of comparative analysis of the proposed classification techniques with alternative approaches

References

[1] BISCHOP, C., M. *Pattern Recognition and Machine Learning*. Springer-Verlag, New York, 2006.

[2] GOODMAN, D., BOGGESS, L., WATRKIS, A. *Artificial immune system classification of multiple-class problems.* Data Mining and Complex Systems, vol. 12, New York, 2002.

[3] GORZAŁCZANY, M., B., RUDZIŃSKI, F. *Measurement data in genetic fuzzy modelling of dynamic systems*. Pomiary Automatyka Kontrola, vol. 56(12), 2010, pp. 1420-1423.

[4] GORZAŁCZANY, M., B., RUDZIŃSKI, F. *A modified Pittsburg approach to design a genetic fuzzy-rule based classifier from data*. Lecture Notes in Artificial Intelligence, vol. 6113, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 88-96.

[5] GORZAŁCZANY, M., B., RUDZIŃSKI, F. *Genetic fuzzy rule-based modelling of dynamic systems using time series*. Lecture Notes in Computer Science, vol. 7269, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 231-239.

[6] GORZAŁCZANY, M., B., RUDZIŃSKI, F. *Accuracy vs. interpretability of fuzzy rule-based classifiers: an evolutionary approach*. Lecture Notes in Computer Science, vol. 7269, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 222-230.

[7] HOWARD, J., HAMILTON, J., NING, S., CERCONE, N. *RIAC: A rule induction algorithm based on approximate alassification.* Technical Report CS-96-06, University of Regina, Saskatchewan, 1996.

[8] KARABATAK, M., INCE, M., C. *An expert system for detection of breast cancer based on association rules and neural network.* Signal Processing, Communication and Applications Conference, 2008.

[9] MITCHELL, T. *Machine Learning*. McGraw Hill, 1997.

[10]NAUCK, D., KRUSE, R. *Obtaining interpretable fuzzy classification rules from medical data.* Artificial Intelligence in Medicine, vol. 16, 1999, pp. 149-169.

[11]PENA-REYES, C., A., SIPPER, M. *A fuzzy-genetic approach to breast cancer diagnosis.* Artificial Intelligence in Medicine, vol. 17(2), 1999, pp. 131-155.

[12]POLAT, K., GUNES, S. *Breast cancer diagnosis using least square support vector machine*. Digital Signal Processing, vol. 17, 2007, pp. 694–701.

[13]SETIONO, R. *Generating concise and accurate classification rules for breast cancer diagnosis*. Artificial Intelligence in Medicine, vol. 18(3), 2000, pp. 205-219.

[14]UBEYLI, E., D. *A mixture of experts network structure for breast cancer diagnosis.* Journal of Medical Systems, vol. 29(5), 2005, pp. 569-579.

[15]WOLBERG, W., H., MANGASARIAN, O., L. *Multisurface method of pattern separation for medical diagnosis applied to breast cytology*. Proc. of the National Academy of Sciences, vol. 87, pp. 9193-9196.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Case Study of VANET Cluster Modeling using Markov-Modulated Birth-Death Processes

*Jelena Revzina
*Transport and Telecommunication Institute, Faculty of Computer Science and Telecommunication, Department of Telecommunications, Lomonosova 1, LV-1019 Riga, Latvia, lena_revzina@tsi.lv

Abstract. Vehicular communication system is a promising technology, which can provide customers with various services from safety alert to in-car entertainment. Due to its huge application potential, it attracts attentions both from academia and industry.
In this paper we describe an operation of the VANET (Vehicular Ad hoc networks) clustering using the theory of Markov-Modulated Birth-Death processes. Area of our interests is an evaluation of some VANET cluster performance characteristics.

Keywords: VANET, cluster, Markov-Modulated Birth-Death Processes

1. Introduction

1.1. Motivation

The modeling of Vehicular Ad hoc networks has attracted much attention of researchers during the last few years. Such particular case of mobile ad hoc network is characterized by a strong mobility of the nodes, a high dynamic and specific topology, a significant loss rate and a very short duration of communication.

The Figure 1 shows the classification of the VANET problems [1].

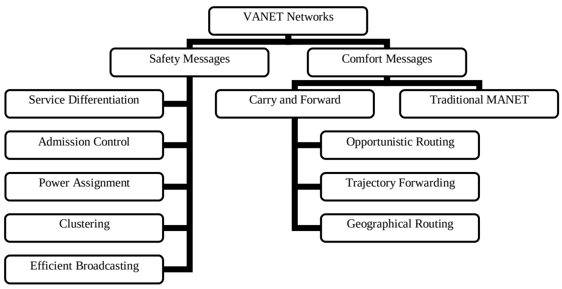


Fig. 1. Networking challenges in VANET.

In our study we consider the mathematical model of the VANET clustering operation.

1.2. Short Description of Mathematical Model

We describe an operation of Vehicular Ad hoc networks (VANET) using the theory of Markov-Modulated Birth-Death Processes [2]. The considered mathematical model is the following.

A Markov-modulated process is defined as two-dimensional continuous-time Markov chain *(X, J)*. So-called Markov component *J* on the finite state space *C = {1, ..., m}*, corresponds to a homogeneous continuous-time Markov chain [3]. This chain is characterized by transition rates $\lambda_{k,j}$, $k, j \in C$. The state $J = k$ is not changed during the exponentially distributed time with the parameter $\Lambda_k = \lambda_{k,1} + \lambda_{k,2} + \dots + \lambda_{k,m}$. After this time the given state is replaced by another state with a probability $q_{k,j} = \lambda_{k,j} / \Lambda_k$. The states of the component *J* changes regardless of the states of component *X*.

If the state $J = j$ is fixed, then the component X on the state space $E = \{1,...,w\}$ behaves as a homogeneous continuous-time birth-death process. In this process the transitions are possible among neighboring states only. The transition rates are follows: $\gamma_i(j)$ for the transition from state i to state $i+1$, $i = 1, 2,..., w-1$; $\mu_i(j)$ for the transition from the state i to state $i - 1$, $i = 2,...,w$. Let $\Gamma_i(j) = \gamma_i(j) + \mu_i(j)$. Is assumed that the Markov chain (X, J) is ergodic.

It is of interest to study the steady state probabilities for the process (X,J): $\pi_{i,j} = P\{X = i, J = j\}, \quad (i, j) \in E \times C$.

2. Proposed Model

We have a cluster of vehicles located and moving freely on certain territory. The flow of vehicles entering the cluster is a Poisson flow with a parameter λ. The sojourn times of the vehicles within the cluster are assumed to be mutually independent random variables, exponentially distributed with the intensity β (therefore, the mean sojourn time is $1/\beta$) [4].

Among the vehicles, a leading vehicle (Head) is singled out, serving as a tool for providing connection between the vehicles being both within the cluster and outside it. Generally speaking, each vehicle entering the given cluster requires a connection to Head, but we can neglect the time needed for this (since it is negligible). Additionally, if Head leaves the cluster, it will be changed to another at time.

Every vehicle can generate a connection request (claim) to the Head. As a result, addressing to Head takes place. The intensity of the one vehicle’s request to the Head equalsν, while the time before the next claim is assumed exponentially distributed, independent of other similar times, the number of vehicles in the cluster, and the Head state.

The Head operates as a single-server queuing system. It can serve only one vehicle at a time. Let the service rate is equal to τ. Should Head happen to be busy at the moment of its being turned to, then the request is directed to the queue from which it will be accepted for service later.

Our task is to determine the characteristics such as mean number of vehicles waiting for communication (pending) to the Head $E(X)$, the probability that the vehicle will have to wait the beginning of a connection to the Head P_w, etc. More generally, we are interested in the probability $\pi_{i,j}$ that j vehicles are in the cluster, while the number of transmitting vehicles or pending connection is equal to i.

The described model is the Markov-type modulated birth-death process (MMBDP). Here, the Markovian component (extraneous factor) J is a number of vehicles in the cluster. The transition rates from state $J = j$ are as follows: λ is the intensity of the transition to state $j + 1$, while $j\beta$ is the intensity of the transition to state $j-1$ $(j > 0)$. Obviously, this process is described by the Markovian queuing system with an infinite number of service places. The stationary distribution with respect to such systems is well known [5], and it is Poisson distribution with the parameter $\rho = \lambda / \beta$. Thus, the probability of j vehicle presence in the cluster

$$\theta_j = \frac{1}{j!} \rho^j e^{-\rho}, \quad j = 0, 1, \dots . \tag{1}$$

The second component X of our two-dimensional process is a number of vehicles that have communication (such vehicles can only be one) or pending connection. If there is a state of $J = j$, then the transition from state $X=i$ to state $i+1$ is realized with the intensity $\gamma_j(i) = (j - i)\nu; \ i \le j; \ i, j = 0, 1, \dots$, while the transition to the state $i - 1$ $(i > 0)$ has the intensity μ that depends neither on i nor on j.

Now, let’s make a few comments to the given description. First, the Markov-modulated process implies the component J changing regardless of the states of component X. Secondly, the changes of both components can not occur simultaneously. In order to provide for these conditions, we assume that the vehicles leaving the cluster neither communicate nor are pending. The only exception is the case where only one communicating vehicle is in the cluster. In this case, should it leave the cluster, then, of course, its communication is interrupted. Note that the number of vehicles being connected or pending is normally much less than the total number of vehicles in the cluster, so the accepted assumptions practically do not affect the accuracy of the calculations.

Now we can apply the method to calculate the stationary state probabilities $\{\pi_{i,j}\}$ as described in paper [2].

3. Computational Procedures

Let's assume that the maximum value of the components of *J* and *X* (the number of their states) is equal to *m* and *w*, *m*<*w*, accordingly. Therefore, the ranges of values of these components will be $C=\{0, 1,.., m\}$, $E=\{0, 1,.., w\}$. We use the conditional probabilities of states of the second component X to calculate the stationary probabilities of states $\{\pi_{i,j}\}$, on condition that the state of the first component *J* is equal to *j*:

$$s_i(j)=P\{X=i\,|\,J=j\}=\pi_{i,j}\,/\,\theta_j\,. \tag{2}$$

Let us recall that the probabilities $\{\theta_j\}$ are calculated according to the formula (1). Let $\Lambda_j=\lambda+j\beta,\ \Gamma_i(j)=\gamma_i(j)+\mu\left(1-\delta_{i,0}\right)$ be the sums of off-diagonal elements of the rows of infinitesimal matrices for the components *J* and *X*. Here $\delta_{i,0}$ is Kronecker symbol: one equals 1 if *i* = 0 and equals 0 otherwise.

For stationary probabilities $\pi_{i,j}=\theta_j s_i(j)$ we have a system of equations as follows:

$$\begin{aligned}&\theta_j s_i(j)(\Lambda_j+\Gamma_i(j))=\theta_j s_{i-1}(j)\gamma_{i-1}(j)+\theta_j s_{i+1}(j)\mu+\\&+\theta_{j-1}s_i(j-1)\lambda+\theta_{j+1}s_i(j+1)(j+1)\beta,\ i\in E, j\in C.\end{aligned} \tag{3}$$

Further we use the following *w*-dimensional column vectors:

$\gamma^{\langle j\rangle}=(\gamma_1(j)\ \ \gamma_2(j)\ \ ...\ \ \gamma_w(j))^T,\ j\le w,\qquad \vec{\mu}=(0\ \ \mu\ \ ...\ \ \mu)^T,$

$\Gamma^{\langle j\rangle}=\gamma^{\langle j\rangle}+\vec{\mu}$, $S^{\langle j\rangle}=(s_1(j)\quad s_2(j)...s_w(j))^T$, as well a diagonal matrix of *w*-th order.

