UNIVERSITY OF ŽILINA

TRANSCOM 2009

8-th EUROPEAN CONFERENCE
OF YOUNG RESEARCH AND SCIENTIFIC WORKERS

PROCEEDINGS

SECTION 3

INFORMATION AND COMMUNICATION TECHNOLOGIES

ŽILINA June 22 - 24, 2009
SLOVAK REPUBLIC

UNIVERSITY OF ŽILINA

TRANSCOM 2009

8-th EUROPEAN CONFERENCE
OF YOUNG RESEARCH AND SCIENTIFIC WORKERS

under the auspices of

Prof. Ing. Ján Mikolaj, PhD.
Minister of Education, Slovak Republic

&

Prof. Ing. Ján Bujňák, PhD.
Rector of the University of Žilina

SECTION 3

INFORMATION AND COMMUNICATION TECHNOLOGIES

ŽILINA June 22 - 24, 2009
SLOVAK REPUBLIC

Edited by Soňa Stránska, Zuzana Dubcová and Peter Brída

ISBN: 978-80-554-0039-6

TRANSCOM 2009
8-th European conference of young research and scientific workers

TRANSCOM 2009, the 8th international conference of young European researchers, scientists and educators, aims to establish and expand international contacts and co-operation. The main purpose of the conference is to provide young 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 research workers 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: Baumann S. (D), Cenek P. (SK), Droździel P. (PL), Janáček J. (SK), Jánošíková Ľ. (SK), Kremeňová I. (SK), Novák A. (SK), Palúch S. (SK), Rievaj V. (SK), Řezáč M. (CZ), Schnieder E. (D), Surovec P. (SK), Žarnay P. (SK)

2. ECONOMICS AND MANAGEMENT

Scientific committee: Bartošová V. (SK), Blašková M. (SK), Březinová O. (CZ), Duncan F.H. (USA), Hittmár Š. (SK), Kos B. (PL), Kucharčíková A. (SK), Lyakin A. (RUS), Rostášová M. (SK), Rybakov F. (RUS), Strenitzerová M. (SK), Strišš J. (SK), Tomová A. (SK)

3. INFORMATION AND COMMUNICATION TECHNOLOGIES

Scientific committee: Bärwald W. (D), Dado M. (SK), Drozdová M. (SK), Hanuliak I. (SK), Keil R. (D), Klimo M. (SK), Kolev P. (BG), Kotsopoulos S. (GR), Madleňák R. (SK), Matiaško K. (SK), Spalek J. (SK), Vaculík J. (SK), Vaculík M. (SK), Vrček N. (HR)

4. ELECTRIC POWER SYSTEMS. ELECTRICAL AND ELECTRONIC ENGINEERING

Scientific committee: Altus J. (SK), Blažek V. (D), Čápová K. (SK), Dobrucký B. (SK), Dodds S.J. (GB), Santarius P. (CZ), Vittek J. (SK)

5. MATERIAL ENGINEERING. MECHANICAL ENGINEERING TECHNOLOGIES

Scientific committee: Adamczak S. (PL), Bokůvka O. (SK), Borkowski S. (PL), Dzimko M. (SK), Guagliano M. (I), Kunz L. (CZ), Meško J. (SK), Nicoletto G. (I), Palček P. (SK), Skočovský P. (SK), Takács J. (H)

6. MACHINES AND EQUIPMENTS. APPLIED MECHANICS

Scientific committee: Dekýš V. (SK), Gerlici J. (SK), Kalinčák D. (SK), Malenovský E. (CZ), Medvecký Š. (SK), Merkisz J. (PL), Nemček M. (CZ), Pawelczyk M. (PL), Sága M. (SK), Zapoměl J. (CZ), Žmindák M. (SK)

7. CIVIL ENGINEERING

Scientific committee: Bouchair H. (F), Dmitrovskaja L. (RUS), Garbuz A. (UA), Ižvolt L. (SK), Melcer J. (SK), Ivanov J. (BG), Teleman E.C. (RO), Vičan J. (SK), Zolotov M. (UA)

8. SOCIAL SCIENCES

Scientific committee: Banáry B. (SK), Baštinec J. (CZ), Cabanová V. (SK), Gulová L. (CZ), Kráľová Z. (SK), Růžičková M. (SK), Šindelářová J. (CZ), Vikoren V. (N)

9. SECURITY ENGINEERING. FORENSIC ENGINEERING

Scientific committee: Danihelka P. (SK), Horák R. (CZ), Navrátil L. (CZ), Poledňák P. (SK), Reitšpís J. (SK), Seidl M. (SK), Šenovský M. (CZ), Šimák L. (SK), Vasiliev D. (BG), Zamiar Z. (PL)

ORGANIZING COMMITTEE

CHAIRPERSON Bokůvka Otakar

EXECUTIVE SECRETARY Vráblová Helena

MEMBERS Bábel J., Belan J., Braciník P., Brída P., Brumerčík F., Dobrotková M., Dubcová Z., Frajtová-Michalíková K., Harušinec J., Hnátová Z., Houšková L., Imrišková E., Jánošíková G., Kačiaková B., Kardoš M., Kormancová M., Krasňan M., Kuzmová M., Ladovský T., Ližbetinová L., Mihalov K., Močková M., Mráz M., Murín M., Neslušan M., Ondrejka R., Pachová S., Pavelková I., Pilát P., Potkan T., Remek Ľ., Smetana M., Sršníková D., Stránska S., Štofková K., Tulejová L., Valentíková E., Varmus M., Vaško A., Vaško M., Záborský M.

SECTION 3 INFORMATION AND COMMUNICATION TECHNOLOGIES

REVIEWERS:

Benčo Miroslav
Brída Peter
Drozdová Matilda
Dúha Ján
Hrianka Miroslav
Hronec Rudolf
Chochlík Matúš
Janáček Jaroslav
Jarina Roman
Klimo Martin
Kovalík Štefan
Kremeňová Iveta
Madleňák Radovan
Martincová Penka
Matis Peter
Miček Juraj
Paralič Martin
Palúch Peter
Počta Peter
Rievaj Vladimír
Schwartz Ladislav
Segeč Pavel
Smieško Juraj
Surovec Pavel
Vaculík Juraj
Vaculík Martin
Vestenický Martin
Vestenický Peter
Wieser Vladimír

Note:

Authors are responsible for language contents of their papers

CONTENTS

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Residential Services, Comprehensive Network Architecture and Devices for Home Networking

*Bohumil Adamec, *Vladimír Hottmar
*University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 8215/1, 01026 Žilina, Slovakia,
{Bohumil.Adamec, Vladimir.Hottmar}@fel.uniza.sk

Abstract. This paper describes the reference network architecture, where the main residential services are identified. It reveals the home reference architecture and places the home network into an end-to-end network architecture. This paper also presents the home access gateway role and reviews some hardware architectures.

Keywords: Multimedia services, interactive services, end-to-end network architecture, home gateway, residential gateway, home networking.

1. Introduction

Connecting each house to broadband access networks represents a great opportunity to offer added-value services and broadband Internet access to residential users. The provisioning of broadband networking services to residential users requires a comprehensive broadband end-to-end digital network infrastructure spread from the service provider to the customer premises. This paper describes the reference network architecture and discusses the modern multimedia services.

The major obstacle in the home-network area is the isolation of residential in-home networks. The requirement for interoperability between isolated home network and interworking between different home appliances over a home-network structure, along with the demand for high-speed Internet access, generated the need for a new device, the home access gateway. This paper presents the home access gateway role and reviews some hardware architectures.

2. Residential Services

A grouping of future home services is shown in figure 1. It consist of home communications, automation, security, small office home office (SOHO) and entertainment services [1; 2; 3].

2.1. Communication Services

Evolution of today's narrowband communication services is expected to be among the first service segments to develop. Future communication services will include applications like shared Internet access from multiple home PCs and value-added voice services, such as voice over IP (VoIP), videophone and so on.

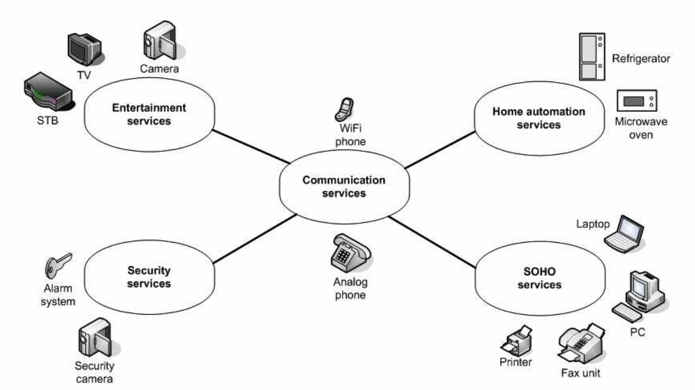


Fig. 1. Future home services.

2.2. Automation Services

Remote and unified control of smart, network-enabled, consumer-electronic devices and appliances will be a key service for future home networks. Modern future home automation services provide the ability to connect to your home network remotely and control your home devices.

2.3. Security Services

The convergence of security systems with home and access networks, along with modern monitoring and sensing devices represents a new market potential. Subscribers will be able to monitor their homes remotely via Internet browsers or mobile PDA, or receive automatic e-mail if something happens.

2.4. Small Office and Home Office Services (SOHO)

The number of telecommuters working from small offices and home offices is increasing rapidly. To work efficiently from home, telecommuters must establish a SOHO network, access the corporate Intranet properly and securely, and establish cost-effective voice communications.

2.5. Entertainment Services

Key service for home entertainment segment is video on demand (VoD) and services that include video streaming such as multiplayer network games.

3. Comprehensive Network Architecture

The provisioning of broadband networking services, including emerging, interactive, streaming multimedia services to residential users, requires a comprehensive broadband digital network infrastructure spread from the service provider to the customer premises [2; 4]. The architecture is organized into three segments:

- The content or service provider segment.
- The central office.
- The home network.

3.1. The Content or Service Provider Segment

The content or service provider segment is responsible for preparing, storing, and manipulating the multimedia content. Videos are received mainly in analog format from various sources (for example satellite connections, analog wireline or wireless broadcasting networks, or cable TV networks). Analog streams are converted to digital MPEG format via MPEG real-time video encoders (RTE). Subsequently, the digital video streams are multiplexed over Internet Protocol via multiplexer (MUX) and distributed over the core network. A service manager host synchronizes the video encoding, multiplexing and distribution process.

3.2. The Central Office

The multimedia content is distributed over the core network to the central office. At the central office are considered various servers:

- A video cache server for temporal caching of the most popular films.
- A session and resource manager server, responsible for setting sessions and allocating the required network resources.
- A billing system server, that provides billing information.

3.3. The home network

Finally, at the user side, customer premises equipment is required to decode and decompress the digital signal and handle upstream communications. Digital set-top boxes coupled with TV screens or multimedia PCs are expected to be appropriate customer premises equipments. The point of connection between the home network and the access network is the digital residential gateway.

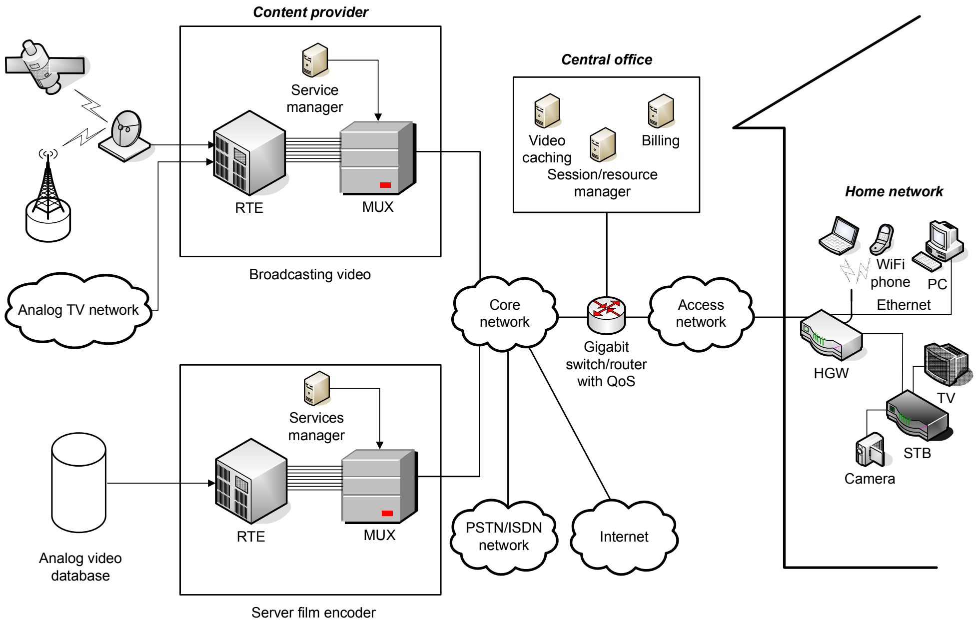


Fig. 2. Comprehensive network architecture.

4. Residential Gateway

Digital home gateway is compact and scalable platform that aggregates multiple services (voice, data, and Internet) and interconnection device between the access and in-home networks providing network termination, device interoperability, and service delivery [5]. Home gateway may be a key element in the deployment of broadband services to residential users. This chapter presents two approaches to the home gateway hardware architecture – one modular and the other compact.

4.1. Modular Hardware Architecture

The modular architecture, shown in figure 3, targets the small offices and home offices and the small-business market.