Furthermore, we will need *w*-th order shifting matrix, causing the vector (row) elements being shifted downwards (shift right) by one position.

Now the system of equations (3) can be written as:

$$\begin{aligned}&\theta_j(\Lambda_j I+diag(\Gamma^{\langle j\rangle}))S^{\langle j\rangle}=\theta_j S_1 diag(\gamma(j))S^{\langle j\rangle}+\\&+\theta_j S_1^T diag(\vec{\mu})S^{\langle j\rangle}+\theta_{j-1}\lambda S^{\langle j-1\rangle}+\theta_{j+1}(j+1)\beta S^{\langle j+1\rangle},\ j\in C,\end{aligned}$$

were *I* – identity matrix, and $diag(\gamma(j))$ - diagonal matrix with the diagonal $\gamma(j)$.

Finally, we have:

$$\begin{aligned}&S^{\langle j\rangle}=(\Lambda_j I+diag(\Gamma^{\langle j\rangle})-S_1 diag(\gamma(j))-S_1^T diag(\vec{\mu}))^{-1}\times\\&\times\left(\theta_{j-1}\frac{1}{\theta_j}\lambda S^{\langle j-1\rangle}+\theta_{j+1}\frac{1}{\theta_j}(j+1)\beta S^{\langle j+1\rangle}\right),\ \ j\in C.\end{aligned} \tag{4}$$

The last formulas allow us to implement an iterative procedure of calculating the conditional distributions $\{s_i(j)\}$. At that, it is necessary to set the initial distributions $S^{\langle j\rangle}=(s_1(j)\ \ s_2(j)\ \ ...\ \ s_w(j))^T,\ j\in C,$ as well as normalize the received transitional probabilities continuously so that the sum of elements $S^{\langle j\rangle}$ would be equal to 1.

The unconditional distribution $\{\pi_{i,j}\}$ is given by formula:

$$\pi_{i,j} = \theta_j s_i(j). \tag{5}$$

It allows one to calculate the average number of vehicles in the queuing system:

$$E(X) = \sum_{j=1}^{m} \sum_{i=1}^{w} i\pi_{i,j} . \tag{6}$$

4. Experimental Results

The calculations were performed by the program written in Mathcad 14. As the initial distribution $S^{\langle j \rangle} = (s_0(j)\ s_1(j)...s_w(j))^T$ for all values *j*, the equally probable distribution assigning the same probability to all possible states: $s_i(j) = 1/(1+w)$ if $j < w$, $s_i(j) = 1/(1+j)$ if $j \le w$.

The program realizing the above-described method of calculation computes stationary distributions (4) and (5) very quickly. The results show the obtained values for the distributions $\{\theta_j\}$, $\left\{S^{\langle j \rangle}\right\} = \left\{(s_0(j), s_1(j)...s_w(j))^T\right\}$ and $\{\pi_{i,j}\}$. The speed of convergence of the used iterative procedure is examined for the mean number of vehicles in the cluster.

The numerical results will be shown at the presentation of the “Transcom 2013” conference.

5. Conclusion

We proposed a well-established probabilistic approach to the cluster description in the VANET networks. One is based on the Markov-Modulated Birth-Death processes.

It is important to emphasize that suggested model does not lose its significance because it can be generalized, allowing one to take into account the different dependencies of the queuing system characteristics on the state of Markov component, i.e., the number of vehicles in the cluster.

The future investigations will be connected with the more detailed research and description of the VANET clustering mechanisms and validation of the proposed model using simulation software VISSIM Car2X.

Acknowledgement

I wish to express my gratitude to my supervisor, Prof. Alexander Andronov, for his invaluable guidance and encouragement in producing this work. The article is written with the financial assistance of European Social Fund. Project Nr. 2009/0159/1DP/1.1.2.1.2/09/IPIA/VIAA/006 (The Support in Realisation of the Doctoral Programme “Telematics and Logistics” of the Transport and Telecommunication Institute).

References

[1] YOUSEFI, S. and others. *Vehicular Ad Hoc Networks: Challenges and Perspectives.* 6th International Conference on ITS Telecommunications Proceedings. June 2006, Chengdu, Sichuan, China. USA: IEEE, 2006.

[2] ANDRONOV, A.M. *Markov-Modulated Birth-Death Processes*. Automatic Control and Computer Sciences,Vol. 45, No. 3, 2012.

[3] PACHECO, A. and others. *Markov-Modulated processes and Semiregenerative Phenomena*. New Jersey: World Sci., 2009.

[4] REVZINA, J. *Analysis of VANET Cluster Performance using Markov-Modulated Processes*. Reliability and Statistics in Transportation and Communication (RelStat’12), Riga, Latvia, 2012.

[5] IBE OLIVER. C. *Markov Processes for Stochastic Modeling* - Academic Press , 2009, pp.185-212.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Intelligent Transport Systems in North European Environment

[*]Miroslav Rusin
[*]University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitna 1, 010 26 Žilina, Slovak Republic, Miroslav.Rusin@fri.uniza.sk

Abstract. This contribution discusses the Intelligent Transport Systems (ITS) in Norway. The realization of the vision of ITS has to involve the government and government organizations, as reflected in the Norwegian national plan. Commercial projects and also the results of national and international research projects are presented. The conclusion includes recommendations for the design of intelligent transport system with respect to the performed review.

Keywords: intelligent transport systems, Norway, deployment

1. Introduction

Intelligent Transport Systems (ITS) is an umbrella term for a variety of electronic, information processing, communication and control technologies that can be combined and used in the transport sector [1]. There is no clear definition of what ITS is and what is not. However, intuitively each ITS must include at least some form of information processing and management of a computer, car or road network to be considered intelligent. ITS can refer to a single technology, integrated system or network of systems. As stated in [1], ITS is neither a monolithic system nor system integration. We can say that "ITS is a multifaceted approach to address the needs of transport". ITS have several functions that perform in a number of ways. They can work with one user or a vehicle or affect the entire road network. ITS can be used to improve traffic safety, traffic flow and capacity, efficiency and productivity of public transport and commercial vehicles, reduce vehicle emissions and consumption of resources. While the potential for ITS to improve the environmental and economic productivity is promising, perhaps the biggest impact of ITS lies in improving road safety. Many systems have been developed with the specific aim to increase passenger protection or the protection of vulnerable road users. Several studies have estimated the potential of ITS for safety. For example, McKeever [2] estimates that 26% of fatalities and 30% of injury accidents could be avoided with a comprehensive in-vehicle infrastructure- and ITS cooperative-based system in the USA.

Another report mentioned in [2] conservatively predicted that full deployment of ITS in OCED countries could lead to the prevention of deaths to 47,000, a reduction in fatalities and accidents with injuries by up to 40%, and the estimated annual savings in accident costs of $194 billion. Economies of fatalities alone are estimated at $73 billion in the USA. Similarly, ERTICO filed a vision for the future of ITS, in which these predictions were stated:

- ITS will contribute significantly to 50% deaths on road.
- 25% reduction in travel time due to ITS.
- 50% reduction in urban centers due to traffic management
- Automatic notification of accidents will result in a 15% reduction in fatalities.
- 40 hours saved in traffic due to an automated system for toll collection.
- 50% reduction of delays due to public transport systems.
- 25% reduction in operating costs due to commercial vehicles fleet management system.

1.1. Legislative Framework

There is no ideal model for the cooperation of the authorities, but there are certain principles that should be applied to services that meet the objectives of transport policy and are financed by public funds or users [3]. Services are designed to meet the diverse needs of the entire value chain. It is necessary to consider the following three dimensions:

- Political and legal responsibility
- Ownership (infrastructure, resources of data systems and information content)
- Operation and services

Public authorities should be legally responsible for systems and services that are initiated and financed by the government, especially if it a system is an open transport system component. Authorities have a particular responsibility to promote transport companies, systems of suppliers and others to build their systems on common architecture, standards, formats and underlying data. ITS raises important questions about the individual's right to privacy because electronic services often include the registration of personal data and store electronic tracks. Use of ITS will result in safer and more efficient operations in service in the transport sector, but conflicts can arise between security and profit on the one hand and considerations of individual integrity on the other.

This means that the right to privacy must be balanced with other interests of consumers. Today's users of the road network are anonymous and have great individual freedom while traveling, while in other sectors such as air transport, security aspects require a high degree of supervision and control individuals. Such considerations may lead to a conflict with an individual's right to privacy. Enhancing safety, traffic management and environmental quality in the transport system will require supervision and control measures, which may not always be compatible with the freedom of the individual [4].

Norwegian legislation is based on internationally recognized principles and imposes specific requirements for the collection and processing of personal data. ITS solution is a challenge to laws, policies and mechanisms to ensure that the guaranteed right to privacy. The transport sector is covered by these provisions and rules and must be taken into account when addressing this issue.

It is extremely important to protect personal data against unauthorized access and abuse, and in many cases it is necessary to introduce measures to protect privacy. These measures must be taken seriously and incorporated into the development of new services right from the start. A solution that takes into account the right to privacy, consumer rights and the need for reliable data and statistics must be found. In determining whether an ITS solution is in conflict with privacy the following factors must be clear:

- The purpose of collecting and storing data
- notification that the data are stored
- The way registered data are managed

When new systems are being introduced, a dialogue with the relevant authorities at an early stage must begin Public transportation management must also take proactive measures with regard to regulations and guidelines that must be examined in relation to new ITS solutions.

We see that ITS has a vast potential in many aspects of transportation, however, its deployment in real environment is not straightforward and there are some obstacles that need to be overcome in such process. The aim of this paper is to look at a success story located in Northern Europe.

2. Norwegian national Plan

Efficient transport is a prerequisite for prosperity and economic growth. The main objective of transport policy is to provide a transportation system that is efficient, accessible, safe and environmentally friendly and meeting the transport needs of society and promoting regional development. For this purpose Norwegian national plan was proposed [5].

Population growth combined with economic growth and rising incomes will result in a strong increase in demand for transport. This leads to a number of problems. Capacity of transport systems must be expanded. The government objective must be fulfilled and the transport system must be more environmentally friendly. Although population growth in urban areas is a major challenge, it also provides opportunities to develop denser urban centers and transportation system in which public transport, walking and cycling absorb growth in local traffic. This will help in reducing emissions and smog from transport. Responsible transport authorities propose three-pronged policy:

In major urban areas, the growth of local transport will be absorbed into public, pedestrian and cycling transport. In the city the priority is given to public solutions and measures to facilitate walking and cycling. The government must provide a financial framework for municipalities and counties to expand their public transport systems. More government funding is needed. Restrictions must be applied to the use of private cars. Examples include congestion charges and higher parking charges.

Measures will be taken to promote regional growth, by creating more robust residential and employment areas. Development of InterCity railway lines in the area and around the main cities and road network in the region are measures that can help to promote regional growth and development. Transport system between cities and regional areas will be improved, which will reduce costs for industry and population. This will help in promoting economic growth and regional development.

National transport network with active nodes will help to ensure that more remote cargo is transported by sea or by rail instead of by road. More effort must be invested in the operation and maintenance of the entire transport network, because infrastructure needs to be more robust to withstand climate change and other challenges. Traffic volumes and increasing demands on reliability, and equipment to be operated and maintained, are still technically complicated. Lack of maintenance over long period has resulted in a growing need for infrastructure renewal. To obtain more reliable and robust infrastructure it is necessary that the standard infrastructure has been brought to a level where availability and security are ensured and wear is minimized. This will help reduce costs for infrastructure owners, users and road operators.

Today, too many people are killed or seriously injured in road traffic. Efforts to improve the operation and safety in all modes of transport must be enhanced. Traffic authorities propose an ambitious goal for the road traffic: in 2024 the number of fatalities should not exceed 100 and the total number of fatalities and serious injuries should not exceed 500.

The 2010-2019 National Transport Plan (NTP 2010-2019), a financial framework for the transport sector has been increased by NOK 100 billion from the level in the previous plan. The increased level of activity represents a major challenge for construction and consulting industries in terms of planning and implementation. Transport authorities cooperate with these sectors in order to promote the development of competencies and expand the implementation capacity to achieve shorter planning periods and new strategies. The long term, predictable funding is requested in order to ensure rational development of transport. With sufficient and predictable funding and effective planning processes, transport infrastructure can be developed to a good level during the 20-year period.

2.1. The implementation of ITS in the specific sub-Arctic areas

ARKTRANS is the Norwegian multimodal framework for ITS [6]. It covers the whole spectrum of the transport sector; its specifications are valid for all modes of transport (road, sea, rail and air, as well as freight and passenger). ARKTRANS provides multimodal (common to all modes of transport) specification of responsibilities, functions, processes, and information flows in the transport sector. ARKTRANS is based on a comprehensive study of the transport sector. All stakeholders, projects and activities that represent different modes and different responsibilities, contributed to this framework. ARKTRANS is used in national and European projects, which focus on modal solutions, and ARKTRANS is continuously updated based on new knowledge and the results of these projects. A common framework for multimodal ARKTRANS supports effective cooperation modal transport solutions. System interoperability, efficient information flow, coordination of transport, etc., are supported using common concepts and specifications that bridge the current differences in semantics and ensure interoperability and efficiency.

3. Design Recommendations

Considering the current ITS deployment in Norway following principles for the future design to be applied in the project can be summarized:

- Intelligent transportation system must be based on the European model of legal acts of the European Union.
- The model must take into account the specific conditions of the regions in which they will be applied, whether it is on the level of municipalities, districts, on a national or international level.
- The model should be designed for the needs of tourism, transportation of goods and public transport.
- This model should facilitate the implementation of an integrated transport policy, whether it is of a ship, air, and rail or road type. Such a model should as far as possible in the given cases facilitate the exchange of incident reports and information concerning the planned service.

Acknowledgement

This contribution/publication is the result of the project implementation:
Centre of excellence for systems and services of inteligent transport II., ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] Ezell S. Intelligent Transportation Systems. 2010.

[2] Bayly M. Review of crash effectiveness of Intelligent Transport Systems. Emergency. 2007.

[3] John Miles, Finella McKenzie, Keen K. THE INTELLIGENT TRANSPORT SYSTEMS (ITS) PRACTITIONERS' GUIDE TO EUROPE. 2011.

[4] Eriksson L, Bjørnskau T. Acceptability of traffic safety measures with personal privacy implications. Transportation Research Part F: Traffic Psychology and Behaviour. 2012;15(3):333-47. doi: 10.1016/j.trf.2012.02.006.

[5] AVINOR. The Proposed Norwegian National Transport Plan, 2014-2023. 2012.

[6] ARKTRANS. ARKTRANS Brochure. 2009.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

The Necessity to Expand Risk Management Systems in IT Projects

*Marko Schwedler
*University of Transport, Faculty of Operation and Economics of Transport and Communications, Univezitna 1, 01001 Žilina, Slovak Republic, marko.schwedler@markoschwedler.net

Abstract. The article considers IT risk management and demonstrates the necessity for an expansion. This necessity is illustrated by statements concerning uncertainties. It is argued that elements of Principal Agent Theory and of Transaction Cost Theory provide approaches to reduce uncertainties.

Keywords: Risk management, Uncertainty, Opportunism, Moral hazard, Hidden information, Hidden action, New institutional economics, Principal agent theory, Asset specificity, Transaction costs

1. Introduction

Enterprises depict their business processes with IT-Systems. SAGE, Microsoft Dynamics and SAP are examples of IT-Systems. These systems are introduced in projects. Project management by product managers forms the leadership framework for realizing and controlling projects. An aspect of project management is risk management. "Project Risk Management includes the processes of conducting risk management planning, identification, analysis response planning, and controlling risk on a project."[1] This risk management will be regarded more closely in this paper.

The definition of project risk has been adopted from the Project Management Institute (PMI): "Project risk is an uncertain event or condition that, if it occurs, has a positive or negative effect on one or more project objectives such as scope, schedule and quality."

While considering project risk, uncertainty will be ignored. This is the starting point of this paper.

2. Objective of This Paper

This article argues the necessity to expand risk management systems that are used in IT projects. Thereby, the PMI's project definition will be applied: "A project is a temporary endeavor undertaken to create a unique product, service, or result."

The expansion refers to the implementation of uncertainty in the timeframe of IT risk management systems. Statements concerning reduction of uncertainities will be formulated concisely.

The theoretical basis for the following is New Institutional Economics (NIE): "[…] the object of NIE is the determination of institutional constructs that take positive transaction costs, imperfect foresight and bounded rationality into account."[2]

The comprehension of uncertainty is related to anticipated behavior. Here, these expectations are bounded rationality and opportunism.[3]

3. Identifying Risks in IT Risk Management Systems

Risk management differentiates between strategic and operative risk management. Strategic risk management considers basic aspects, cultural aspects and methodological procedure.

Operative risk management can be divided into the following components and/or phases:

- identification
- evaluation
- control
- supervision.

The identification of risks can be realized through self-assessment, process analyses, system analyses and fault analyses of earlier errors. Further possibilities to identify risks are creativity and benchmarking techniques as well as exchange of experience.

Self-assessment can be effected through individual interviews, workshops, surveys and employee suggestion systems. In addition to IT damage reports, the analysis of damage reports regards non-IT damages and external damage reports.[4]

The evaluation contains the determination of an occurrence probability for risks and their impact assessment. Determining occurrence probability for uncertainties is not possible. Consequently, ex-ante and ex-post opportunism must be assumed. Therefore, basic measures and rules are to be framed.

An explicit observation of uncertainty is not discernible in the classic IT risk management. No more so is consciously dealing with bounded rationality.

4. Expanding IT Risk Management Systems

PMI claims that unknown risks cannot be managed in advance. PMI recommends including a management reserve.[1] This assertion will be questioned here.

Among the unknown risks that are identified as uncertainties are theft, secure programming, the use of current programming guidelines, consciously writing incorrect invoices, nonexistent competences, insufficient work input, installation of obsolete and insecure software and withholding necessary technical measures.

These uncertainties constitute not only a cost factor but also a security risk. The discussions about Swiss banks demonstrate the scope of data theft. Delays in project due-dates can lead to competitive disadvantages. Thus, a professional, proactive risk management that consciously factors in uncertainties is necessary.

This risk management already begins in considering whether an IT project should be realized internally or externally.

The relevant criteria for the evaluation are asset specificity and transaction costs.[3] Williamson identifies a shift towards internal organization when asset specificity increases. Human asset specificity as the main characteristic of asset specificity is reflected in the sum of operational process expertise, customer needs, technical implementation and knowledge of correlations between varying operational processes. The higher the asset specificity, the sooner an internal organization—that is, an internal project—is realized. Thereby, the uncertainty concerning staff qualifications can be reduced since the staff has been hired by the enterprise. Observing working staff further reduces the uncertainty concerning knowledge and motivation. The danger of exploiting an "aware" employee can be averted since this employee remains in the enterprise and a common interest in the firm can be assumed.

If it is decided to realize a project externally, the construction of the preliminary agreement is to be understood as a phase of uncertainty reduction. The problem of asymmetrical information is to be heeded. Uncertainty is shown here in the "actual" intentions and/or competences of the external

IT bidder. He/she appears as Agent (A), the client as Principal (P). The Principal-Agent theory offers itself as resolution of uncertainty, screening and signaling models. Signaling models offer agents the option to reveal agent-specific information, that is, to signal.[5] A screening model can be understood as a situation in which the principal draws up a self- conceived contract that is then offered the agent. By accepting or rejecting the offered contract, information is revealed by the agent to the principal.[6]

In case of an external realization, team production is to be observed. The design of team production is therefore the next component to be regarded. Team production is present if a number of types of production factors have been introduced to the production process, if separable outputs of the introduced activities cannot be measured and if all introduced factors do not belong to one participant alone.[1] The danger in team production is the shirker problem, seen in the slight factor commitment of a participant. As a result, a project constellation can arise in which employees act in a project and thus also invoice although they have performed no services.

A solution that the theory offers is the specialization of a participant in supervising team members. Idlers can be removed from the project upon identification. [7]

5. Conclusion

The article has made the problem of uncertainties in IT projects topic of discussion. The necessity against the background of the consequences of disregard has been underlined. Within in-house training and further education, enterprises should implement aspects of the Principal-Agent theory that are based on realistic anticipated behavior. A form of contract compatible with incentives should follow.

The expansion of the IT risk management system is an initial, necessary step to a project management that factors in realistic anticipated behavior. A further step is the installation and ensuing communication of contractual components relevant to incentives. This will be the topic of a further article.

References

[1] Project Management Institute (2010): A GUIDE TO THE PROJECT MANAGEMENT BODY OF KNOWLEDGE. 5th ed.

[2] Richter, Rudolf; Furubotn, Eirik G. (2010): Neue Institutionenökonomik. Eine Einführung und kritische Würdigung. 4th, rev. and exp. ed. Tübingen: Mohr Siebeck.

[3] Williamson, Oliver E. (1990); Streissler, Monika: Die ökonomischen Institutionen des Kapitalismus: Unternehmen, Märkte, Kooperationen. Unternehmen, Märkte, Kooperation. 1. Aufl. Tübingen: Mohr Siebeck (64).

[4] Seibold, Holger (2006): IT-Risikomanagement: Oldenbourg Wissenschaftsverlag.

[5] Roiger, Manuela Barbara (2012): Gestaltung von Anreizsystemen und Unternehmensethik: Eine norm- und wertbezogene Analyse der normativen Principal-Agent-Theorie 2007. Ed.: Deutscher Universitätsverlag.

[6] Jost, Peter-J (ed.) (2001): Die Prinzipal-Agenten-Theorie in der Betriebswirtschaftslehre: Schäffer-Poeschel.