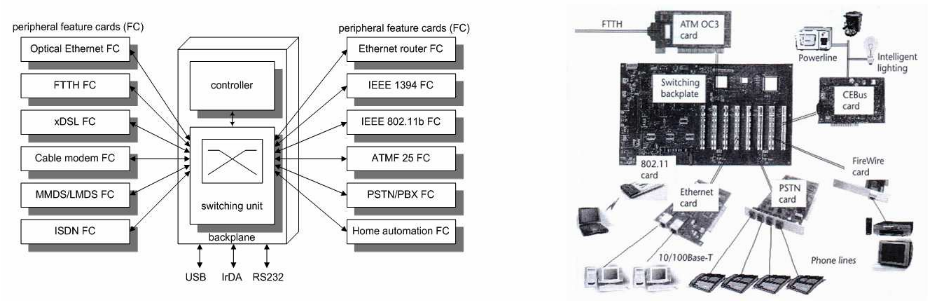


Fig. 3. Modular hardware architecture.

It consists of a switching backplane, which interconnects various peripheral feature cards. Each feature card provides two-way communication between network devices and is responsible for physical layer transmission and control, analog-to-digital signal conversions and vice versa, protocol translations, and data communications with the backplane in a common data format. The backplane provides adequate switching capacity to support most existing or forthcoming indoor and access interfaces. A controller feature card provides the overall supervision and control. It controls the switching unit, coordinates the signaling communications between the feature cards and provides for a single point of home gateway management. Finally, a number of narrowband interfaces (such as USB, Bluetooth) and input/output interfaces (such as LCD display, console, IR/RF remote control) may be directly attached either to the backplane or to the controller. The design philosophy behind the modular architecture is to create an expandable home gateway that can accommodate any viable interface or protocol. Whether, the feature card is simple or quite complex, the modular home gateway is able to support it. Moreover any new, emerging or evolutionary interface can be introduced on-demand, according to user needs or service provider offerings [2]. Although the modular design is a very flexible and expandable, it may lead to a quite expensive and complex multiprocessor system.

4.2. Compact Hardware Architecture

As an alternative, compact home gateway architecture is shown in figure 4. The compact home gateway targets only the residential-user market [2]. In the compact home gateway version, a set of modules and interfaces is considered mandatory or default. This set includes the RAM and flash memories, some standard interfaces such as USB, Ethernet and at least one access interface. Vendor differentiation will mainly be based on the processor speed, system functionality, selection of the additional modules and the access and in-home network interfaces.

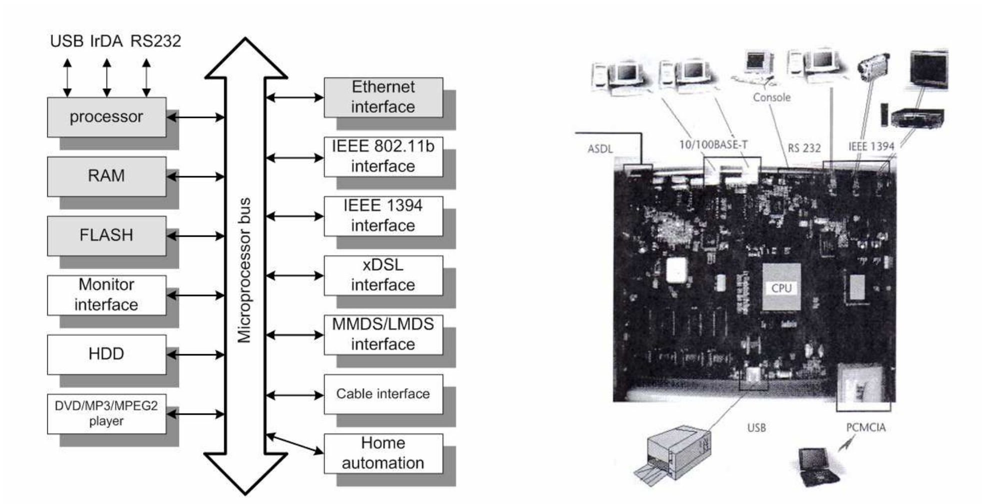


Fig. 4. Compact hardware architecture.

Of course, the compact residential gateways are less modular and scalable; however, they are also cheaper and this is a very important factor in the consumer-electronics area. The compact architecture is shown in figure 4.

Acknowledgement

This project was supported by the VEGA grant No. 1/0375/08 of Ministry of Education of the Slovak Republic, EU.

References

[1] ZAHARIADIS, T. B. *Interactive multimedia services to residential users.* IEEE Communications Magazine, Vol. 35, No. 6, pp. 61-68, June 1997.

[2] ZAHARIADIS, T. B. *Home networking technologies and standards.* Boston, London: Artech House, 2003.

[3] PAJDUŠÁKOVÁ, D. *Current HFC networks, they architecture and 3P service.* In proceedings of 9-th international conference RTT, ISBN 978-80-227-2939-0, September 2008.

[4] HOTTMAR, V., SCHWARTZ, L., TRSTENSKY, D. *End user box for interactive cable television*. IEEE Transaction on Consumer Electronics, Vol. 53, pp. 412-416, May 2007.

[5] LAVRENCE, V. *Digital gateways for multimedia home networks.* Telecommunications System Journal, August 2003.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Radio Map Framework for GSM Positioning

*Peter Brída, Jozef Benikovský
* University of Transport, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia, peter.brida@fel.uniza.sk, jbenikovsky@gmail.com

Abstract. This paper introduces radio map framework for creation of GSM (Global System for Mobile communications) network radio maps. These radio maps are structured in the way that they can be utilized later on as a data source for fingerprinting-based localization methods. Attributes of fingerprinting localization method as well as particular radio signal parameters are explained. The framework handles data measurement, data transfer and data visualization scenarios. Data are basically various radio signal parameters coupled with position coordinates. These data make up a database, which allows centralized storage and opens a path to service-oriented localization architecture in the future. Visualization provides means to verify radio map quality, distribution and density in perceptible and user friendly way as well as means to present and publish achieved results. Entire system is aimed to provide usability by multiple users and further optimization. Centralization allows easy implementation of changes and system maintenance at one place.

Keywords: Localization, fingerprinting, radio map, GSM, received signal strength.

1. Introduction

There was undoubtedly great progress made in the area of localization systems and services over last few years and current market provides numerous navigation and localization devices. Location based services attract more subscribers every year (see [4]) and have great potential. However, there are also drawbacks of this technological advance. Growing number of standards, technical recommendations and resource exhaustion (e.g. radio wave frequency spectrum) reduces possibilities for development of new and independent technologies.

This work is aimed to show solution that uses existing technologies and algorithms and combines them into **radio map framework** for localization in GSM network. The framework is basically a set of measurement algorithms, communication interfaces and database of various signal parameters usable by fingerprinting method to provide localization itself. Fingerprinting method is briefly described in chapter 1.1 below. Framework architecture is described in Chapter 2.

1.1. Fingerprinting

There are many ways to determine position of mobile station within cellular network, for instance cell identification, angle of arrival, time of arrival, received signal level or their mutual combinations [3]. But all of these suffer from inaccuracy caused by estimation of distance between Mobile Station (MS) and Base Stations (BSs) influenced by radio channel effects such as multipath propagation or delay spread.

The fingerprinting method even benefits from aforementioned radio channel properties. It is based on unique "fingerprints" of radio signal properties measured at certain spots. Accuracy of the method in radio networks is according to [1] and [2] determined by two factors. Firstly, signal properties vary very much at relatively small area. For instance, in few meters range, signal from a BS can attenuate, lose or be replaced with stronger one. Secondly, these signals are relatively constant in time. It allows data gathering and their use in future. These fingerprints form the radio map shown in Fig. 1 and described below.

A disadvantage of the method is sensitivity for environment changes – movement of pedestrians and cars, construction and demolition of buildings or weather conditions – they altogether affect signal properties. It is necessary to update the map, but basically, buildings and walls affect the signal most of all and therefore update is not needed very often.

Fingerprinting method consists of two steps. At first, radio map for particular area is created. It is basically a database of spots with known position (coordinates) coupled with various signal properties, e.g. received signal level (RxLev), signal angles or propagation time. This step is called *offline* phase. The offline phase is to be facilitated by the radio map framework presented in this paper.

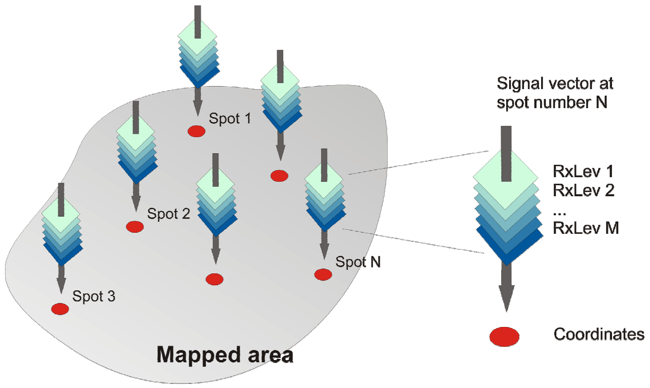


Fig. 1. Radio map for fingerprinting using RxLev.

After the radio map is created, second phase can take place. MS measures signal properties at unknown spot. Then the radio map is searched to find a best match from existing spots. This step is called *online* phase. The online phase is not part of this paper.

2. Radio map framework architecture

The radio map framework consists of two parts - gauge subsystem and radio map itself. Gauge subsystem is used to measure various signal parameters and transfer them to central database. Radio map is actually the central database and gathers all information needed for fingerprinting online phase.

2.1. Gauge subsystem

Fig. 2 shows components that make up gauge subsystem. *BS_1*, *BS_2* and *BS_M* in Fig. 2 represent GSM base transceiver stations. *GPS satellite* represents Global Positioning System (GPS) network utilized to read "precise" position for radio map spots. *Gauge* is a device used to measure GSM as well as GPS signals, interpret them and transfer (upload) measurement results to *communication server*. The *communication server* then calls *radio map database* interfaces to append new results to database and gets notified if there were any errors.

Measurements were performed using pocket computer HP iPAQ hw6515d with Microsoft® Windows CE® 4.21 and Microsoft® .NET Compact Framework 2.0 installed. This device is able to measure GPS position as well as to utilize Subscriber Identity Module Toolkit (SIM TK) to measure GSM signal parameters. Radio map database is stored in Oracle®

Express 10.0 database, which is free to use but has limited capabilities. This allows data access using standardized Structured Query Language (SQL).

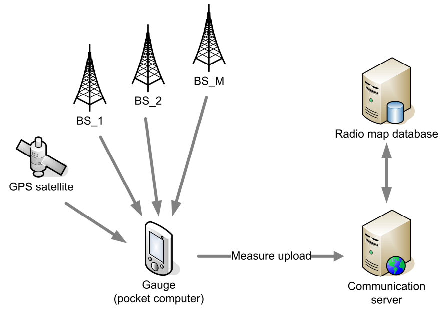


Fig. 2. Gauge subsystem.

2.2. Radio map

Radio map consists of spots. Every spot represents point in real world, where measurement was performed. Tab. 1 shows a snip of the radio map database.

Spots are uniquely identified by **spot_id**, which is its database identifier and is not related to GSM or GPS network. Position of the spots is defined by their position coordinates, which are **latitude** (in degrees), **longitude** (in degrees) and **altitude** (in meters). As shown in Tab. 1, columns with same spot_id have same position coordinates. There can be up to 7 GSM signal measurements performed for every spot depending on a number of available base stations. There are 3 measurements for spot 28914 and 2 measurements for spots 28915 and 28916 shown in Tab. 1. Parameters measured from GSM network are:

- **RxLev** – received signal strength (in dBm) of BS under measure,
- **CI –** cell identity, which is unique identifier of BS in GSM network. It is not always available, therefore BCCH and BSIC are measured as well,
- **BCCH –** Broadcast Control Channel number,
- **BSIC –** Base Station Identification Code. Together with BCCH uniquely identifies BS in GSM network.

SPOT_ID	28914	28914	28914	28915	28915	28915	28915
RxLev	-73	-67	-78	-79	-65	-73	-78
CID	20872	24343	24342	22401	39893	20872	24342
BCCH	3	13	8	10	20	3	8
BSIC	29	27	27	24	26	29	27
Latitude	49,20206	49,20206	49,20206	49,20204	49,20204	49,20204	49,20204
Longitude	18,72063	18,72063	18,72063	18,72065	18,72065	18,72065	18,72065
Altitude	390	390	390	391	391	391	391

Tab. 1. Radio map database example from Žilina-Bánová, Slovakia, GSM operator Orange Slovensko.

3. Future works

Future works can be focused on radio map extrapolation, which would allow estimation of radio signal properties at unknown spots as well as comprehensive visualization options. This would also allow better accuracy in areas of radio map with weaker or missing signal information.

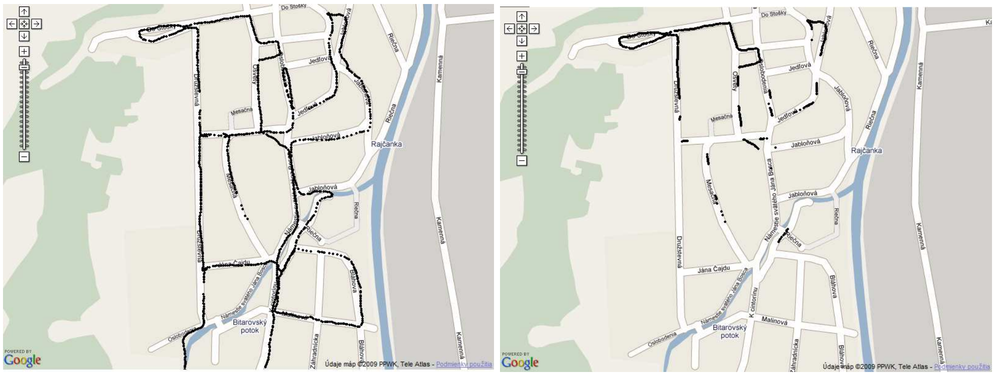
The development of fingerprinting online phase component would enable radio map framework to become usable service-oriented solution for localization in GSM networks.

For positioning in areas like buildings, where GPS signal is not available in general, extension of the framework could be proposed. It would allow measuring signal parameters without need of GPS signal. This would have allowed indoor and outdoor positioning using only one device, if localization service had been implemented.

4. Conclusion

Proposed solution shows combination of existing technologies to create robust framework for further localization development. It provides means to create radio map database of any area covered by GSM signal with no dependency on certain mobile network operator. By use of Oracle® Express database and corresponding SQL support, it gives a hand to advanced data research.

The radio map framework data visualization, which is made by web server and Google Maps™ API, allows displaying of all spots from radio database at once as shown in Fig. 3 a). Spots are shown as black dots. Besides, it allows picking one BS and displaying all spots where signal of this particular BS was available as shown in Fig. 3 b).



a) b)

Fig. 3. a) Radio map database with all spots; b) Radio map database filtered for BS with CID=4F81, BSIC=29 and BCCH=15.

Acknowledgement

The work on this paper was partially supported by the grant VEGA 1/4065/07 of Scientific Grant Agency of the Slovak Republic and by institutional grant 28/103150 of Faculty of Electrical Engineering, University of Žilina.

References

[1] OTSASON, V. *Accurate Indoor Localization Using Wide GSM Fingerprinting*. Master's thesis, Tartu, 2005.

[2] BENIKOVSKÝ, J. *Vytvorenie systému na určovanie polohy v reálnej sieti GSM - Creation of System for Localization in Real GSM Network*. Master's thesis, Žilina, 2008. (in Slovak)

[3] BRIDA, P. *Location technologies for GSM*. Proceedings of 5th European Conference of Young Research and Science Workers in Transport and Telecommunication TRANSCOM, June 2003, Žilina, pages 119-122. ISBN 80-8070-081-8.

[4] *Location Based Services Forecast from ABI Research*, cit. [2008-11-12], available on internet: http://www.twine.com/item/11djpjdj2-172/location-based-services-forecast-from-abi-reserach

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Internet Communication Influence on Insurance Companies

[*]Maciej Czaplewski
[*]University of Szczecin, Faculty of Management and Economics of Services,
ul. Cukrowa 8, 71-004 Szczecin, Poland, maciej.czaplewski@wzieu.pl

Abstract. The economic insurance market seems to have a potential proclivity towards Internet, e-business and e-commerce usage. The use of electronic solutions creates the conditions for winning new customers, increasing the loyalty of current customers and in consequence allows to strengthen the market position of a company. The article describes the main e-business strategies as well as related innovative distribution systems.

Keywords: E-business models, e-commerce, insurance market.

1. Introduction

Without any doubt the insurance sector is an essential component of financial markets and plays a very important role in almost every national economy. Many companies from the insurance sector are functioning mostly in a traditional way, which means that their customers are reached mostly by their own offices and other external structures (e.g. insurance agents, brokers and other dealers). But nowadays – due to the increasing liberalization process - the insurance companies are forced to function on more and more competitive market and subsequently they need to fight even harder for their customers.

Such conditions cause, that insurance companies are trying to benefit from the use of a relative new invention, which is Internet. The Internet has an important part to play within the new concepts of disaggregated value chains of “manufacturers” and “distributors”. The most obvious applications apply to distribution with the Internet acting as both - a distribution channel and as an enabler for other more integrated distribution models. The Internet plays also a significant role in other areas, such as data exchange and customer services. In consequence the new technology can enable companies to enter into the world of e-business. [1]

Electronic Business, commonly referred to as "eBusiness" or "e-Business", may be defined broadly as any business process that relies on an automated information system. Today, this is mostly done with Web-based technologies. [7]

Although e-business is relatively new, it is already having a profound impact on business strategy and operations. It is not limited to messaging but instead it includes ways of helping companies meet their goals, both in the marketplace as well as in the back-office. Moreover e-business can be used as well as in the strategic and in the operational dimension.

From one side companies believe that e-business can help them serve their existing as well as potential customers better, which is nowadays a key strategic goal. It’s even more important due to the fact, that many enterprises today are reorganizing to become centered around customers. The new technology can help to provide services “anywhere and anytime”. [3]

On the other side the technology of e-business itself is enabling companies to accomplish new operational goals. Managing technology and quality of information remain key operational concerns for managers, and they see process streamlining and cost control/efficiency as best ways to gain competitive advantage. [3]

2. Main e-commerce strategies

In order to benefit from the use of Internet and e-business, the companies have to adapt also an appropriate e-commerce strategy. Observation of e-commerce utilization on the insurance market shows that their business activities could be measured by:

- the way of e-commerce engagement. According to this factor, one can divide insurance companies:
    - selling services only through electronic channels;
    - acting in traditional way and setting up separate organizational units that are providing services with electronic channels only;
    - doing business in both, traditional and electronic way, within one organization, with different grade of internal coordination of these activities;
- the method of implementation of e-commerce value-added chain. This allows to divide companies to those which are:
    - concentrating only in selected segments of the value-added chain;
    - implementing e-commerce in the whole added-value chain.

Juxtaposition of the two, above mentioned, ways of measuring business activities of companies implementing e-commerce, allows to point out four basic e-commerce strategies. The figure number 1 shows the four basic e-commerce strategies pursued by the companies on the insurance market.

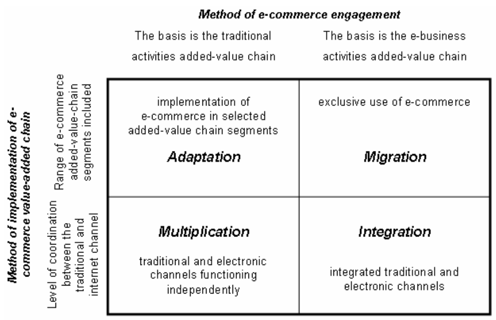


Fig. 1. Basic strategies of e-commerce implementation. (Source: B. Birkhofer, E-Commerce als innovativer Absatzkanal. Ein entscheidungsorientiertes Modell, Rosch-Buch, Schesslitz 2001, p. 118.)

3. Basic e-business models in the insurance sector

As it was stated earlier in the article, the effects of e-business and e-commerce are the subject of intense debate in the insurance industry, although actual translation into solutions is still in its infancy. Various e-business models are emerging and compete with traditional insurers. Newly established internet insurers are in the process of implementing the new possibilities provided by technology and testing innovative business models. [5]

Basically it looks that nowadays for companies that want to use e-business technology and operate on the insurance market there is scope for three different business models: [1]

- Independent eIntermediary,
- Insurers’ own websites,
- eInfomediary.

The independent eIntermediary business model takes the traditional agent or broker role from the office and phone-based distribution and applies it to the Internet. Under this model, the user provides his details and requirements and the online broker provides a number of quotes and suitable products for the consumer. Subsequently the customer can choose from the options proposed by the broker. There are already many companies in the Internet that are using this model. One of the best examples can be found in the USA – InsWeb, but there are also numerous European sites coming online and using this model, such as Screentrade or Ironsure.

The second business model provides transactional capabilities on the insurer's own website. Many established insurers noticing the advantages of e-business technologies and the advancements made by the insurance market "newcomers" using Internet have also begun to restructure their business systems and set up additional online sales channels.

Currently all (or almost all) companies acting on the insurance market have created their own website. One can find very simple as well as very complex websites, but irrespective of that, the basic application is enabling the customers access to information about company and its products as well as cheap and ubiquitous communication.

The more sophisticated websites allow customers also: [4]

- to perform a simple needs-analysis,
- get quotes for insurances,
- buy products that may interest them,

and after acquiring a product:

- to get pending information about their insurance policy,
- manage their accounts values,
- report an coverage accident and allow claim management.

Moreover the use of Internet-based information systems within an enterprise creates the opportunity to maximize efficiency and effectiveness of its internal processes. Issues that can be improved in his way are for example:

- more systematic use of the available data at the right time of business process,
- more systematic information storage, which makes them available at another step in the process, at another time or another place for different employees,
- better connections of various information systems,
- storage of the data just in one place and subsequently a cheaper and easier maintenance and keeping them up to date. [6]

However, re-engineering traditional business processes is expensive and often meets with considerable opposition from within the company itself. Creation of an internal information system is relatively easier, because the agents and other company employees can feel the benefits of such a system quite quick. But the main challenge for insurers is to adapt the existing agent-based model of distribution to the new Internet reality and convince the agents that there is no threat for them from the side of Internet-based distribution.

Nowadays the companies try to encourage more of the commodity business and commodity transaction to go directly through the Internet, reducing cost and commissions for the agents. At the same time, the agents have to focus on more added-value and complex products like variable universal life insurances, which will help to compensate for he loss of earnings from simple products like motor insurances.

The third business model combines the traditional channels with the Internet. The model works around the Internet being used simply as an information source, which then refers the consumer to an existing adviser in order to provide more tailored advices and carry out the transaction. Examples of this type of model are websites that provide insurance product information from experts as well as from independent consumers who can write their mind about an insurance company or its products.

The Internet represents an ideal medium for companies to enhance customer servicing and give individual information at any time, but one can foresee that the eIntermediary model will develop into the main distribution network for commodity-type insurance on the Internet. The independent eIntermediary can offer more benefits to the consumer than a website of one of the insurances companies because it compares more products at one time and can tailor better services for the customers.

But nowadays it is hard to compare different products offered by different companies. That is why most customers who wants to buy their insurances online are using insurance companies own websites which have more and more facilities.

4. Conclusion

It is obvious that the Internet has a great influence on the insurance industry and many companies are adapting it to their business processes. But many of them see the Internet just as a channel of distribution. The Internet can be seen as both less and more than a channel of distribution. [4]

On the one hand, the Internet is less than a channel because it often needs to hand off to other channels to close the deal. Customers may find the company via Internet, but the vast majority will want to talk to an advisor by phone or direct in person before decide to buy. That's why Internet can be perceived in many cases more of a lead-generation tool than a distribution channel.

On the other hand, the Internet is more than a channel in that the Internet is helpful tool for all channels of distribution – for getting quotes, doing illustrations, getting the right forms, viewing pending policy status downloading data, etc.

Moreover the internet enables new entrants to the market to avoid the expensive and lengthy process of setting up traditional sales networks. Established insurers are thus facing growing competitive pressure. [5] In order to stay on the market and to be more competitive in the future they are adapting the Internet and e-business, which enables them to link their internal and external data processing systems more efficiently and flexibly, to work more closely with suppliers and partners as well as to better satisfy the needs and expectations of their customers.

Nowadays it seems that insurance companies (no matter if new entrants or established ones) should choose a market strategy, which from one side bases on both ways of functioning - the traditional one as well as on the one that uses the new technologies.

References

[1] BEERLI A. J., Facilitating conditions for effective knowledge transfer In New Business models, Difo-Druck GmbH, Bamberg 2001.

[2] BIRKHOFER B., E-Commerce als innovativer Absatzkanal. Ein entscheidungsorientiertes Modell, Rosch-Buch, Schesslitz 2001.

[3] BROWN S.A., Customer Relationship Management. A strategic imperative in the world of e-business, Willey & sons, Toronto 2000.

[4] HOLLEN K. E., Re-engineering the insurance business system using e-business in financial services. In: The insurance revolution: new frontiers in the convergence of financial services, seminar proceedings vol. 2001, Vienna 2001.

[5] Swiss Re, sigma 5/2000.

[6] VERBOOM M.,VAN IWAARDEN J., VAN DER WIELE T., A transparent role of information systems within business processes: A case study. Erim Report Series, September 2004.

[7] <http://en.wikipedia.org/wiki/E-business>

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Sentence Melody Analysis for Speech Production in the TTS System

*Mária Čerňanská, *Ondrej Škvarek
*University of Zilina, Faculty of Informatics and Management Science, Department of InfoCom Networks, Univerzitna 2, 01026 Žilina, Slovakia, {maria.cernanska, ondrej.skvarek}@fri.uniza.sk

Abstract. This paper deals with melody analysis of Slovak language sentences. It presents method of sentence melody contour extraction from voice recordings. Extracted contours will be used at the concatenative text-to-speech system for speech prosody modification. Method based on normalized autocorrelation function (program Praat) is used to determine overall pitch contour – aggregation of sentence melody and shorter segments melody (words, syllables). To diminish the number of incorrectly detected values, software parameters are set according to speaker's voice characteristics. Overall pitch contour is further analyzed; values exhibiting extraordinary jumps are removed. Finally, smoothing method is used to remove short-term melody variations and extract melody contour of the sentence. Smoothing method computes values in equidistant time events, which makes cluster analysis (finding types of similar melody sentences) possible.

Keywords: text-to-speech, melody contour, autocorrelation, Praat, smoothing, weighted moving average

1. Introduction

Digitalization has influence to many areas of speech processing, like VoIP [10, 11], new sampling methods [12] etc. In this paper we will focus on text-to-speech synthesis.

Text-to-speech (TTS) systems based on concatenative synthesis method achieve good synthesized speech quality. Concatenative synthesis method is based on selection of voice units from the TTS database. Units are concatenated into one waveform, and then prosody properties (melody, intensity) are applied. This paper describes method of sentence melody contour extraction in Slovak language. Overall melody contours are determined by the method based on normalized autocorrelation function [7], and then sentence contours (suitable for prosody generation and further analysis) using proposed smoothing methods are obtained. To analyze large number of voice recordings, Praat scripts were programmed.

2. Applying Prosody at the Speech Synthesis Produced by TTS-KIS

System TTS-KIS [1, 2, 3, 4, and 5] developed at the Department of InfoCom Networks uses concatenative synthesis (Fig. 1) to produce speech. The basic speech units (segments), stored in the TTS-KIS database are "diphones" (diphone – the signal from the mid point of a phone to the mid point in the next phone). Incoming text is converted into phonetic transcription, and then couples of phonemes are formed. Corresponding diphones are chosen from the system database and they are concatenated into one waveform. Finally, prosody properties (melody and intensity contours for words and for the whole sentence) are applied. In the paper we focus on the sentence melody contour extraction.