[7] Erlei, Mathias; Leschke, Martin; Sauerland, Dirk (2007): Neue Institutionenökonomik. 2nd ed.: Schäffer-Poeschel.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Graphical User Interface for Automatic Driver Loading

*Roman Slabý, *Martin Stankuš, * Radim Hercík, *Vilém Srovnal Jr.
*VŠB – Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, Department of Cybernetics and Biomedical Engineering, 17. listopadu 15, 70833 Ostrava - Poruba, Czech Republic, {roman.slaby, martin.stankus, radim.hercik, vilem.srovnal1}@vsb.cz

Abstract. This article describes the possibilities of drivers loading in QNX real-time operating system. Driver loading is based on information about cards connected to the bus and information stored in the system configuration database. Drivers under QNX are not loaded automatically but are loaded by shell script or console command. This motivates the effort to create a configuration tool capable of automatic loading of individual drivers with preset parameters. The parameters can be changed at any time in a graphical user interface. Shorter time needed for configuration of individual drivers leads to increased efficiency of system deployment and significant financial saving.

Keywords: QNX, Drivers, Fieldbus, PCI cards

1. Introduction

QNX operating system is multi-platform, real-time operating system that uses microkernel architecture and is developed by QNX Software Systems. QNX is a Unix-like operation system compliant with POSIX specification [1].

Graphical configuration system is designed to manage software support for industrial communication adapters in the QNX operating system. Individual industrial communication networks differ in needed software support and methods of communication. Graphical configuration system unites these different requirements for software support and configuration.

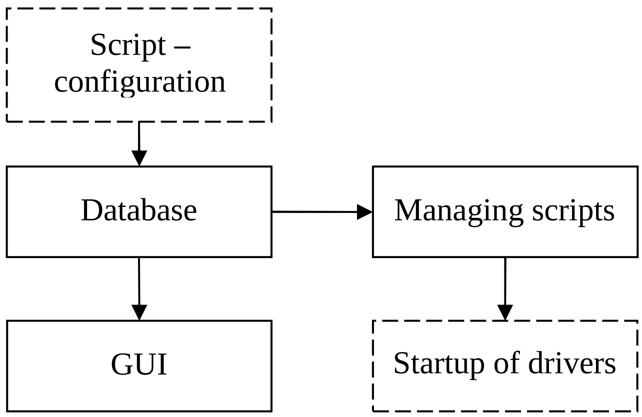


Fig. 1. Block diagram of the automatic driver loading.

Graphical configuration system can be divided into several functional blocks. The core of the system is a database that stores all configuration data. This database is implemented using SQLite technology. The data from the database are displayed and modified in the graphical user interface. Each industrial communication adapter is provided with dedicated space in the graphical user interface. Script management module performs operations with individual scripts, such as calling of scripts with appropriate parameters. These operations are based on the data from the database.

Initial configuration is loaded into database by scripts specific for every type of industrial communication adapter.

2. Identification of Connected Cards

Graphical configuration system allows setting of configuration parameters for supported cards and is able to load the drivers with appropriate parameters for these cards. Identification of cards begins with generation of data file containing output of "enum-pci" command. Information describing individual cards are separated from the data file and compared with the information stored in the database. Sequence of driver loading is shown in Figure 2.

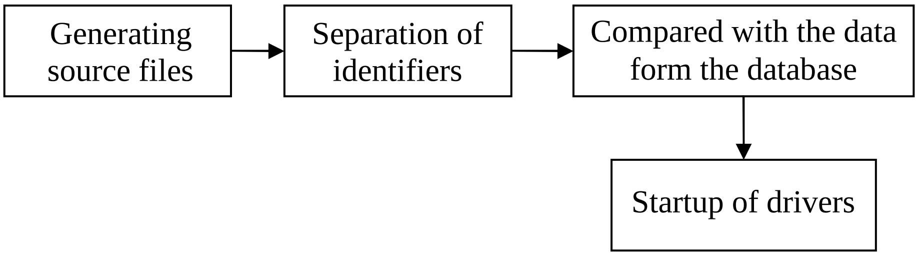


Fig. 2. Identification of the PCI card.

Separated data containing the identifiers of all available cards are compared with the data stored in a device database. Identifiers are information needed to uniquely identify every attached card. These identifiers include Vendor ID, Devices ID, Subsystem Vendor ID and Subsystem ID. These values are stored in the PCI card configuration space. Vendor is unique 16 bit number used to identify manufacturer of the card. This unique number is assigned for a fee to every PCI device manufacturer by PCI consortium, which involves more than 800 companies. The Board of Directors of this consortium is composed of representatives of the most important technological companies, such as Intel, Microsoft, IBM, AMD, Agilent Technologies and more. Device ID is assigned by deice manufacturer. This 16 bit value is unique within particular Vendor ID scope and is stored in Device ID register. There is possibility to use Subsystem Vendor ID and Subsystem ID to further specify card's manufacturer and model of the card [5].

Unlike databases build on a client-server model, where the database server is running as a dedicated process, SQLite is only a small library that is linked with the application. Each database is stored as a separate .dbm (Database Manager) file, where the data are stored in equally sized blocks using a simple primary key. Database for the configuration system contains two tables. These tables are named DEVICES_INFO and DRIVER. Table DRIVER contains information about executed driver and its parameters. Table DEVICES_INFO contains list of supported cards (devices) and data elements to uniquely identify the card. Besides this it contains STATUS flag determining whether a given card is present or not [3] [4].

DEV_ID	VENDOR	DEVICE	SUBVENDOR	SUBSYSTEM	DEC_CH1	DEC_CH2	DEC_CH3	DEC_CH4	STATUS
1	3756	8	3756	8	2	0	2	0	1
2	4277	36912	5280	42	1	0	0	0	0
…	28	1	28	1	1	0	0	0	0

Tab. 1. Contents of the table DEVICES_INFO.

Driver loading is realized by script management block. This block continuously monitors database transactions. Actual driver loading or unloading is implemented by shell scripts. Calling of these scripts is based on database monitoring by script manager block. Script filenames, paths to scripts and scripts parameters are stored in a database table named DRIVER.

DRV_ID	COMMAND	PARAMETERS	LOAD
1	"./sbin/runme.sh "	"-x 500 -t 10"	1
2	"./sbin/runme1.sh "	"-x46 -t 12"	0
…	"./sbin/runme.sh "	"-x 25 -t 150"	0

Tab. 2. Contents of the table DRIVER.

Database content is read after start of the system. This is followed by loading of the driver for every card, which is described by a valid entry in the database and is also physically connected to the computer. Database content can be dynamically modified by user changes through GUI. From a technical view the application is notified by calling of callback function. This callback function identifies the record in the database which is related to change and performs update of this record. Termination and reloading of driver with new parameters is performe d as well.

Drivers are unloaded before termination of system operation. Individual scripts for driver termination are called with appropriate parameters.

3. Identification of Connected Cards

Photon is a graphical user interface (GUI) environment for the QNX realtime operating system and QNX-based applications. As a GUI, Photon provides a flexible, easy-to-use environment for you to interact with your computer. Like QNX itself, Photon is built around a small microkernel. This modular architecture makes Photon fast, flexible, and inherently capable of network-distributed computing. It's designed to fit in embedded systems [2].

The graphical user interface is designed for supported cards. Addition of another card is possible after extension of graphical user interface.

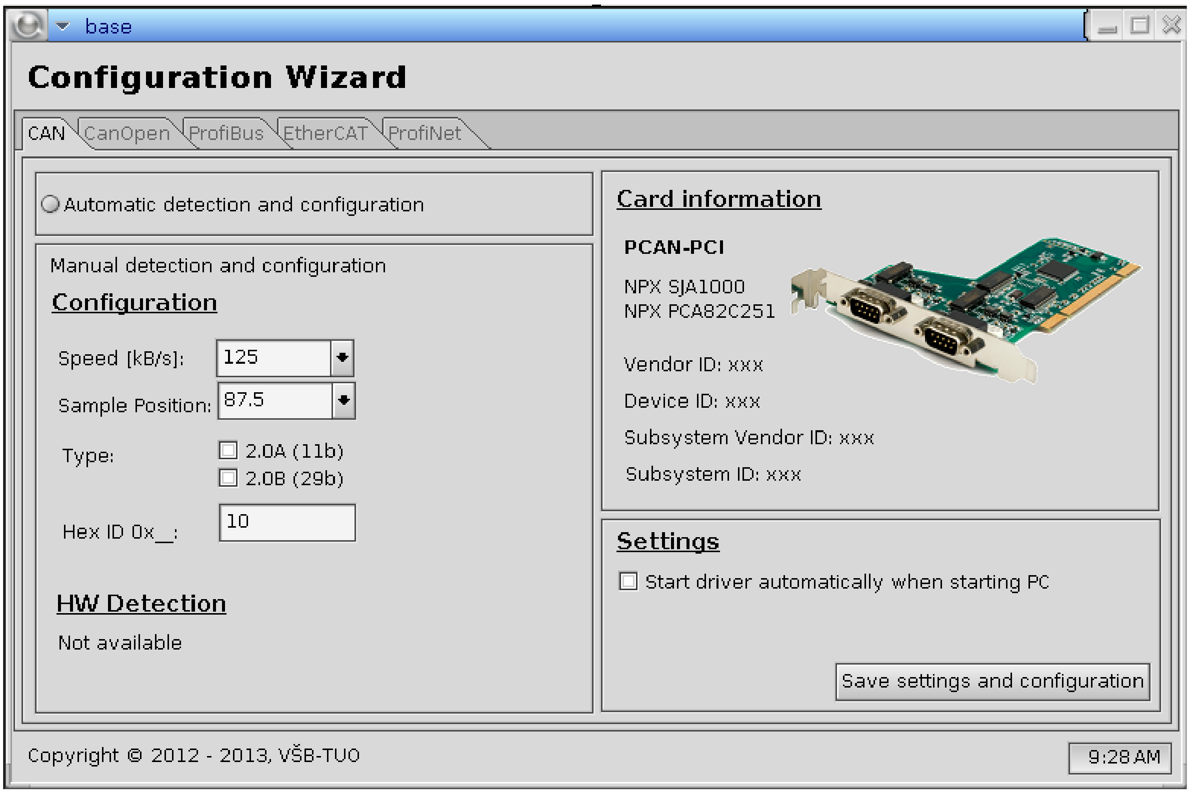


Fig. 3. Graphical user interface for configuration of the CAN driver

4. Conclusion

Automatic driver loading is matter of course for users of the Windows operating system. However, this is not true for users of QNX operating system. It is necessary to load the driver by shell script or console command under this operating system. The goal of described system is to clarify and facilitate initial setup of drivers including their loading. Our graphical configuration system currently supports only limited number of communication cards but can be extended. This extension can be especially useful for systems like SCADA. Communication via fieldbus networks often require setting of specific parameters related to bus communication. The disadvantage of our system is the need to adjust the graphical interface whenever the set of supported cards is changed. This could be solve by dynamic modification of graphical user interface with layout of interface describes by additional data.

Acknowledgement

This work was supported by project TAČR TA01010632 and Grand- aided student R. Hercík, Municipality of Ostrava, Czech Republic.