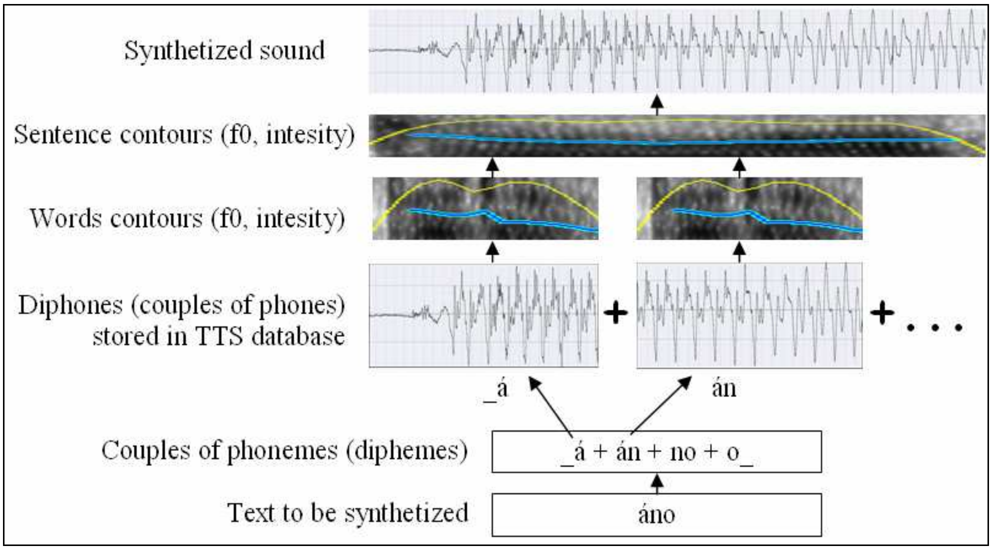


Fig. 1. Concatenative method of speech synthesis in the TTS-KIS system, based on speech synthesis unit “diphone”.

3. Obtaining of Overall Melody Contour Using Program Praat

We use program Praat [8] to determine overall melody contour for each one of recorded sentences. Program detects voiced and unvoiced parts of the speech and determines melody contour (pitch contour, glottal frequency, f0) during voiced intervals (of Fig. 2).

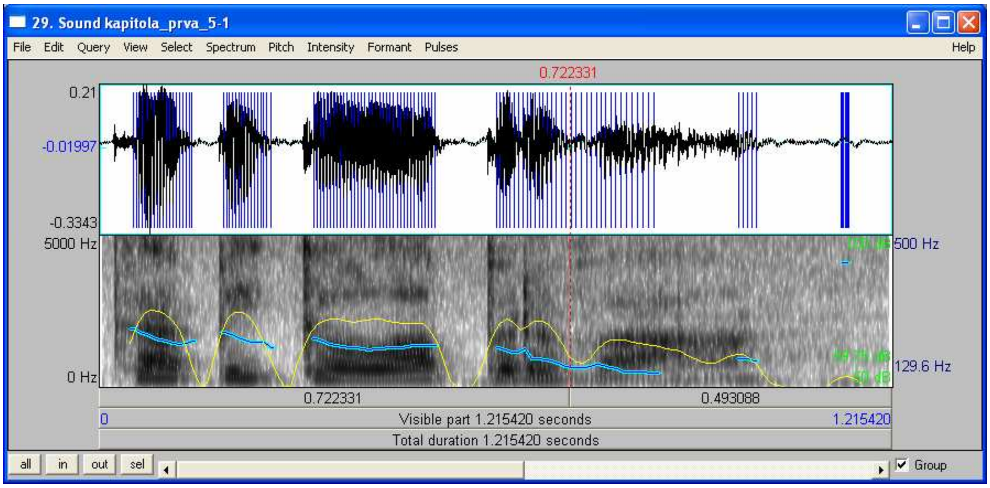


Fig. 2. Melody and intensity contour of sentence ”Kapitola prvá” (upper part of the figure – waveform with glottal pulses, lower part – blue dicontinuous curve – melody in Hz, yellow curve – intensity in dB).

Pitch contours at Fig. 2 were computed by “To Pitch (ac)” method [9]. This method implements algorithm [7] to detect periodicity of speech signal. Speech signal is multiplied by window function, and then autocorrelation function is computed. Resulting values are divided by autocorrelation function of the window.

At each time event more f0 candidates are determined. For each candidate preference is computed (Fig. 3). Finally, algorithm chooses candidates forming optimal path along the sentence (preferences and parameters: Octave Cost etc. [7, 9] are taken into account).

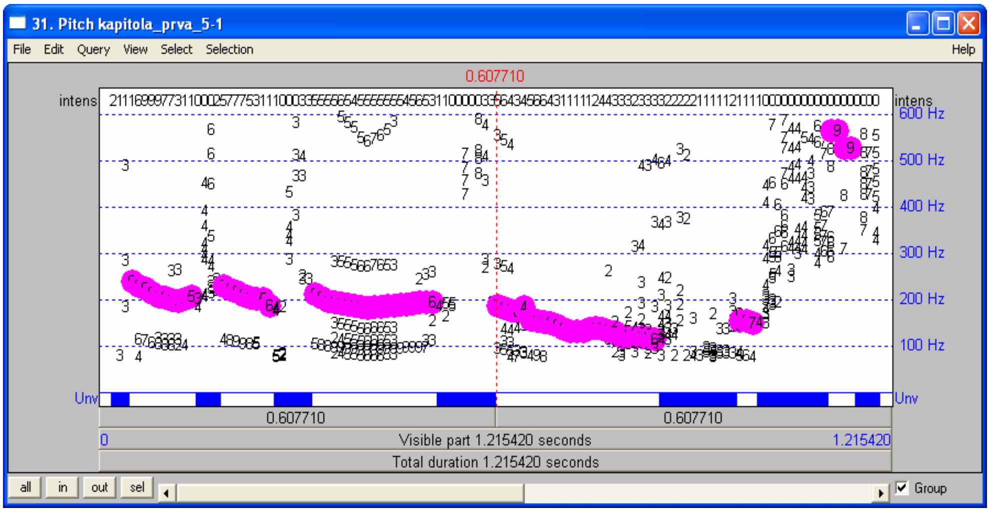


Fig. 3. Melody contour computed by “To Pitch (ac) “method of program Praat (default settings: analyzed frequency range 75 Hz – 600 Hz).

To eliminate incorrectly determined high frequencies (Fig 3, upper right cornet) and low frequencies (f0/2), analyzed frequencies range was changed according to the speaker characteristics to 100 Hz – 350 Hz (Fig. 4).

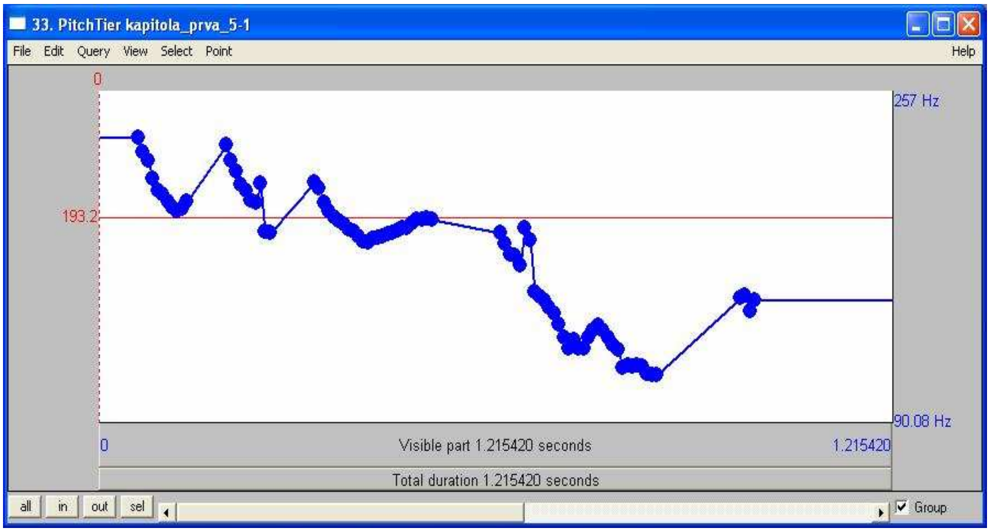


Fig. 4. Melody contour obtained after restriction of analyzed frequencies to range 100 Hz - 350 Hz (Praat: we extracted the PitchTier object from sound Manipulation object).

Melody contour was further analyzed; values exhibiting extraordinary jumps (80 Hz and more) were removed.

4. Melody Contour Smoothing

Overall melody contour obtained by program Praat is composed of the sentence melody and of shorter segments melody contours (words, syllables). Smoothing method is used, to remove short-term melody variations and to extract sentence melody contour. We designed smoothing methods computing pitch values in equidistant time events, which make cluster analysis (finding types of similar melody sentences) possible:

1. **Moving average (given interval length).** We divide duration of the sound (sentence sound) into "m" equal length intervals (m – denotes required number of smoothed values), then we compute smoothed value in the center "t" of each time interval as follows. Values of overall melody contour belonging to interval (t – d, t + d) are taken and arithmetic mean is computed (d – denotes the window size in seconds). See Fig. 6 and 7, contour 1.
2. **Moving average (given number of points).** "N" values of original melody contour are taken (n/2 values to the left and n/2 values to the right from the time event "t"), and then arithmetic mean is computed. See Fig. 6 and 7, contour 2.
3. **Weighted moving average (given number of points).** This method is similar to the second method. Values of overall melody contour are weighted, then weighted average is computed: $(y1{*}w_1 + y_2{*}w_2 + ... + y_n{*}w_n) / (w1 + w_2 + ... + w_n)$, where i = 1, 2 ... n, y_i – overall melody contour values, w_i – weights, $w_i = a * | t - t_i | + b$, $a, b \in R$. (See Fig. 5; Fig. 6 and 7, contour 3).
4. **Exponential smoothing**. Overall melody values "f_i" (i = 1, 2 … n) are taken, smoothed values $O_{i+1} = \alpha * f_{i+1} + (1 - \alpha) * O_i$ are computed, where parameter $\alpha \in \langle 0, 1 \rangle$. Final values in equidistant time events we obtain by linear interpolation. See Fig. 6 and 7, contour 4.

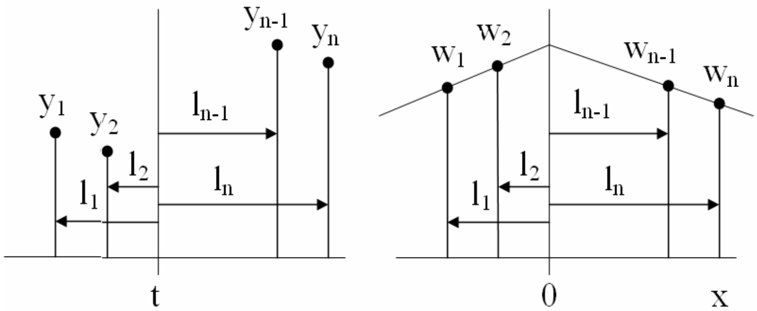


Fig. 5. Computing weights in Method 3: Weighted moving average (given number of points).

To exclude short-term melody variation, we set the length of interval d = 0.3 sec (method 1), number of points n = 15 (method 2 and 3), parameter α = 0.4 (method 4). Required number of contour points "m" was chosen proportional to the number of words in the sentence.

On unvoiced interval, method 1 exhibits high jump in frequency (Fig. 7, contour 1), also method 2 does not grow continuously (constant trend, Fig. 7 contour 2).

Method 3 was designed to improve the second one method. This method achieved continuous growth.

Given low value α, method 4 underestimate long falling melody trend. If larger values α are used, this method does not remove short-term melody variations.

We chose method no. 3 to be the best one and we compute melody contours of recorded sentences.

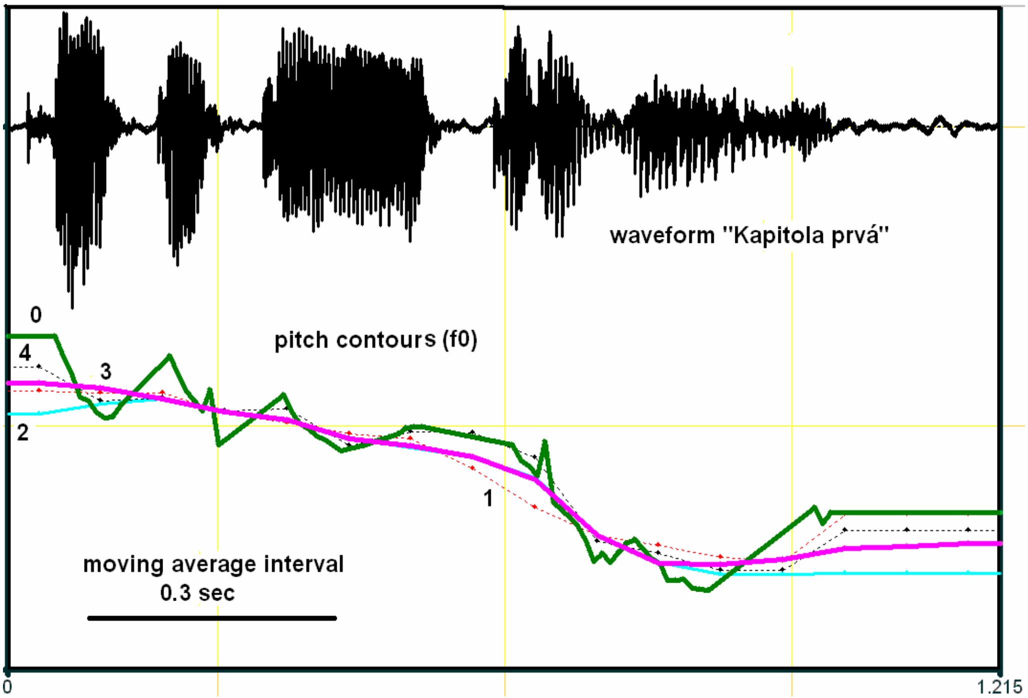


Fig. 6. Choosing the best method of smoothing. Contours of the sentence “Kapitola prvá”. (0 – contour from Praat, 1 – Moving average (given interval length), 2 – Moving average (given number of points), 3 – Weighted moving average (given number of points), 4 – Exponential smoothing.)