References

[1] KOLNICK, Frank. *The QNX 4 real-time operating system: programming with messages in a distributed environment*. [new ed.]. Willowdale, Ont: Basis Computer Systems, 1998. ISBN 09-219-6001-8.

[2] G KRTEN, By Rob. *Getting started with QNX Neutrino 2: a guide for realtime programmers*. Kanata, Ont: PARSE Software Devices, 2001. ISBN 09-682-5011-4.

[3] KREIBICH, Jay A. *Using SQLite*. 1st ed. Sebastopol, CA: O'Reilly, 2010. ISBN 05-965-2118-9.

[4] OWENS, Michael. *The definitive guide to SQLite*. New ed. Berkeley, CA: Apress, 2006. ISBN 15-905-9673-0.

[5] ABBOTT, Doug. *PCI bus demystified*. Eagle Rock, Va.: LLH Technology Pub., c2000, viii, 198 p. ISBN 18-787-0754-X.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

3D Shape-Motion Detection

*Peter Sykora, *Robert Hudec, *Miroslav Benco
*University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitna 2, 01026 Žilina, Slovakia, {peter.sykora, robert.hudec, miroslav.benco}@fel.uniza.sk

Abstract. A new concept of shape-motion detection is proposed in this paper. Shapes and motion plays important role in human-computer interaction. Firstly, disparity map is obtained using Microsoft Kinect system. After capture an input image and its corresponding disparity map is preprocessed. Next, the preprocessed disparity map is segmented. Then, the feature and corresponding points were extracted. Finally, the 3D image is obtained. This proposed method can be applied to detect actions, which are composite from multiple shapes of the objects and different motions. This system can be used in real applications.

Keywords: 3D shape, motion detection, Microsoft Kinect, disparity map

1. Introduction

Feature extraction is the first step towards an abstraction, to gain content-related information from the signal or make two signals comparable by criteria beyond sample-wise comparison. Usually, the output of feature extraction is sets of parameters which allow classifying the signal based on feature properties. [1]

Object shape features are very powerful when used in similarity search and retrieval. This is because the shape of objects is usually strongly linked to object functionality and identity. Humans can recognize characteristic object solely from their shape – an illustration that shape often carries semantic information. This property distinguishes shape from some other elementary visual features, such as color or texture. A large body of research has been devoted to shape-based recognition, retrieval and indexing [2]

As shown in [3], motion is a typical phenomenon to be studied in computer vision and robotics fields. The visual analysis of human actions, behaviors or activities inherently targets the underlying motions. Besides, motions are also typically a kind of constitutive element of robot tasks to be learned by robots.

The outline of the paper is as follows. In the next section, an overview of shape-motion detection system is introduced. The image segmentation algorithm is described in section 3. In section 4, the process of feature point's extraction using SIFT and SSD algorithms is presented. Finally proposed method is introduced in section 5, and it is followed by conclusion in Section 6.

2. Relation to Previous Work

In [4], they proposed a scale invariant descriptor using VMDs. This descriptor is able to handle a shape, the parts of which are drawn using different scales. This descriptor can handle some changes in shape, but only to stretch of the shape.

For simple hand gesture, recognition based on shape can be used [5]. But the weakness of this method is that they define certain parameters and threshold values experimentally since it does not follow any systematic approach for gesture recognition, and maximum parameters taken in this approach are based on assumption made after testing number of images.

As shown in [6], there are many works focused on global motion. In terms of local motion, much work concerns human activity. Their major contribution lie in the combination of local motion related parametric recovery and moment – based motion features to derive the concept of "salient frames" for distinguishing motion sequences.

This [7] study proposed a method to detect characteristic posture to constitute sign language word using Motion Quantity from the movement of the sign language speaker who photographed it in digital video camera. However, this method can detect gestures only by characteristic posture in it.

3. Image Segmentation

Since objects of interest are often located on chairs or fixed at walls, the space is segmented into a support surface and a background surface. Goal in image segmentation is to separate background areas of the image from foreground regions of motion that are of interest for human tracking. For calculating the segmentation GrabCut algorithm is used. GrabCut is an interactive image segmentation technique based upon Gaussian Mixture Model (GMM) and iterative energy minimization. The energy minimization is optimized by graph cut algorithm. In this paper, GrabCut is used on the disparity space image to get object coarse silhouette [8] By GrabCut image segmentation unnecessary data are removed. Summary of GrabCut algorithm is divided into following stages:

1. The GMMs are initially created for the foreground and back ground classes.
2. Each pixel in foreground class is assigned to the GMM component in the foreground GMM and each pixel in the background class is assigned to GMM component in the background class.
3. New GMM parameters are learned from the pixel sets which are created in previous set.
4. Using min cut (max flow) to estimate the segmentation and find the new foreground and background classification of pixels.
5. Apply border matting.
6. User editing to improve the result.

4. Feature Calculation

The feature extraction algorithm improves precision of final shape recognition. In order to find the feature points, the SIFT descriptor was used. SIFT image features provide a set of features of an object that are not affected by many of the complications experienced in other methods, such as object scaling and rotation.

4.1. SIFT Descriptor

Scale Invariant Feature Transform (SIFT) [9] is a local descriptor of image features, which is insensitive to illuminant and other variants and is usually used as sparse feature representation. The first step computes the locations of potential points of interest in the image by detecting maxima and minima of a set of Difference of Gaussian (DoG) filters applied at different scales all over the image. Then, these locations are refined by discarding points of low contrast. An orientation is then assigned to each key point based on local image features. Finally, a local feature descriptor s computed at each key point. This descriptor is based on the local image gradient, transformed according to the orientation of the key point to provide orientation invariance [10].

5. Proposed Method

The proposed method relies on the computation of the disparity map [11] from an image to extract disparity information using the Kinect camera system, which interprets 3D scene information from a continuously projected infrared structured light [12]. The proposed approach is implemented on the AMD Opteron 2.3 GHz. The experiments have been implemented in C programming language and CUDA platform [13]. The motion detection algorithm is implemented using the OpenCV real-time computer vision library. The block diagram in Figure 1 gives an overview of the software architecture of the proposed motion system.

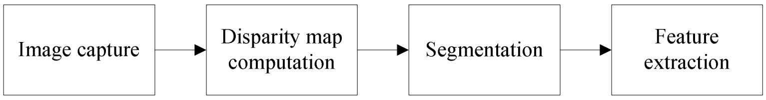


Fig. 1. Block diagram of the proposed method.

After input image (see Figure 3a) is captured, disparity map (see Figure 3b) is calculated and segmented. It is necessary for object extraction from map. Then, the feature and corresponding points are extracted. The feature extraction algorithm reduces data size of disparity map and improves precision of shape recognition. Finally, the image description using SIFT algorithm is obrained. Steps 1-4 for every frame of video sequence will be repeated for specific length of time.

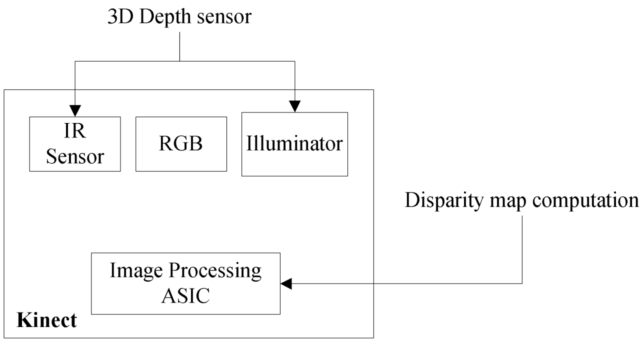


Fig. 2. Microsoft Kinect block diagram.

The basic architecture of Kinect camera system is shown in Figure 2. The RGB sensor is a regular camera that streams video with 8 bits for every color channel. The depth sensing system is composed by an IR emitter and CMOS sensor. The IR light source projects structured light, which is then captured by the CMOS image sensor, and decoded to produce the disparity image of the scene.

a) b)

Fig. 3. Kinect sensor color image (a) and disparity map (b).

The disparity (see Figure 3b) value is encoded with gray values the darker a pixel, the further away the point is to the camera in space. The black pixels indicate that no depth values are available for those pixels. This might happen if the points are too far (and the depth values cannot be computed accurately), are too close (there is a blind region due to limited fields of view for the projector and the camera), are in the cast shadow of the projector (there are no IR dots), or reflect poor IR lights (such as hairs) [14].

6. Conclusion

We presented a simple and effective shape-motion system for combination with Kinect disparity estimates, which can be applied without any further hardware requirements. We described a proposed method that allows for improved correspondence matching between RGB and IR cameras. This detection can improve not only human-machine interaction, but computer vision as well. However this study is still incomplete, there are many issues to be discussed. In future, we plan to perform experiments with more complex segmentation and feature extraction algorithms in order to compare the presented approach with other existing algorithms.

Acknowledgement

The work presented in the paper has been supported by the Slovak Science project Grant Agency, Project No. 1/0705/13 "Image elements classification for semantic image description".

References

[1] OHM, J. R. *Multimedia communication technology.* Springer-Verlag Berlin Heidelberg New York, ISBN 3-540-01249-4, 2004.

[2] MANJUNATH, B. S., SALEMBIER, P., SIKORA, T. *Introduction to MPEG-7.* John Wiley & Sons LTD, England, ISBN 0-471-48678-7, 2003.

[3] WU, S., LI, Y. F., ZHANG, J. *Motion descriptor: a motion trajectory signature.* Springer, IEEE International Conference on Information and Automation June 22-25, China, ISBN 978-1-4244-3608-8, 2009.

[4] UEAOKI, K., IWATA, K. *A novel shape descriptor based on von mises distributions.* Second World Congress on Nature and Biologically Inspired Computing, Japan, ISBN 978-1-4244-7376-2, 2010.

[5] PANWAR, M. *Hand gesture recognition based on shape parameters.* International Conference on Computing, Communication and Applications (ICCCA), ISBN 978-1-4673-0270-8, 2012.

[6] TONG, X., DUAN, L., XU, CH., TIAN, Q., LU, H. *Local Motion Analysis and Its Application in Video based Swimming Style Recognition.* 18th International Conference on Pattern Recognition, ISBN 0-7695-2521-0, 2006.

[7] TAKAI, M., *Measurement of Motion Quantity from human movement and detection of the sign language word.* International Conference on Advanced Mechatronic Systems (ICAMechS), ISBN 978-1-4673-1962-1, 2012.

[8] DONGGE, G., YONG, Z., YULE, Y., GANG, H. *Human segmentation based on disparity map and GrabCut.* International Conference on Computer Vision in Remote Sensing (CVRS), ISBN 978-1-4673-1272-1, 2012.

[9] MIKOLAJCZYK, K., SCHMID, C. *Scale and Affine Invariant Interest Point Detectors.* International Journal of Computer Vision (IJCV), 2008.

[10] KAMENCAY, P., BREZNAN, M., JELSOVKA, D., ZACHARIASOVA, M. *Sparse disparity map computation from stereo-view images using segment based algorithm.* 22nd International Conference Radioelektronika (RADIOELEKTRONIKA), ISBN 978-1-4673-0659-1, 2012.