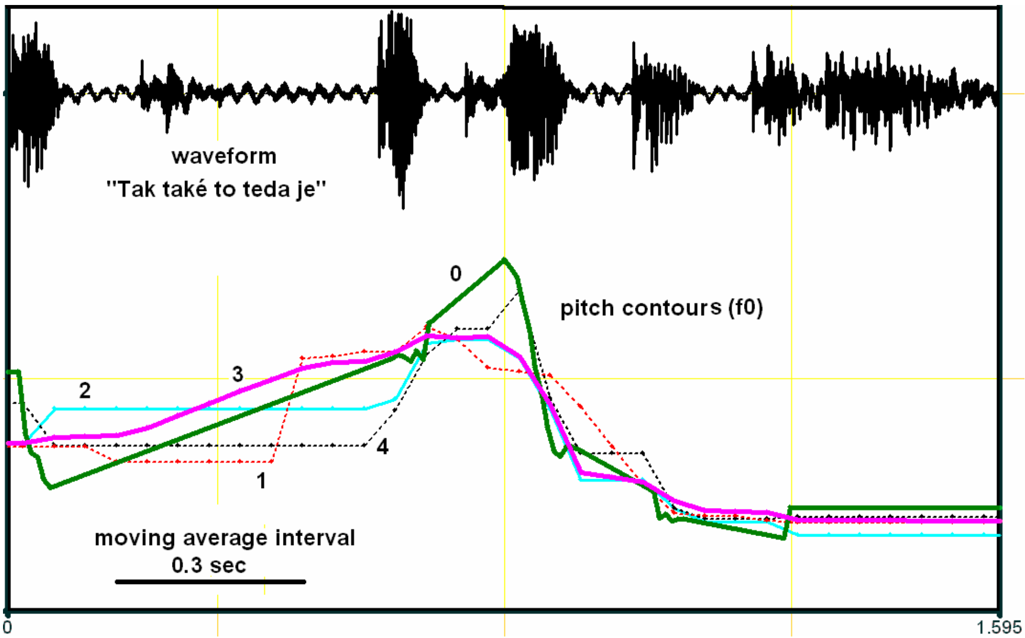


Fig. 7. Choosing the best method of smoothing. Contours of the sentence “Tak, také to teda je”. Stress is on the word “také” (0 – contour from Praat, 1 – Moving average (given interval length), 2 – Moving average (given number of points), 3 – Weighted moving average (given number of points), 4 – Exponential smoothing).

5. Conclusion

We presented method for sentence melody extraction from voice recordings in Slovak language. Method “To Pitch Tier” (program Praat) is used to obtain overall melody contour. Proper parameter values were set, to eliminate incorrect detection of f0. To extract sentence melody from overall melody contour, different smoothing methods are examined. The “Weighted moving average (given number of points)” method is chosen as the best one and it is used to compute melody contours for recorded sentences. To deal with large number of sound files, scripts in the Praat program were implemented. Obtained contours can be used at the TTS-KIS system to add prosody to the generated speech.

Contour values are determined at given number of equidistant time intervals, which makes cluster analysis possible. Our future work will focus on topics: melody comparison of different length sentences; finding of typical contours for different sentence types [6] in Slovak language; melody analysis of shorter speech segments (words, syllables); intensity contour analysis.

References

[1] ČAKY, P., KLIMO, M., MIHÁLIK, I., MLADŠÍK, R. Text-to-speech for Slovak language. *Text, Speech and Dialogue: 7th International Conference, TSD 2004, Brno, Czech Republic, September 8-11, 2004: proceedings.* Berlin: Springer, 2004. ISBN 3-540-23049-1. P. 291-298.

[2] TTS system version 1 demo (developed at the Department of InfoCom Networks, Faculty of Informatics and Management Science, University of Zilina, Slovakia). http://tts.kis.fri.utc.sk/. (Taken: March 18 2009.)

[3] MIHÁLIK, I. *Kompresia reči v TTS* [Ph.D. thesis]. Žilinská univerzita v Žiline, Fakulta riadenia a informatiky, Katedra informačných sietí. Žilina, 2008. P. 108. (In Slovak)

[4] KLIMO, M., MIHÁLIK, I. Extending annotational SSML into structural content for speech synthesizer. *Scientific Papers of the University of Pardubice: 12 (2006): Series B* - The Jan Perner Transport Faculty. - Pardubice: Univerzita Pardubice, 2007. ISBN 978-80-7194-985-5. P. 193-199.

[5] SELEP, P. Diphone concatenation method for speech synthesis. *TRANSCOM 2007: 7-th European conference of young research and science workers: Section 3: Information and communication technologies.* Žilina: University of Žilina, 2007, pp. 221-224. ISBN 978-80-8070-693-7

[6] KRÁĽ, A. *Pravidlá slovenskej výslovnosti.* Slovenské pedagogické nakladateľstvo Bratislava, 1996. (In Slovak)

[7] BOERSMA, P. *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound.* Institute of Phonetic Sciences 17: 97–110. University of Amsterdam, 1993.

[8] Program Praat main web page. http://www.fon.hum.uva.nl/praat/. (Taken: March-18 2009.)

[9] Program Praat manual. Methods “Sound: To Pitch...”, “Sound: To Pitch (ac)...” – purpose, arguments, algorithm. http://www.fon.hum.uva.nl/praat/manual. (Taken: March-18 2009.)

[10] BACHRATÁ, K., BOROŇ, J., ROSIVAL, B. New methods proposal for VoIP. *Scientific Papers of the University of Pardubice : 12 (2006): Series B* - The Jan Perner Transport Faculty. - Pardubice: Univerzita Pardubice, 2007. ISBN 978-80-7194-985-5. - P. 201-210.

[11] KLIMO, M., BACHRATÁ, K., SMIEŠKO, J., URAMOVÁ, J. *Teória IP telefónie.* Žilinská univerzita, 2009. ISBN 978-80-8070-915-0. P. 379. (In Slovak)

[12] KLIMO, M., BACHRATÁ, K. Sampling of function generating shift invariant spaces. *Scientific papers of the University of Pardubice : The Jan Perner Transport Faculty. 10 (2004): Series B.* - Pardubice: Univerzita Pardubice, 2005. ISBN 80-7194-779-2. P. 165-175.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Effective Bandwidth at Work Conserving Link

*Jakub Daubner
*University of Žilina, Faculty of Management Science and Informatics, Department of Info Communication Networks, Univerzitná 8215/1, 01026 Žilina, Slovakia, {jakub.daubner}@kis.fri.uniza.sk

Abstract. In this paper, we describe how can be the theory of effective bandwidth applied at dimensioning the capacity of work conserving link to guarantee the maximal packet loss probability for given stochastic input process. We suppose that stochastic input process is independent and identically distributed.

Keywords: effective bandwidth, work conserving link, independent and identically distributed process

1. Introduction

Dimensioning different network elements to guarantee different *quality of service* (*QoS*) parameters of communication networks is very actual problem. This problem cannot be solved easily, because there are many different network elements and many different input processes for which we can define many QoS parameters. So it is hard (or impossible) to solve such problem in general.

Theory of effective bandwidth can solve some parts of this problem. It focuses on stochastic input processes and on few types of network elements like work conserving link, multiplexer and router. The theory also allows us to connect these basic network elements to arbitrary acyclic networks.

In the present paper we will focus at problem how to guarantee packet loss probability for work conserving link with constant capacity. To avoid any misunderstanding, first, we define all used notations.

2. Preliminaries and Notation

We consider a discrete-time system with time indexed by $t = 0,1,2,\dots$ We describe a discrete-time arrival process of a traffic source by a sequence of random variables $A \equiv \{A(t)|t = 0,1,2,\dots\}$, where $A(t)$ is the cumulative number of arrivals by time t. We assume that there are no arrivals at time 0, i.e. $A(0) = 0$, and that $A(t)$ is increasing. Let $a(t) = A(t) - A(t-1)$ be the number of arrivals at time $t = 1,2,\dots$ Following the terminology in communication networks, we call an arrival a packet. Packets are assumed to be of the same size.

We say that sequence of random variables $\{a(t)|t = 1,2,\dots\}$ is independent and identically distributed if for any times t,s and any values x,y it holds that $\Pr(a(t) > x|a(s) > y)) = \Pr(a(t) > x)$ and $\Pr(a(t) > x) = \Pr(a(1) > x)$, respectively. Shortly, we say that such sequence (or process) is i.i.d.

Work conserving link with constant capacity c (Figure 1) is network element, which has infinite buffer for arrival packets and which can serve maximum c packets per unit of time (per time slot). We denote number of packets in buffer (queue length) at time t by $q(t)$. We denote the output from a work conserving link by $B \equiv \{B(t)|t = 0, 1, 2, \ldots\}$.

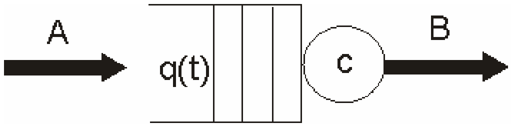


Figure 1. Work conserving link with constant capacity c

The queue length satisfies Lindley equation

$$q(t+1) = (q(t) + a(t+1) - c)^+, \tag{1}$$

where $(x)^+ = \max\{0, x\}$. In the case that $q(t) + a(t+1) \geq c$, the number of packets in the buffer at time $t+1$ is simply the sum of the number of packets at time t and the number of arrivals at time $t+1$, subtracting those c packets that depart at time $t+1$. In the case that $q(t) + a(t+1) < c$, the becomes empty at $t+1$.

For i.i.d. process $\{a(t)|t = 1, 2, \ldots\}$ we can define effective bandwidth function $\alpha(\theta)$ following

$$\alpha(\theta) = \frac{1}{\theta} \ln E[e^{\theta a(1)}]. \tag{2}$$

In [2] is proved that $\alpha(\theta)$ is increasing for $\theta > 0$ and is bounded between average rate and peak rate of process $a(1)$. Moreover, in [2] is proved following upper boundary for queue length of work conserving link. If $c > \alpha(\theta)$, then

$$\Pr(q(t) > x) \leq \beta(\theta) \cdot e^{-\theta x}, \tag{3}$$

where $\beta(\theta) = (1 - e^{\theta(\alpha(\theta) - c)})^{-1}$ (note that $\beta(\theta)$ does not depend on x). It means, that probability, that queue length is more than x, decay at least exponentially fast with exponent $-\theta x$.

In [2] is also proved following limit behavior of $\Pr(q(t) > x)$. If $c = \alpha(\theta)$, then

$$\lim_{x \to \infty} \frac{1}{x} \ln \Pr(q(t) > x) = -\theta. \tag{4}$$

This means, that upper boundary from (3) is tending to real value of $\Pr(q(t) > x)$ as $x \to \infty$.

3. Application of the theory of effective bandwidth

3.1. Traffic source with binomial distribution

Suppose we have a traffic source which generates one packet per time slot with probability p and zero packets per time slot with probability $1 - p$. The effective bandwidth function for such process is

$$\begin{aligned} \alpha(\theta) &= \frac{1}{\theta} \cdot \ln(p \cdot e^{1\cdot\theta} + (1-p) \cdot e^{0\cdot\theta}) \\ \alpha(\theta) &= \frac{1}{\theta} \ln(pe^{\theta} + 1 - p). \end{aligned}$$

Let $p = 1/2$, then we have

$$\alpha(\theta) = \frac{1}{\theta} \ln(\frac{1}{2} e^{\theta} + \frac{1}{2}), \tag{5}$$

which is plotted on Figure 2. Note that $\alpha(\theta)$ is tending to average rate $1/2$ as $\theta \to 0$, and it is tending to peak rate 1 as $\theta \to \infty$.

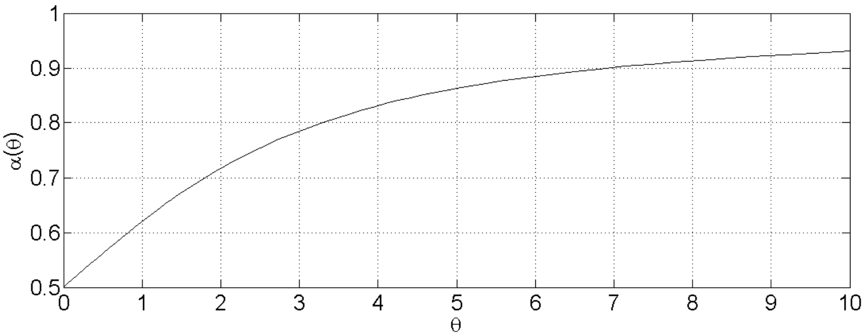


Figure 2. Plot of effective bandwidth function for concrete binomial distribution

Let $\theta = 3$, then we have $\alpha(\theta) = 0.7851$. Let $c = 0.8$, then from (3) we have

$$\Pr(q(t) > x) \leq 22.9455 \cdot e^{-3x}. \tag{6}$$

For example for $x = 10$ we have that $\Pr(q(t) > 10) \leq 2.1472 \cdot 10^{-12}$. This means, that if we set the capacity of work conserving link to 0.8, than the probability that queue length will be more than 10 is less than $2.1472 \cdot 10^{-12}$.