[11]KAMENCAY, P., BREZNAN, M., JARINA, R., LUKAC, P., ZACHARIASOVA, M. *Improved depth map estimation from stereo images based on hybrid method.* In Radioengineering Journal, Vol. 21, No. 1, ISSN 1210-2512, 2012.

[12] Kinect for Windows SDK, http://msdn.microsoft.com/en-us/library/hh855347.aspx [online, 22.03.2013]

[13] CUDA Toolkit documentation, http://docs.nvidia.com/cuda/index.html [online, 22.03.2013]

[14]ZENG, W. *Microsoft Kinect Sensor and Its Effect.* MultiMedia IEEE , vol.19, no.2, ISSN 1070-986X , 2012.

TRANSCOM 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

Fast Exact String Pattern-Matching Algorithm for Fixed Length Patterns

Ľuboš Takáč
University of Žilina, Faculty of Management Science and Informatics, Department of Informatics,
Univerzitná 8215/1, 010 26 Žilina, Slovakia, lubos.takac@fri.uniza.sk

Abstract. In this paper we describe a dynamic in-memory data structure and an algorithm for fast string pattern matching. Algorithm supports input strings patterns with fixed length which contain any number of undefined characters. We use dynamic multiway tree data structure for indexing with optional deep tree setting. Developed algorithm is implemented in java class FastStringPatternSearch which is easy to use.

Keywords: string pattern searching, dynamic tree index, in-memory data structure, fixed length pattern-matching.

1. Introduction

We are developing generator for word search game. Our generator was slow because of the complexity of string pattern searching in dictionary (about 230 thousand words). We did not find any fast solution for this problem. Existing approaches for example using database was slow because of accessing a hard drive. Using basic indexing, hash or tree mapping is inappropriate for pattern searching and using full-text indexing was also slow. Because of these reasons we develop an own algorithm using dynamic tree especially for this purpose.

The requirements for our solutions are:

- In-memory fast data structure with random access.
- Support for string pattern with fixed length (fixed length means that matched string must be the same length as string pattern) with any number of undefined characters (example of patterns, '--g---h-', '-a-p--g', etc.).
- A matched word must be chosen randomly.
- A matched word must be returned only once.

We need a class with input of dictionary string array, method for searching with input pattern and the method for resetting structure that will allow using all used words again. See UML diagram of class in Fig. 1.

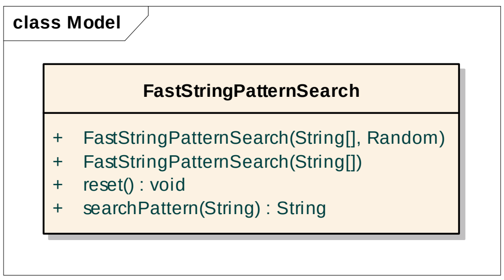


Fig. 1 UML diagram of the desired class FastStringPatternSearch.

The paper is structured as follows: Section 2 (Complexity of problem) explains the complexity of problem. In section 3 (Developed Data Structure and Algorithm) we describe our developed algorithm and data structure. Section 4 (Testing) shows the testing and results of word search game generator using our approach with various parameters and Section 5 (Conclusion) concludes the paper.

2. Complexity of Problem

The basic algorithm for searching strings in array lies in comparing all strings in array with searched string. If we define N as array length, the complexity of this algorithm is O (N). It means than we need at maximum N steps to find out if the string array contains the searched string or not. There also exist better algorithms for example binary search where the string array is sorted and the complexity is O ($\log_2$ (N)) or hash table where the complexity may be even better. With a good hash function in some cases O (1). However, this is not the case for string patterns. If we do not know before creating data structure which positions in pattern are undefined and can be replaced with any character we cannot use binary search nor hash table. There exist also brute force solutions. We can generate all potential combinations of patterns from dictionary that may exist and map it with dictionary strings. In this way we have complexity of pattern searching O ($\log_2$ (N)) using binary search or even better using hash table. But this solution requires a lot of memory. For example string 'PATTERNS' has 256 patterns combinations that can match it. See table 1.

Number of undefined positions	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Example of pattern	PATTERNS	PATT-RNS	PA-TE-NS	-AT--RNS	P-T-E--S	-A--ER--	--TT----	----E---	--------
Binomial coefficient $\binom{n}{k}$	$\binom{8}{0}$	$\binom{8}{1}$	$\binom{8}{2}$	$\binom{8}{3}$	$\binom{8}{4}$	$\binom{8}{5}$	$\binom{8}{6}$	$\binom{8}{7}$	$\binom{8}{8}$
All combinations count	1	8	28	56	70	56	28	8	1
Total combinations count	**256**								

Tab. 1. Computation of all combinations of patterns from eight characters string 'PATTERNS'.

Such a solution will be very fast but memory intensive and also initializing necessary data structures takes some time. We develop an approach which is the compromise between brute force solution and slow basic solution and you can choose required complexity by a parameter.

3. Developed Data Structure and Algorithm

The main data structure in developed algorithm is three-dimensional array of nodes with linked list of strings (more accurate linked list of string id, for saving memory). At first we preprocess input string array (dictionary). We remove duplicate strings, we shuffle array that strings will be randomly ordered, we determine the maximal and minimal word length and we create an alphabet of used characters of all strings in the dictionary. After that we create three-dimensional array of nodes with linked lists. First dimension is alphabet dimension, the second is word length dimension and the third is character position dimension. We put all of the words from dictionary to nodes of the three-dimensional array according to their characters, length and characters positions. For example we put the word "NAUTICAL" into nodes [N][8][1], [A][8][2], [U][8][3], [T][8][4], …, [L][8][8]. See example in Fig. 2. When we will search for pattern "...TI..L", we only look into one of the nodes list [T][8][4] or [I][8][5] or [L][8][8] where we find "NAUTICAL". The reason why we have 8 copies of word "NAUTICAL" is that we do not know before searching which pattern combination will be searched for. We have to find word "NAUTICAL" either using pattern "N……." but also using patterns ".A……", "..U…..", "…T….", ".AU..C.L", etc.

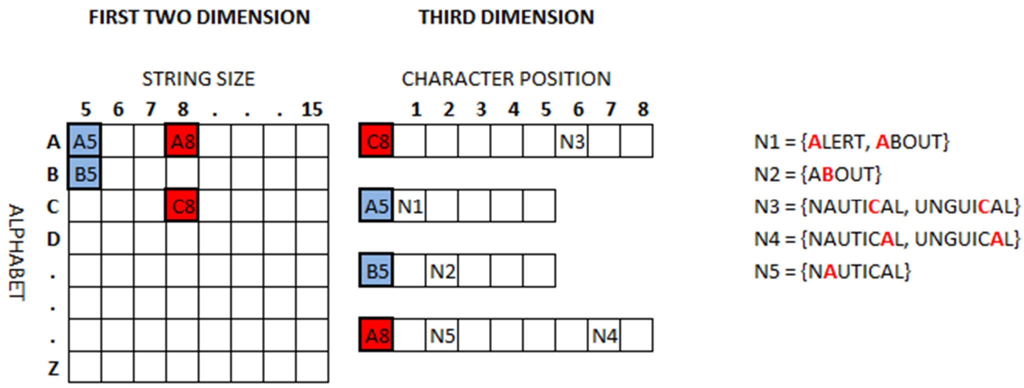


Fig. 2. Example of three-dimensional array. For example string 'NAUTICAL' must be included in node 3 because the length of 'NAUTICAL' is 8 and the character 'C' is at position 6 in the word.

This way we can speed up string pattern searching in our case above 300 times. We tested such approach with dictionary with about 230 thousands of strings. The complexity of searching string pattern with naive algorithm is in the worst case O (N) = 230 thousands steps. In our implemented data structures the nodes of three-dimensional array have linked lists with averagely 690 strings (the node with the maximal length of linked list has 6 839 strings). It means that the worst case to find string pattern is 6 839 steps and the average is 690. We did a test on word search game generator which confirmed above 300 times acceleration (4 Testing, Table 2).

So fast algorithm is anyway for us too slow and we had to improve it. Complexity of algorithm is depending on linked list size. The maximal linked list size of data structure is the worst case of searching pattern. We had to limit linked list size in data structure. We divided all nodes with linked list greater than parameter MaxListSize into new nodes. These new nodes are children of the nodes in the main three-dimensional array. This procedure is applied recursively on new nodes as well until all leaf nodes (note that no leaf nodes in tree may have size greater than MaxListSize and it does not affect complexity of searching) have linked list size less than or equal than MaxListSize parameter. Children of each node are stored in two-dimensional array with dimension character position and dimension alphabet. Nodes are divided in following way:

In previous example we have node N1 which has list size 2 (ALERT, ABOUT). We want to divide it to have all lists with maximum 1 word. We split it based on remaining 4 characters. We create child node with two-dimensional array with dimension alphabet and character positions (4). We put the word "ALERT" into nodes [L][2], [E][3], [R][4], [T][5] and word "ABOUT" into nodes [B][2], [O][3], [U][4], [T][5]. See example of splitting node 1 in Fig. 3.

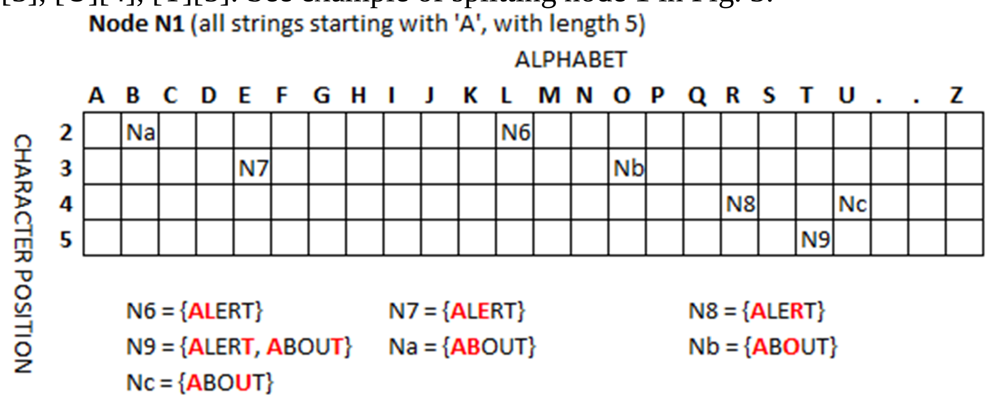


Fig. 3. Example of splitting node 1 (N1 corresponds to the previous figure). Node 1 is splitted into new 7 nodes according to the position and characters of words from node 1. Note that the character position starts at 2 because the first character was indexed in parent node.

But we have still node N9 with 2 words in its list. We have to split this node again based on remaining 3 characters. We create children to node N9 as two-dimensional array with dimension alphabet and character positions (3). We put the word “ALERT” into nodes [L][2], [E][3], [R][4] and word “ABOUT” into nodes [B][2], [O][3], [U][4]. See example of splitting node 9 in Fig. 4.