3.2. Traffic source with Poisson distribution

Suppose we have a traffic source where $a(1)$ has Poisson distribution with parameter $\lambda = E[a(1)]$. From Possion distribution we know, that for all $k = 0, 1, 2, \dots$ it holds

$$\Pr(a(1) = k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!}. \tag{7}$$

The effective bandwidth function for such process is

$$\begin{aligned} \alpha(\theta) &= \frac{1}{\theta} \cdot \ln\left(\sum_{k=0}^{\infty} e^{\theta k} \cdot \frac{\lambda^k \cdot e^{-\lambda}}{k!}\right) \\ \alpha(\theta) &= \lambda \frac{e^{\theta} - 1}{\theta}. \end{aligned}$$

Let $\lambda = 1/2$, then we have

$$\alpha(\theta) = \frac{e^{\theta} - 1}{2\theta}, \tag{8}$$

which is plotted on Figure 3. Note that $\alpha(\theta)$ is tending to average rate $1/2$ as $\theta \to 0$, and it is tending to ∞, because Possion process has no peak rate.

Let $\theta = 0.85$, then we have $\alpha(\theta) = 0.7880$. Let $c = 0.8$, then from (3) we have

$$\Pr(q(t) > x) \leq 98.7657 \cdot e^{-0.85x}. \tag{9}$$

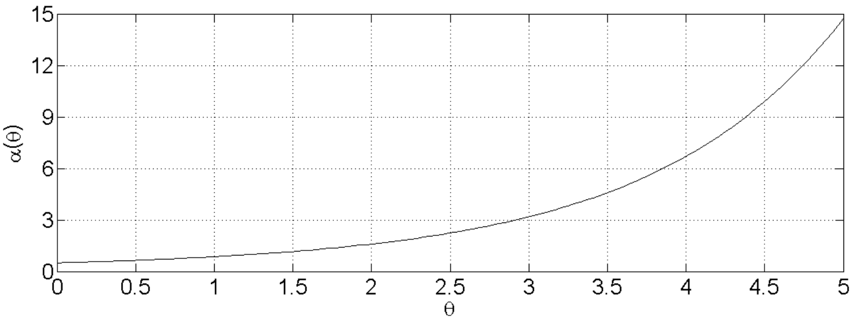


Figure 3. Plot of effective bandwidth function for concrete Poisson distribution

For example for $x = 10$ we have that $\Pr(q(t) > 10) \leq 0.0201$. This means, that if we set the capacity of work conserving link to 0.8, than the probability that queue length will be more than 10 is less than 0.0201.

Notice that the upper boundary from (6) is better than upper boundary from (9) (for the same capacity $c = 0.8$). This means that we can guarantee packet loss probability better for binomial process with mean value $1/2$ from first example, than for Poisson process with same mean value from second example. It is true, because variance of first process is $1/4$ and variance of second process is $1/2$. So the second process has bigger standard deviation than first process. Notice, that standard deviation is measure of how is random variable spread around its mean value.

4. Conclusion

In this paper has been demonstrated how can be the theory of effective bandwidth used for dimensioning the capacity of work conserving link to guarantee packet loss probability. Goal of future work will be to extend this theory to other network elements, especially on a token bucket.

References

[1] BACHRATÁ, K.: *Effective Bandwidth for Deterministic Networks*, Communications, Vol. 4, pp. 78-82, 2003

[2] CHANG, CH. S.: *Performance guarantees in communication networks*, Springer-Verlag, 2000.

[3] KLIMO, M., BACHRATÁ, K., SMIEŠKO, J., URAMOVÁ, J.: *Teória IP telefónie*, EDIS-vydavateľstvo ŽU, 2009

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Accurate Shot Boundary Detection for Slovak TV News Cutting

[*]Zuzana Dubcová, [*]Róbert Hudec, [*]Miroslav Benčo, [*]Peter Lukáč
[*]University of Zilina, Faculty of Elektrotechnical Engineering, Department of Telecomunication and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia, {zuzana.dubcova, robert.hudec, miroslav.benco, peter.lukac}@fel.uniza.sk

Abstract. In this paper, an overview of Shot Boundary Detection (SBD) principles as an essential element for various video-processing technologies with specially focus oriented on the cut detection is presented. A lot of scientists have been concentrated on creating accurate shot boundary detection algorithms in recent years. However a truly accurate method of cut detection still eludes researchers in general. On the other side, the basic type of abrupt transition called "cut or hard cut" is better detected than types of gradual transitions like dissolve, fade out/fade in or wipe. Moreover, in Slovak TV news streams, generally, the hard cuts are widely used.

Keywords: group of picture (GOP), shot boundary detection (SBD), cut, dissolve, fade in, fade out, wipe

1. Introduction

In this paper, the algorithm of video cut detection is introduced. Our study on semantic analysis of digital media falls under the domain of Information and Communication Technologies. Digital video collections are growing rapidly in both the professional and consumer environment, and are characterized by a steadily increasing capacity and content variety. Since searching manually through these collections is tedious and time-consuming. The development of video archiving and retrieval systems are based on the algorithms for video content analysis. These algorithms are built around the models bridging the gap between the syntax of the digital video data stream and the semantic meaning of that stream. The task is to find each shot boundary in the test collection and identify it as an abrupt or gradual transition where any transition which is not abrupt, is considered gradual.

The paper outline is as follows. In the next sections, an overview of the video structure and the video transition types are introduced. The algorithm theories of the SBD (Shot Boundary Detection) are presented in section no. 3 and the feature extraction for accurate cut detection in section no. 4, respectively. Finally, the brief summary of achieved results and future tasks are discussed in conclusion.

2. Problem background

In content-based video analysis, video can be divided to the six structural levels, namely, key frames, frames, shot, scene/view, story/event and video sequence, as shown in Figure 1.

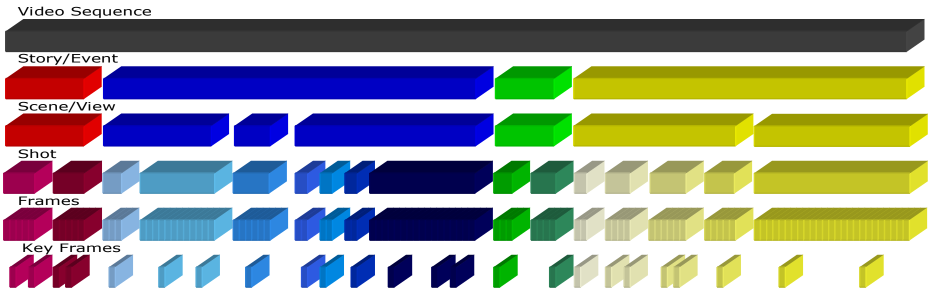


Fig. 1. Six structured levels for content based video analysis.

Digital video is also organized into frames from several frames up to 60 per second. Above the frame, the next largest unit of video both syntactically and semantically is called the shot, [3] . There are various types of the shot transitions, namely:

- abrupt transition of shots (cuts), Figure 2a,
- gradual transition of shots:
    - dissolves and fades in/out (two scene crossing), Figures 2b-c,
    - wipes (special graphical transitions), Figure 2d.

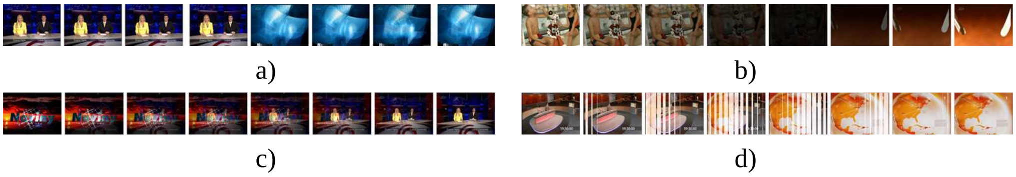


Fig. 2. The shot transitions, a) cut, b) fade in/out, c) dissolve, d) vertical wipe.

3. Algorithms of shot detectors

Each method for cut detection works on a two-phase principle:

- Scoring - Each pair of consecutive frames of a digital video is given a certain score that represents the similarity/dissimilarity between these two frames.
- Decision - All scores calculated previously are evaluated and a cut is detected if the score is considered high.

The major techniques used for SBD are differences in luminance or color values of pixels, statistical measures, histograms, edge classifications, level of compressions, motion vectors, textures, level of segmentations, low to high ratio of DCT coefficients DCT (Discrete Cosine Transform), and a many more. Nowadays, there are complex algorithms generally using the bottom principles.

3.1. Pixel differences

Equations in the manuscript are of the style “equation”. Equations are numbered as shown in the following example: The easiest way to detect if two frames are significantly different is to count the number of pixels that change in value more than some threshold. This total is compared against a second threshold to determine if a shot boundary has been found [1]. This method is sensitive to camera motion. Very important is the color primaries composition which defines the color space. The human eye is more sensitive to gray level intensities than to color. A lot of color spaces are not based on the linear tristimulus principle but on the luminance and chrominance components, e.g. YUV, YCbCr, YPbPr, YCC, HSV, HMMD, etc. [2]. In these cases, the differences in luminance components can by measured only.

3.2. Statistical differences

Statistical methods expand on the idea of pixel differences by breaking the images into regions and comparing statistical measures of the pixels in those regions. Kasturi and Jain [1] compute a measure based on the mean and standard deviation of the gray levels in regions. This method is reasonably tolerant of noise, but is slow due the complexity of the statistical formulas. It also generates many false positives (i.e., changes not caused by a shot boundary).

3.3. Histograms

Histograms are the most common method used to detect shot boundaries. The simplest histogram method computes gray level or color histograms of the two images. If the bin-wise difference between the two histograms is above a threshold, a shot boundary is assumed. In [5], used the color histogram change rate to find shot boundaries and gray level histogram differences in regions, weighted by how likely the region was to change in the video sequence.

3.4. Compression differences

The differences are in the size of JPEG compressed frames to detect shot boundaries as a supplement to a manual indexing system, or shot boundaries by comparing a small number of connected regions - uses differences in the DCT coefficients of JPEG compressed frames as their measure of frame similarity, thus avoiding the need to decompress the frames [1]. It was obtained by sampling the frames temporally and using a form of binary search to find the actual boundary.

3.5. Edge tracking, edge change ratio (ECR)

ECR attempts to compare the actual content of two frames [1]. It transforms both frames to edge pictures, i.e. it extracts the probable outlines of objects within the pictures. Afterwards, it compares these edge pictures using dilatation to compute a probability that the second frame contains the same objects as the first frame. ECR is one of the best performing algorithms for scoring. It reacts very sensitively to hard cuts and can detect many soft cuts by nature. In its basic form even ECR cannot detect soft cuts such as wipes as it considers the fading-in objects as regular objects moving through the scene. Yet, ECR can be extended manually to recognize special forms of soft cuts.

3.6. Motion vectors

Motion vectors are determined from block matching to detect whether or not a shot was a zoom or a pan or they are extracted as part of the region-based pixel difference computation described above to decide if there is a large amount of camera or object motion in a shot [5]. Because shots with camera motion can be incorrectly classified as gradual transitions, detecting zooms and pans increases the accuracy of a SBD algorithm. Motion vector information can also be obtained from MPEG compressed video sequences. However, the block matching performed as part of MPEG encoding selects vectors based on compression efficiency and thus often selects inappropriate vectors for image processing purposes.

4. Feature extraction

4.1. Evaluation criteria

The accurate boundary detections are evaluated with the ground truth in terms of precision and recall. Recall is defined as the amount of desired items that are retrieved. Precision is defined as the amount of retrieved items that are desired items. Recall measures the ability to present all relevant items, while precision measures the ability to present only relevant items. Recall and precision are in the interval of [0,1]. The recall and precision measures are defined as follows:

$$recall = R = C/(C+M); \qquad precission = P = C/(C+F). \tag{1}$$

F1 is a combined measure that results in high value if, and only if, both precision and recall result in high values. The formula for the calculation of the F1 may be written as follows:

$$F_1 = 2PR/(P+R). \tag{2}$$

The symbols stand for: C - the number of correctly detected cuts, M - the number of not detected cuts and F - the number of falsely detected cuts.

For the SBD algorithms testing, the annotated test video contained a total of 277 frames and 23 shot transitions (14-abrupt and 9-gradual transitions) was created. In experiments, are 3 hard cuts (yellow blocks), 3 dissolves (green blocks), 3 fades out/in (blue blocks) and 3 wipes (red blocks), as shown in Figure 3 were established. There are three Slovak TV news streams: STV 1, MARKIZA, JOJ, where the abrupt dominant transition are localized.

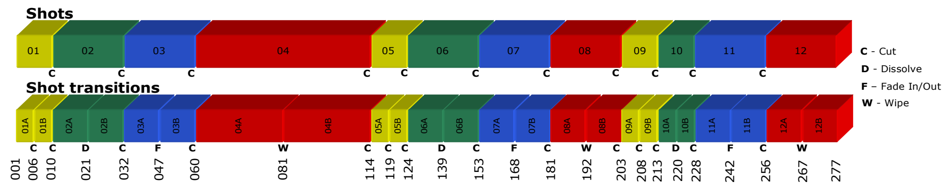


Fig. 3. Shot transition structure of the test video.

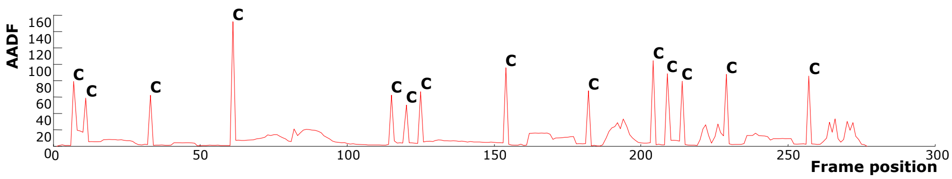


Fig. 4. Averaged Absolute Frame Difference feature.

The cut detection of annotated test video is shown in Figure 4, where symbol C constitutes the shot transition – cut. The choice of the detection threshold is decisive for the success of the cut detection. Is used the difference, which exceeds high threshold – hard cut.

The estimated difference between consecutive frames is commonly used to decide whether there is cut at frame k. Is used a threshold, if the frame difference from frame k-1 to frame k exceeds a given value t (threshold), then there is cut at this position.

5. Conclusion

In this work is the brief description of the method for video cut detection - the SBD algorithms testing, an overview of shot boundary detection and the possible algorithms of shot detectors. Scene - cut detection is usually the initial step and an important part of video segmentation, which has various applications in a variety of fields such as video retrieval, indexing, analysis, semantic description and compression. The future goal is to find the best discriminator that maximizes the overall accuracy of the transitions detection in three Slovak TV news streams.

Acknowledgement

This paper has been supported by the VEGA grant agency no. 1/4066/07, "New systems and principles of semantic description and retrieval of multimedia content".

References

[1] John S. BORECZKY, J.S., ROWE, L., A. *Comparsion of video shot boundary detection techniques*, Computer Science Division – EECS, Journal of Electronic imaging.

[2] KEITH, J. *Video Demystified*, Fourth Edition, 960 pages, Publisher: Elsevier-Newnes, 2005, ISBN: 0750678224.

[3] KRAAIJ, W., OVER, P., IANEVA, I., SMEATON, A.,F. *Trecvid 2006 ,An overview*, (2006) TRECVID 2006 - an overview. In: TRECVid 2006 - Text REtrieval Conference TRECVid Workshop, 13-14 November 2006, Gaithersburg, Maryland.

[4] NACI, U, HANJALIC, A. *Spatiotemporal Block Based Analysis of Video for Fast and Effective Shot Transition Detection and Identification,* Trecvid 2005.

[5] IONESCU, B., LAMBERT, P., COQUIN, D., Buzuloiu, V. *The Cut Detection Issue in the Animation Movie Domain*, Journal of Multimedia, vol. 2, no. 4, august 2007

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Electronic Commerce and Its Legal Rules in CR

* Dagmar Frendlovská
* VUT Brno, faculty of business, information institute, Kolejni 2906/4,61200, Brno, Czech Republic,
d.frendlovská@seznam.cz

Abstract. The article presents electronic commerce in the Czech Republic, and features current legal documents the aim of which is to govern, administer and check e-commerce. It says that ecommerce has to follow some rules as well as other kind of businesses.

Keywords: E-commerce, legal rules for standard business, legal rules for e-commerce, global trade, and global competition.

1. Electronic commerce in 21st century

Electronic ways of shopping appear to be both extremely effective and progressive ways of direct purchase and sale in the present times. The terms electronic commerce or e-commerce represent a wide range of facilities. E-commerce has become integral part of the world and domestic economies proving a rapid annual growth by over 100 per cent. The examples of e-commerce activities are as follows: purchase orders by means of ´Electronic Exchange of Dates´ (EDI), the use of fax and email, the credit card, the internet and on-line services. All of these means make it possible to place contracts on the ´electronic market´, which is, however, comparable with the ´classical market´[2].

2. Aspects of electronic commerce

Each way of doing business is supposed to be liable to certain legal rules, which allows for the observation of rightness or legitimacy of each attempt. It is possible to say that the internet has not been subject to any specific arrangements yet, and that is why there may not exist many documents focused on e-commerce regulations or specifications in the Czech Republic.

Nevertheless, e-commerce has become one of the most needed and one of the most interesting areas noticed and frequently used by a large number of experts, specialists and scientists. All this opens a great chance for companies to invest, to develop their activities and to become successful entrepreneurs [1].

3. Joint statement of EU and USA concerning electronic commerce

The importance of the whole of e-commerce phenomenon can be illustrated by a large number of conferences, declarations and statements that have been reflecting the most recent processes of investigations and innovations in the area. The global e-business significance was evident as early as 10 years ago when the Joint Statement of the EU and the USA Concerning

Electronic Commerce was released on 5th October 1997. Then, e-business was presented as part of perspective future based on the following ideas [3]:

- Global electronic commerce will become a significant and essential driving force of the 21st century world economy.

- E-commerce will offer new job opportunities in most parts of the world.

- Small companies will be able to do business worldwide in a wide range of goods and services.

- E-commerce will increase productivity in all sectors of economy. It will facilitate multilateral trade, and will support the existence of new business areas and new forms of marketing, purchase and sale.

Large global competition (i.e. among companies of all sizes with low-cost investments) will result in a larger variety of choice on the consumer side and in this way will boost economic activities and innovations (e.g. logistic aspects will play a more important role in e-commerce in terms of mail order trade which already influences further logistic development).

4. The subject matter in view of JUDr. Jakub Šváb

JUDr.Jakub Šváb presents his own view of the subject matter in the article of 10th April, 2003 as follows: „Currently, electronic commerce as such is liable to no legal rule. Certain regulations take e-commerce into account, be it the Commercial Code or the Electronic Signature Act. However, there is no comprehensive norm so far which would govern both rights and duties dealing with e-commerce. On top of it, the legal nature of the Internet seems to be rather complicated. The Internet is neither a tangible thing nor a purely incorporeal good, such as franchise or asset value. That is why the “traditional law“ based classifications of some negotiations carried out via internet may appear difficult[3].

5. General view of legal documents for e-commerce

From the general standpoint we may say that the appropriate protection of intangible assets as well as legal interests should be ensured by uncompromising and unanimous application of the existing legal documents. However, so far dispensable decision making court which would standardize the solution of current problems and situations that appear on the Internet by introducing model patterns used in similar cases would be a most welcome body to be established [3].

The collection of the most important legal documents and rules in the Czech Republic which in my opinion deal with e-commerce comprises the following:

- the Electronic Signature Act. 227/2000
- Personal data security Act. 256/1992
- Intellectual property security acts
  - Copyright/Authors´ Act.121/2000
  - Patents´ Act . 527/1990

6. Rules for e-commerce and standard business

As a rule, the regulations that apply to both standard business and e-commerce are identical. In case an entrepreneur does business without having the right trade licence, he or she will offend the law. It will make no difference whether it was “only“ an e-commerce transaction or not. In the same sense, the legal mode of contracts entered by means of guaranteed electronic signature is the same as the legal mode of contracts entered by classical means. All business relations will abide by the rules of the Commercial Code.

The UN commission for international trade and law (UNCITRAL) prepared a so called model legal code for e-commerce. However, the model legal code has not become an obligatory rule of law yet. There is a possibility, though, that parties involved in the contract may accept the proposed rule of law as obligatory [3].

7. Conclusion

The systematic use of e-commerce in the present times offers a large proportion of advantages and many ways of its use on the one side, on the other side there appear worries, states of uncertainty, and also expectations and visions that should be dealt with by means of specifically e-commerce defined rules and regulations, as well as particular laws. Such rules, regulations and laws do not exist so far, and that is why it is essential to follow the existing documents valid in standard business. However, in view of worldwide increasing electronic commerce activities the need of specifically e-commerce defined legal documents is becoming more and more urgent.

References

[1] DVORAK,J.,KRIZ,J., DVORAK,J. *Electronic Business.Lecture note* VUT Brno, faculty of business 2005.

[2] DVORAK,J. *Progress of electronic business and education of officers in public service.* IN: Sborník příspěvků z konference “Vzdělávání úředníků ve veřejné správě”. Olomouc 2005. s. 22-28., CDV Univerzity Palackého, Olomouci. ISBN 80-244-1142-3.

[3] SVAB,J.*Electronic business*. 2003 [online]. [cit. 2008-04-12].
http://www.businessinfo.cz/cz/clanek/orientace-vpravnich-ukonech/elektronicky-obchod/1000818/7013/

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Automobile Industry and Economic Crisis

[*]Peter Gajdzik, [**]Marián Šulgan
[*]University of Žilina, Faculty of Operation and Economics of Transport and Communications, Univerzitná 8215/1, 010 26 Žilina {Peter.Gajdzik@volkswagen.sk}
[**]University of Žilina, Faculty of Operation and Economics of Transport and Communications, Univerzitná 8215/1, 010 26 Žilina {Marian.Sulgan@fpedas.uniza.sk}

Abstract. This paper brings information about the automobile industry in the situation when automobile factories fall into economic problems.

Keywords: Automobile factory, Restructuring, Cost reduction, Innovation.

1. Introduction

The European Union belongs to the world automobile manufacturers producing 18 millions automobiles annually, i.e. approximately one-third of the world production. The first problems within the automobile industry have appeared during the last half of the year as a result of consumer resistance in countries like France and Spain.

Facts: Since this year automobile factories started to have been influenced by economic crisis completely, therefore the member states of the European Union adopt survival plans. Behind all problems there is consumer resistance to sale of new automobiles. In 2008 the number of sold cars in the European Union decreased by 1, 2 millions. In January 2009 there was 27% decrease compared to the previous year [2].

All these data only confirm that automobile industry is in a recession and automobile sales curve indicates downtrend. The current situation makes a pressure on the automobile producers who are forced to rationalize production due to both reduction in orders and decline in sales.

When automobile factories wish to be further competitive and survive at this unfriendly times they have to focus on the following activities:

1. Restructuring
2. Cost reduction
3. Innovation and model offensive

2. Company restructuring

The main factor that is to be applied within restructuring is absolute pragmatism. It is necessary to treat each stage of the project in consideration of the specific and precisely defined goal; otherwise the given changes fail to achieve the desired effect.

In general, the process of restructuring is understood to be some specific methodological approach to relocation of company's resources. In addition, as one of the possible means of management changes includes rights, liability and correlation of all subjects within the business activity. It is just optimization of these two parts that constitutes a basis of any internal and external restructuring efforts. The whole restructuring process of the company is divided into stage of consolidation and development, each with its further subdivisions. First of all, company (concern) is reduced that makes it more effective in order to expand gradually again later. The consolidation stage is characterized by reduction in activities and searching for the internal reserves of the company.

At the present crisis the automobile factories trouble with lack of funds and surplus production capacities. Therefore, within the consolidation process liquidation is to be ensured firstly and then outlook and cleansing of economy has to be taken. In other words, it is necessary to know producer's situation, what is the source of his profit but also of loss, prices of products and the capacity of company. Following it deficiencies, mostly within orders, are to be eliminated in order to ensure everything on time. Then, reorganization of human resources should follow, i.e. modification of organization structure, what relates to employee saving but also to reallocation of responsibilities and competences. The last stage of the consolidation process is represented by motivation of people; mainly it belongs to the tasks of the subject who manages the whole change. As soon as the consolidation process is completed everyone within the company should be fully employed. The noticeable output is cost saving.

The development stage requires both vision and strategy. Firstly, it is important to define what kind of product the manufacturer wishes to produce in the future. Then, within the second step of the development stage comes up to the modification of management system, most often the company passes to the procedural management, eventually the already existing management adapts to new conditions. As to the next step, the producer should focus on revenues and margins because profit is essential for the functionality. The fourth step is focused on investing in either technologies or innovations. The last step of the development stage can be marked as acquisition - the company decides, for example to buy some competitor or already existing factory; in fact, when searching for the subject convenient for association it tries to gain a new strategy investor.

3. Cost reduction

The most effective way how to increase market competition is to reduce price that also means to reduce costs. The costs should be mainly reduced within the following areas:

Production and logistics

Transport of the individual components to consumer represents the largest cost-share of the total logistics costs. Larger distance results in cost increase. On the other hand, the more suppliers and the lower distance make large savings. To the typical example belong automobile factories Porsche AG – all their important suppliers reside in distance of max. 200km from the factory. In case of such distances small component deliveries in the real time JIT (just in time) are taken into consideration. Implementation of JIT philosophy into the production process results in increasing labour productivity, reduction in both reserves and storage accommodation, and quality improvement. At the present time the application of JIT strategy within realization of inter-companies supplies represents an essential advantage for large consumers. In fact, specific order seems to be only some kind of "foreign storage on wheels with components" that leads up to its assembly line.

Personnel costs

Personnel costs within the automobile industry represent 15 up to 20 % of the total costs. However, during the current economic crisis their amount increases because even though the company does not produce it is obliged, under the Labour Code, to bear the costs of the permanent employees. Both contract employees and job agencies represent one of the possible solutions because they offer opportunity to respond to changes in market requirements. In reality, labour costs (per 1 employee) of contract employees are higher than of permanent employees. And in case of production stoppage the absolute pragmatism belongs to the key factors of the restructuring projects because anytime within the projects takes into consideration specific and precisely defined goal, being further elaborated into detailed fragmental steps which are, despite occasional necessary recondition with respect to parties concerned, fulfilled in a precise way. However, they vanish due to changes in market demand - the agencies are paid for the realized services having been ordered in advance.

4. Innovation and model offensive

The period when automobile industries experience drop in automobile sales is suitable for innovations as well as their introduction to practice. What makes the situation easier is the fact that introduction of innovation within the production process at the time of maximum demand is ineffective due to losses arising from introduction of innovation. Therefore, from the economic point of view the period suitable for introduction of innovations is rather in times of the economic crisis when production capacities do not indicates high usage level. Existing factory units suffice to meet demand but also prepare for new, more technological and advanced units. In case of the economic crisis end, all these newly created factory units will take advantages of it opposite to their competition. Currently, automobile factories focus mainly on concept of so-called “hybrid power” that combines either the power of petrol or Diesel engine with electromotor and therefore offers the engine power of the above standard, moreover, with low emission and improved economics of combustion. However, this conception remains dependent on petroleum fuels.

Transportation is responsible for approximately 71 % of petroleum consumption in EU. Automobile transport is 98% dependent on petroleum fuel, remaining 2% comprises of other power fuels (mainly electricity, gas). European Union tries to reduce sector dependence on petroleum fuel, therefore set the following target – until 2020 to substitute 20% of the original fuels for alternatives [1]. One of the possible alternatives is hydrogen. The main advantages are:

- belongs to the most often element in the world,
- is an universal energy bearer - can be made from any source,
- it can reduce both dependence on petroleum and CO2 emissions and pollution caused by transport – reach a null value is possible in practice

Automobiles based on electromotor driven by hydrogen fuel elements are more effective, they use 40-60% of monopropellant energy.