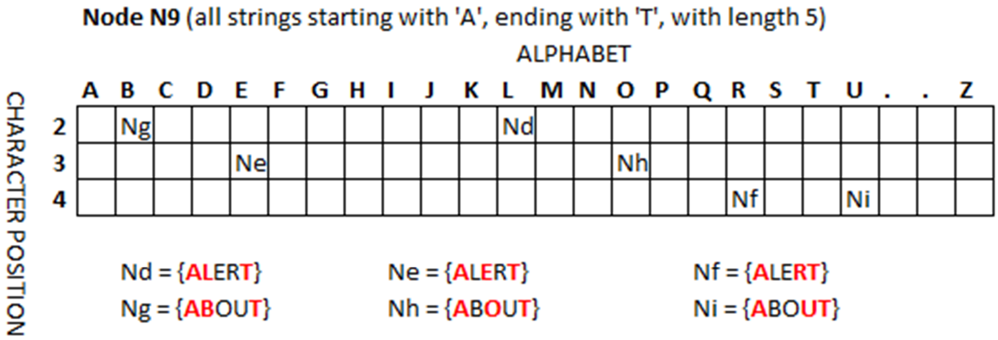


Fig. 4. Example of splitting node 9. Node 9 is splitted into new 6 nodes according to the position and characters of words from node 9. Note that character position starts at 2 and ends at 4 because the first and last characters were indexed in parent node and parents parent node.

Now we have all children node lists with size 1, so we do not have to split it again. Searching in such a structure is described in next section (3.1. Searching Algorithm). In this way, we split all other nodes and we reduce the complexity according our requirements and possibilities.

Such a data structure (multiway tree) is dynamic and flexible because there are new nodes created only when it is necessary. This multiway tree has linked lists on all nodes (not only on leafs) and the complexity depends on the tree depth and linked list size on tree leafs. The depth of the tree is maximally as the maximal string length of preprocessed strings and the linked list size on leaf nodes is given by the parameter MaxListSize.

3.1. Searching Algorithm

Searching algorithm in such data structure is very easy. We can search strings based on patterns in these 2 steps (example of pattern ‘..T.ER..’):

1. Get the first defined character, pattern length and the position of first defined character (T, 8, 3). Get a node of three-dimensional array data structure at [character][length][position] ([T][8][3]). Continue to step 2 with this node.

2. If a node is null, string with this pattern does not exists. **– END.**
   If a node is not null and a node has not children (leaf node) or pattern has no further defined characters, find the first string in a node list which matches the pattern. Return founded string or null if no string matches the pattern. **– END.**
   If a node is not null and a node has children (not leaf node), take the next defined character in pattern (E, position 5) and access two-dimensional array of children nodes of node at element [position][character] ([5][E]), go to step 2 with the given node.

3.2. Mixing Strings

Mixing string is ensured in searching. Each time when a searching algorithm is used and a string which matches the pattern is found on linked list, we remove this string to the end of this list. This operation takes constant operation time and does not significantly affect the complexity of searching. The next requirement of the algorithm is that searched strings cannot repeat. We can ensure it by excluding used strings from linked lists when they are founded. But a string may occur in a lot of nodes, not only in node where we found it and moreover, when we generate more than one word game we need the whole dictionary again and creating and preprocessing of new data structure would take a lot of time. Because of this we create a simple array of boolean values which indicates if *i-th* string in dictionary is used or not. The only changes in searching algorithm is that we check it when we are searching in nodes linked list and after we found unused string matching the pattern we write it to this array. Both operations have constant complexity and do not slow down searching significantly. By resetting the array values we can use all words from dictionary again and even by using the same patterns we get other strings because of mixing lists of strings.

3.3. Algorithm Complexity

Complexity of searching algorithm depends on parameter MaxListSize. If we set MaxListSize to N the resulting complexity of our algorithm for pattern P with length L is O (L+N) in the worst case. The average complexity is much lower. Setting parameter MaxListSize to low values cause higher memory requirements (see 4. Testing in the next section). The challenge is to set this parameter according to the application needs.

4. Testing

We tested implemented algorithm on word search game generator. We generated five thousands word search games of size 25x25[1] with dictionary of size 224 942 words and we used developed algorithm with a different parameter of MaxListSize. We monitored data structure initialization time, generating time and memory consumption. For testing we used HP ProBook 6550b with configuration Win 7 Professional 64bit, Intel® Core ™ i5 CPU M450 2cores 2.40GHz, 4GB RAM, Java 7. Testing results are shown in Table 2. For comparison, the rate of the naive algorithm is at the bottom of the table.

MaxListSize	Initializing time (s)	Generating time (s)	Memory consumption (MB)
Unlimited	1,508	989,643	86
5000	2,726	839,294	101
1000	4,843	400,539	265
500	7,062	324,728	340
100	16,141	279,410	808
Naive algorithm[2]	0,095	381 073,600	15

Tab. 2. Table with testing results. Memory consumption includes only resources for data structure.

5. Conclusion

We develop a dynamic in-memory data structure and an algorithm for fast string pattern matching. Algorithm supports input strings patterns with fixed length which contain any number of undefined characters. We have achieved more than thousand times acceleration in comparison with naive algorithm for pattern searching. We use dynamic tree data structure indexing with optional deep tree setting. Developed algorithm is implemented in java class FastStringPatternSearch which is easy to use.

References

[1] EDMONDS, J., *How to Think about Algorithms*, Cambridge University Press, 2008, ISBN 978-0-521-84931-9.

[2] HORSTMANN, C., *Big Java*, 3rd edition, John Wiley & Sons, Inc., 2008, ISBN 978-0-470-10554-2.

[3] The Java™ Tutorials, 2013, available at http://docs.oracle.com/javase/tutorial/.

[1] Number of rows and columns of word search game matrix.

[2] Basic solution, without using developed data structure. All strings are in a linked list and pattern is step by step compared with strings in list until an acceptable string is found.

Transcom 2013, 24-26 June 2013
University of Žilina, Žilina, Slovak Republic

The Impact of Radiowave Polarization, Frequency and Rain Intensity on the Satellite Signal Reception in the Area of Kielce City

*Jacek Łukasz Wilk
*The Kielce University of Technology, Faculty of Electrical Engineering, Automatics and Computer Science, Al. Tysiąclecia Państwa Polskiego 7,
25-314 Kielce, Poland, jwilk@tu.kielce.pl

Abstract. In this paper we will present the fundamental information about the international research project COST IC0802, which is the point of reference in this study. "The main benefit of this […] COST Action will be (prepare – note J.W.) the integrated assessment of the radio channel issues relevant to a GIN providing communication, positioning, meteorology and remote sensing services. Concrete outputs will include a set of required dedicated models and tools, as well as a common experimental database of measured results suitable for modelling, validation and system simulation" [1]. The article covers attenuation of satellite line-of-sight radio links according as rain intensity, frequency and polarization of radiowaves in the region of Kielce city in Poland in accordance with the recommendations of ITU. The long-term 1-min average rain-rate characteristics was estimated to be 34,4 mm/h. Collected data were used to determine the impact of the rainfall intensity, polarization and radio waves frequency – including the frequencies used in a group of satellite Hot Bird – on the quality of received satellite signal in this area (the chosen experimental data are presented).

Keywords: Atmospheric attenuation, Influence of rainfall intensity, Radiowave polarization and frequency on the propagation of microwave signals, Solar activity.

1. International Research Project COST IC 0802

The Kielce University of Technology was the member of the international research project COST Action IC0802 – *Propagation Tools and Data for Integrated Telecommunication, Navigation and Earth Observation Systems*, whose aim is to analyse the impact of weather conditions on the quality of satellite transmission [2]. Development works on COST Action IC0802 focused on the free space propagation, meteorology and developing a coordinated set of models in order to improve the design and performance of Global Integrated Networks (including *GMES & Disaster Management and Relief*). The target architectures include mobile and fixed, satellite and terrestrial communication systems (including optical links), satellite navigation systems and Earth Observation systems [1]. The establishment of this system seems to be a valuable initiative to deliver the integrated satellite services to users at the highest level. The activities of the research groups of the project COST IC0802 by gathering the specialists from European countries from various academic centers, and many institutions, such as: ESA-ESTEC, TEC-EEP, DLR, ONERA and multi-faceted (depending on weather conditions) analysis of the electromagnetic wave propagation in free space tends toward the tangible benefits for society. It can be used for manufacturers and users of the satellite equipment (TV, radio, satellite telephones, GPS systems and many others) [1, 2]. It turns out that the data transmitted by satellite systems due to free space propagation are significantly damping (especially in the troposphere, where weather conditions on the Earth are build). "The action will also expand the European network of experts by multidisciplinary interaction and will create the competence and support needed for the development and deployment of such a system. The collaboration between remote sensing and radiowave propagation experts will improve modelling by assessing the physical fundamentals of radiowave propagation using experimental climatic data and including results from new Earth observation missions and new Numerical Weather Forecast models […] Up to now, radio channel modelling has been performed separately for each type of […] (services or links e.g. satellite navigation systems, terrestrial mobile communications, fixed satellite communications, remote sensing systems, and so on – note J.W.)" [1]. See also: COST Actions 255, 258 and ESA-Projects in collaboration with COST Action IC0802 (D. Nörenberg, *A Novel Ground-Based Microwave Radiometer for High Precision Atmospheric Observations Between 10 and 90 GHz (ATPROP - Atmospheric Propagation and Profiling System)*; A. Danklmayer, *Study of Tropospheric Propagation Effects...*; P. Cannon, *Measurements and Simulation of VHF and UHF...*; J.M. Riera, J.M García-Rubia, P. García-del-Pino, A. Benarroch, *Derivation of Rain Attenuation from Experimental Measurements of Drop Size and Velocity Distributions;* N. Jeannin, L. Castanet, X. Boulanger, F. Lacoste, *Study and modelling of tropospheric attenuation…*) [2]. Ongoing research may be used in the further to analyse the propagation of microwave satellite signal (on the direction of the satellite-Earth and Earth-satellite) in the area of Kielce with the best possible quality (the lowest attenuation) including the ionosphere phenomena. The research may acquire special significance in the long term – for example, at a time when the Sun reaches the peak of it activity (may 2013). In early 2010, NASA sent probe into orbit to predict changes in solar activity in the future. The probe will be monitoring the movement of the masses inside the Sun and magnetic field changes. According to the scientists, solar activity will be increased until May 2013. Many of them said that the Sun would damp microwave waves more effectively from geostationary satellites. The results will be used to determine the link budget.

1.1. The Main Scientific Goals of the Project

To ensure satisfactory operation of satellite links in the direction of the satellite-to-earth it is essential to the power of the transmitter. In practice, this power depends on many factors such as geographical coordinates, terrain, frequency, aperture antennas, the type and location of the antenna, the signal attenuation in space – in the atmosphere, the required level of availability of bandwidth. A problem is to estimate the effect of the troposphere on radio-wave propagation in a conductive environment. A properly calculated link budget requires consideration of extreme weather conditions (e.g. precipitation) in accordance with the recommendations of ITU. In practice, precipitation and other adverse phenomena (e.g. storms) – may cause temporary interruptions and lack of communication between the terminal and the satellite [3, 4]. So it is necessary to determine the required transmitter power (or antenna performance) as a function of this factors. This research include the analysis of the conditions of propagation of electromagnetic waves in an open environment and problem of free space propagation as an ideal, open and free of any obstacles transmission medium, a completely homogeneous which is not absorbing electromagnetic field energy. One of the scientific goals also include the impact of weather conditions in troposphere on the quality of satellite signal reception – especially influence of precipitation, solar activity, clouds and atmospheric gases on the propagation of microwave satellite signals. The most important effects in this layer of the atmosphere are tropospheric refraction and tropospheric attenuation (due to precipitation, fog etc.). The emitted radio wave is scattered by the air particles in the troposphere and its energy is absorbed by these molecules. The quality of the signal is also degraded by the natural sources of noise, such as atmospheric discharges, ionization noise, thermal noise in large part [4]. Moreover relationship between signal loss (due to absorption) and received frequency and relationship of signal attenuation versus frequency was thoroughly investigated. It depends on the local rainfall intensity (according to ITU-R model). The aim of this study was to analyse a geometric model of the beam on the satellite path in rainy weather. Attenuation caused by rain was analysed at the 0.01 GHz intervals. The polarization effects would be also analysed and quality parameters would be determined. In practice, the signal quality can be estimated on the basis of a number of qualitative indicators which we can determine by the evaluating microwave links. In addition, we built a specialized test set-up for measurements of the quality of microwave signals. It will be connected to testing station and can be used to measure actual microwave links in rainy weather in the area of Kielce. In addition it can be used to measure the influence of solar activity on the quality of the satellite signal. So the modelling results can be compared with actual real-time data. The project has the following main research goals: (1) Determination of satellite signal attenuation due to rain in the area of Kielce; (2) The comprehensive analysis of the influence of the weather conditions on the quality of the satellite signal transmission (an increase in the signal noise, level and degradation etc.); (3) Determination of relationship between the frequency of the electromagnetic wave and the polarization of the radio waves and determination which of the polarized waves are more attenuated (influence of hydrometeors); (4) Observation if increase of solar activity can be related to changes in frequency and the intensity of rainfall in the area of Kielce. In practice, this research focused on optimizing the reception of the microwave signal transmitted from the group of geostationary communications satellite Hot Bird. The research work „The analysis of factors affecting the propagation of microwave signals in the troposphere and satellite reception in terms of rainfall in the area of Kielce", which is connected with this topic and realized by The Kielce University of Technology can solve the following research problems: (1) The aspects analysis of the use of the geostationary orbit for the propagation of microwave; (2) The research into the propagation of microwaves in the troposphere in the area of Kielce (the conditions of microwave propagation in the troposphere in Kielce); (3) Modelling and actual measurements of microwave links in rainy weather in the area of Kielce; (4) Work out of research results and its interpretation.

1.2. Importance of the Research Work

Research is concerned with acquiring knowledge about:

I. The analysis of the geostationary satellite systems. In addition, the frequencies of operations of satellite systems will be discussed and free space propagation (ideal and normal).

II. The analysis of the radiowave propagation in the troposphere in the area of Kielce (e.g. tropospheric refraction, the natural sources of the microwave signals attenuation – particularly the influence of the rainfall on the propagation of the microwave signals).

III. Study and modelling of microwaves links tropospheric attenuation in terms of rainfall in the area of Kielce. The aim of this study is to analyse the influence of the intensity of rainfall in the area of Kielce on the signal attenuation and relationship between frequency and the polarized microwaves.

The second aim is to establish data and develop new methods with application of model of polynomial regression. From a practical point of view the technical parameters changes change the quality parameters of the received satellite signal [5, 6, 7, 8]. The research can be used for the selection of appropriate equipment in order to get the best of reception microwave signals from the group of Hot Bird satellite in the area of Kielce and estimate the signal margin which can compensate the adverse weather conditions. It can be also used to minimize the risk of loss the satellite signal. To our knowledge this is the first study of this type in Poland [2].

1.3. Research strategy

The subject of research was forced the use of several research strategies. The usability of each method has been evaluated through its use to achieve the specified aims of researches. The elements of system analysis have been used in conceptual analysis and research strategy because the aim of this project is to analyse the impact of weather conditions on the propagation of satellite signal in the troposphere in the area of Kielce. The theoretical aspects use of the geostationary orbit for the microwave propagation and the conditions for propagation in the troposphere was used in the task I and II (method of literature review and method of individual cases). The other methods which will be used for this project are: (1) Experimental method and statistical analysis was used to model of microwave links in rainy weather in the area of Kielce; (2) Comparative method was used to comparative analysis of the results of measurements; (3) Institutional-legal method (The research was conducted in accordance with the recommendations of ITU).

2. Measurement Examples in the Area of Kielce City (*$R_{0,01}$*=34,4 mm/h)

Increase in the frequency affects on the increase in the signal attenuation L_d, noise increase due to precipitation W_{sz} and downlink degradation *DND* for each polarization (horizontal and vertical).

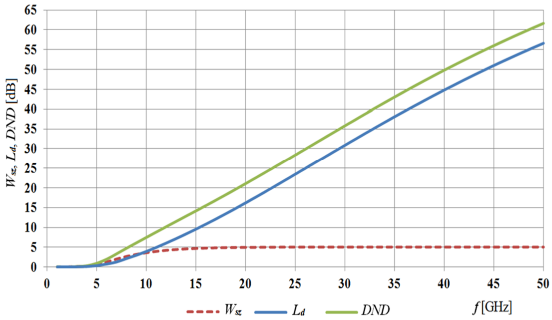


Fig. 1 W_{sz}, L_d, *DND* [dB] vs frequency [GHz] (pol. H)

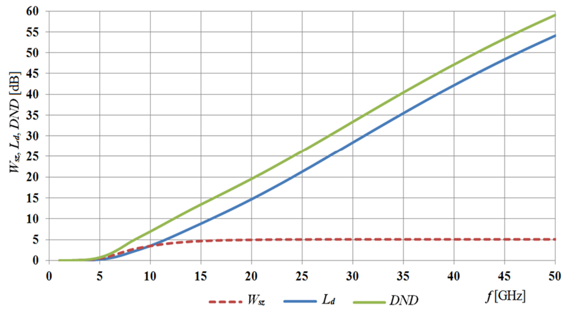


Fig. 2 W_{sz}, L_d, *DND* [dB] vs frequency [GHz] (pol. V)

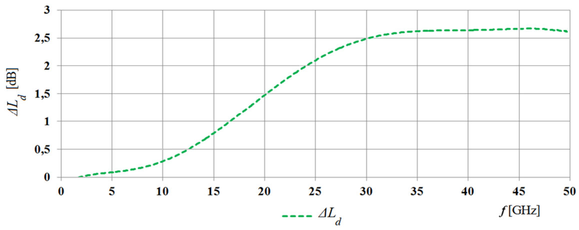


Fig. 3 Increase in L_d [dB] vs frequency [GHz] (pol. H-V)

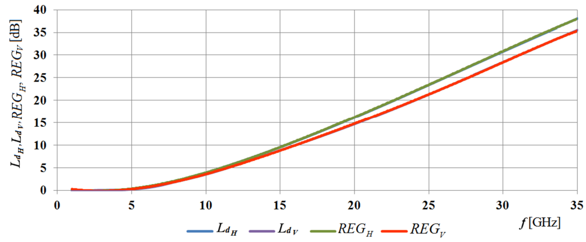


Fig. 4 Multiple results by using the models of polynomial regressions).

This model uses bilinear interpolation to obtain an improved evaluation for a selected location from it's grid neighbours and is based on 40 years of records collected by ESA. The proposed graph of attenuation due to rain and rain attenuation statistics this city in the range of frequency 1-35 GHz is the following:

(1) for horizontal polarization:

$$REG_H = -7{,}612 \cdot 10^{-9} f^6 - 4{,}933 \cdot 10^{-8} f^5 + 6{,}594 \cdot 10^{-5} f^4 - 4{,}287 \cdot 10^{-3} f^3 + 0{,}131 f^2 - 0{,}629 f^1 + 0{,}731;$$

(2) for vertical polarization:

$$REG_V = -4{,}967 \cdot 10^{-8} f^6 + 3{,}753 \cdot 10^{-6} f^5 - 5{,}066 \cdot 10^{-5} f^4 - 2{,}864 \cdot 10^{-3} f^3 + 0{,}122 f^2 - 0{,}637 f^1 + 0{,}779.$$

In practice, signal attenuation depends on the radiowave polarization, frequency and rain intensity [4]. The maximum absolute difference in attenuation calculated on the basis of the achieved formula REG_H and actual data Ld_H for horizontal polarization is equal 0,32 dB (for frequency 35 GHz) and equals 0 dB (for 4,6 GHz, 8,3 GHz, 8,4 GHz, 8,6 GHz), whereas for vertical polarization is equal 0,53 dB (for frequency 35 GHz) but in most cases, it does not exceed the value of 0,1 dB. Because the achieved results (fig. 4) coincide with the attenuation curves for horizontal (fig.1) and vertical polarization (fig.2), it can be assumed that they provide a good estimate of L_d attenuation in the city of Kielce.

References

[1] Memorandum of Understanding for the implementation of a European Concerted Research Action designated as COST Action IC0802 'Propagation tools and data for integrated Telecommunication, Navigation and Earth Observation systems', http://w3.cost.eu/fileadmin/domain_files/ICT/Action_IC0802/mou/IC0802-e.pdf

[2] WILK J. Ł., Współpraca naukowa w ramach Projektu Europejskiego COST IC0802, [w:] Wschód i Zachód w wymiarze globalnym. Doświadczenia z przeszłości a perspektywy na przyszłość, (red.) M. Miłek, G. Wilk-Jakubowski, R. S. Brzoza, Wydawnictwo Stowarzyszenia Współpracy Polska-Wschód. Oddział Świętokrzyski, Kielce 2012

[3] WILK J. Ł., Wybrane zagadnienia dotyczące szumów w komunikacji satelitarnej, [w:] Rola Informatyki w Naukach Ekonomicznych i Społecznych. Innowacje i implikacje interdyscyplinarne, pod red. T. Grabiński, Wydawnictwo Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach, t. II, Kielce 2010

[4] WILK J. Ł., Naturalne źródła szumów w transmisji satelitarnej, [w:] Rola Informatyki w Naukach Ekonomicznych i Społecznych. Innowacje i implikacje interdyscyplinarne, pod red. T. Grabiński, Wydawnictwo Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach, t. II, Kielce 2010

[5] WILK J. Ł., The influence of the dish efficiency on the quality of a satellite signal, [w:] TRANSCOM 2011, 9th European Conference of Young Research and Scientific Workers, Proceedings, Section 3 – Information And Communication Technologies, EDIS-Žilina University publisher, Žilina 2011

[6] WILK J. Ł., The influence of the antenna aperture on the quality of a satellite signal, [w:] XII International PhD Workshop OWD 2010, pod red. G. Kłapyty, PTETiS, vol. 28, Wisła 2010

[7] WILK J. Ł., The influence of the antenna parameters on a digital satellite signal reception, CEEPUS CII-CZ-0404-01-1011-M-43894, Brno 2011

[8] WILK J. Ł., The measurment processing of satellite signal, CEEPUS CII-CZ-0404-01-1011-M-43894, Brno 2011

TRANSCOM 2013

Proceedings, Section 3

Published by University of Žilina
First Editions
Printed by EDIS-Žilina University publisher
Printed in 65 copies

ISBN 978-80-554-0692-3