Battery-driven vehicle represents other kind of alternative. Power source is usually accumulator that has to be charged from the external source one day before driving. Trailing throttle depends on accumulator capacity. Accumulator belongs to the weakest elements of this alternative whereas its capacity does not allow the similar kilometrical power as combustion engines. Technical break within the battery-driven vehicles development got started with development of electricity, control systems and new types of accumulators (for example SCIB batteries), but mainly with fast increase in energy costs, i.e. fuels – mainly petroleum and gas. To the main advantages belongs also simpler automobile design with location of power directly in wheels, electronics and accumulator set located in the floor part. Automobile body construction can be therefore constructed in a more variable way within the same car chassis.

5. Conclusion

Automobile industry belongs to the key factors in economics of the European Union states, therefore when automobile factory falls into economic problems it is necessary to:

- provide the key changes within the company direction
- change the structure of the company
- eliminate or reconstruct some company's components
- encourage self-confidence and work on innovations

Only those companies which are able to adapt new conditions of the market are able to survive. Whereas authorized capital for both moderation of immediate crisis in times of indebtedness of company and investment support to fixed assets or sales has only short-term effect; on the other hand, development, research and implementation of new knowledge to mass production help to ensure company survival as well as its perspective growth in the future.

References

[1] VEGA Project 1/0398/08: *Negative-Externality Reduction in Road Transportation, Zilina University, FPEDAS, 2008-2010.*

[2] www.euractiv.sk, *March 16th, 2009.*

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Multiple Beam Antenna Arrays Synthesis Using Different Array Weightings Functions

*Ján Haring, *Norbert Majer, *Peter Poloha, *Rudolf Hronec
*University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 01026 Žilina, Slovak Republic,
{Jan.Haring, Norbert.Majer, Peter.Poloha, Rudolf.Hronec}@fel.uniza.sk

Abstract. A linear phased antenna array is a system compound from simply radiators (dipoles, microstrip antennas) that together form desired radiation pattern. Phased array antennas consist of multiple stationary antenna elements that are fed coherently and use variable phase or time-delay control at each element to scan a beam to given angles in space. Variable amplitude control (adjusted using weight functions) is sometimes also provided for antenna pattern shaping. Arrays are sometimes used in place of fixed aperture antennas (reflectors, lenses) because the multiplicity of elements allows more precise control of the radiation pattern, thus resulting in lower sidelobes or careful pattern shaping. However, the primary reason for using arrays is to produce a directive beam that can be repositioned (scanned) electronically. An array can form multiple narrow beams towards different directions.

Keywords: Linear antenna array, radiation pattern, antenna element, phase control, sidelobes, pattern shaping, multiple beam, spectrum weights.

1. Introduction

A phased antenna array uses an array of antenna elements. Signals induced on each array element are combined to form a single output of the antenna array. This process of combination of signals from different array elements is known as beamforming.

For a given array the beam may be steered to different directions electronically. Beam steering can also be accomplished by appropriately delaying the signals before combining. This can be realized by means of signal phase shifters.

The term adaptive antenna is used for a phased antenna array when the weighting on each array element is applied in a dynamic fashion. The amount of weighting on each channel is not fixed at the time of the array design. In other words, the antenna array pattern adapts to the situation and the adaptive process is under control of the steering system.

The term smart antenna incorporates all situations in which a system is using an antenna array and the antenna pattern is dynamically adjusted by the system as required. Thus, a system employing smart antennas processes signals induced on a sensor array [2].

The term multibeam array defines an antenna array that can form multiple narrow beams towards different directions. This multibeam array may be used in adaptive antenna systems or smart antennas systems. So we can receive or transmit the signal to several directions with the same antenna array.

1.1. Multibeam Arrays

Multibeam array can form multiple beams towards different directions. For example, it is desired to form three beams towards the steering angles θ_1, θ_2, θ_3.

The weights for such a multibeam array can be obtained by superimposing the weights of a single broadside array, say $w_{(n)}$, steered towards three angles [3].

Equation (1) described corresponding scanning phases for three angles.

$$\varphi_i = kd\cos\theta_i \,. \tag{1}$$

For i = 1, 2, 3. Equation (2) described array factor for multibeam antenna array steered toward three angles θ_1, θ_2, θ_3.

$$AF(\theta) = \sum_{n=0}^{N-1} w_{(n)} e^{+jn(kd\cos\theta + \varphi_1)} + w_{(n)} e^{+jn(kd\cos\theta + \varphi_2)} + w_{(n)} e^{+jn(kd\cos\theta + \varphi_3)} \,. \tag{2}$$

Where :

k – wave number,

d – distance between antenna elements,

φ – scanning phase.

The basic broadside array weights $w_{(n)}$ can be designed to achieve a desired sidelobe level or beam width.

2. Multibeam Array Synthesis

We use combination of Woodward – Lawson frequency – sampling design method and Fourier transform method for multibeam array synthesis. The steps in the algorithm are:

We expressed the beam pattern in φ coordinates for three steering angles θ_1, θ_2, θ_3.

$$B(\varphi) = \sum_{i=1}^{3} \left(e^{-j\left(\frac{N-1}{2}\right)(\varphi-\varphi_i)} \sum_{n=0}^{N-1} w_n^* e^{jn(\varphi-\varphi_i)} \right), \quad -\frac{2\pi d}{\lambda} \le \frac{2\pi d}{\lambda} \,. \tag{4}$$

The beam pattern is sampled at:

$$\varphi_k = \left(k - \frac{N-1}{2} \right) \frac{2\pi}{N}, \quad k = 0,1,\ldots N-1 \,. \tag{5}$$

Where N is number of antenna elements. We use equation (6) to determinate $B(k)$.

$$B(k) = B^*(\varphi_k) e^{-j\varphi_k\left(\frac{N-1}{2}\right)} \,. \tag{6}$$

We find b_n as the IDFT of $B(k)$ using equation (7).

$$b_n = \frac{1}{N} \sum_{k=0}^{N-1} B(k) e^{jkn\frac{2\pi}{N}} \,. \tag{7}$$

We use equation (8) to find new array weights for multibeam array.

$$w_n = b_n e^{-jn\pi\left(\frac{N-1}{N}\right)} \,. \tag{8}$$

2.1. Array Weightings Functions

There are several classes of array weightings that are equivalent to various windows or tapers used in the spectral analysis of time series [4]. The utilization of array weightings causes that the height of 1st side lobe decreases and the main beam width increases. We used these array weightings functions for the most commonly used array weightings are there:

Name	Mathematical expression	Weight parameter
Uniform weighting	$w_N = \frac{1}{N}$	
Cosine weighting	$w_i = \sin\left(\frac{\pi}{2N}\right)\cos\left(\pi\frac{i}{N}\right)$	$-\frac{N-1}{2} \le i \le \frac{N-1}{2}$ (for all weightings)
Raised cosine weighting	$w_i = c\left[p + (1-p)\frac{\pi i}{N}\right]$	$c = \frac{p}{N} + \frac{(1-p)}{2}\sin\frac{\pi}{2N}$ $p = 0{,}2$
Hamming weighting	$w_i = g_0 + g_1\cos\left(2\pi\frac{i}{N}\right)$	$g_0 = 0{,}53836$ $g_1 = 0{,}46164$
Kaiser weighting	$w_i = I_0\left(\beta\sqrt{1-\left(\frac{2i}{N}\right)^2}\right)$	$\beta = 3$
Blackman – Harris weighting	$w_i = a_0 + a_1\cos\left(\frac{2\pi i}{N}\right) + a_3\cos\left(\frac{4\pi i}{N}\right)$	$a_0 = 0.42$ $a_1 = 0.5$ $a_3 = 0.08$

Tab. 1. Spectrum weight functions

3. Experiments

We compare monobeam antenna gain and multibeam antenna gain for three steering angles θ_1, θ_2, θ_3. The monobeam antenna array is steered to the direction with maximum incoming signal strength. The frequency is 2 GHz.

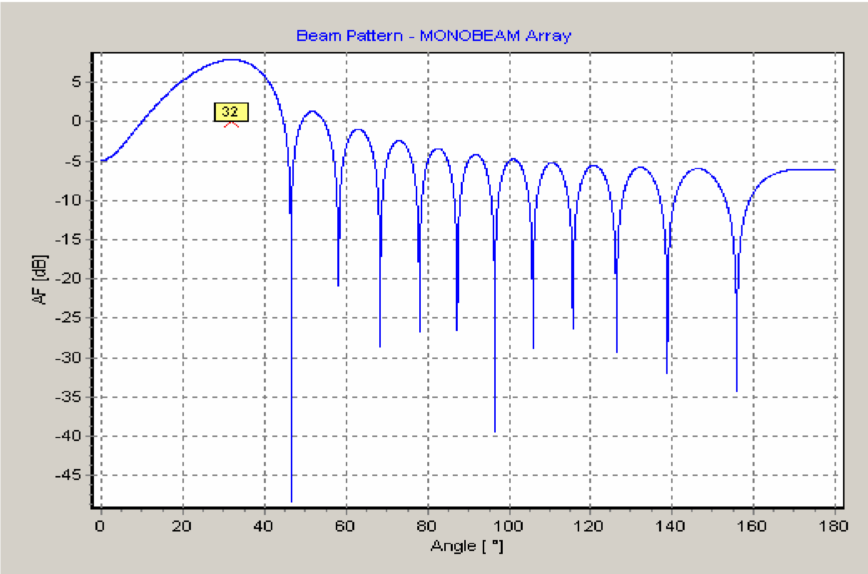


Fig. 1. Monobeam array beam pattern.

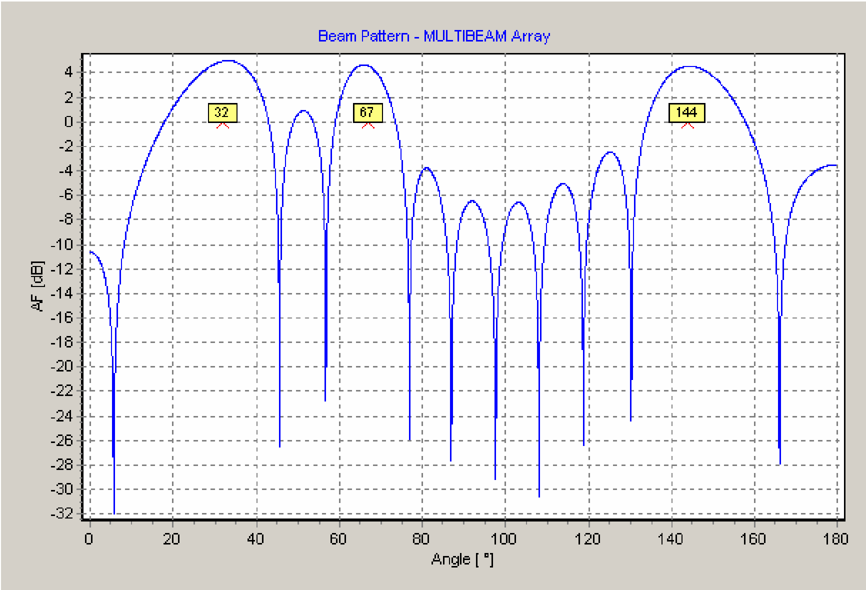
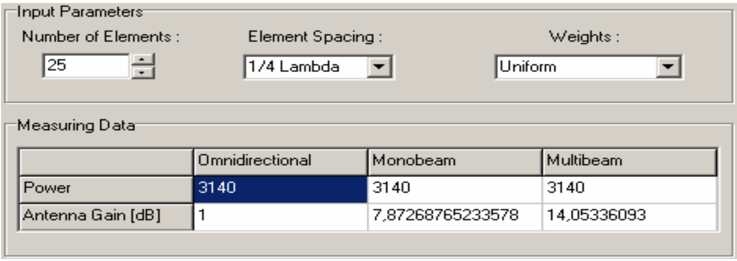


Fig. 2. Multibeam array beam pattern.



	Omnidirectional	Monobeam	Multibeam
Power	3140	3140	3140
Antenna Gain [dB]	1	7,87268765233578	14,05336093

Fig. 3. Input parameters and measured data.

4. Conclusion

In this paper we describe multibeam antenna array theory, simulate this system and describe antenna gain. Antenna array simulations are created by the help of program Antenna Arrays Simulator. This program was created in Borland Delphi 7. The final results are described in Fig. 1, Fig. 2 and Fig. 3.

Acknowledgement

The work on this paper was supported by the grant VEGA 1/4067/07 of Scientific Grant Agency of the Slovak Republic and by institutional grant 28/103150 of Faculty of Electrical Engineering, University of Zilina.

References

[1] GODARA, CH. *Smart Antennas*. New York: CRC Press LLC , 2004.

[2] MAILLOUX, R. *Phased Array Antenna Handbook*. Norwood: Artech House, 2005.

[3] ORFANIDIS, S. *Electromagnetic Waves and Antennas*. www.ece.rutgers.edu/~orfanidi/ewa, 2008.

[4] VAN TREES, H. *Optimum Array Processing – Part IV*. New York: Wiley, 2002.

Comparison of IT Service Management Approaches

*Zuzana Hnatová, *Iveta Kremeňová, *Margita Majerčáková
*University of Žilina, Faculty of Operation and Economy of Transport and Communications, Department of Communications, Univerzitná 1, 010 26 Žilina, Slovakia,
{Zuzana.Hnatova, Iveta.Kremenova, Margita.Majercakova}@fpedas.uniza.sk

Abstract. IT executives implement best practices because they need to increase IT predictability and efficiency, reduce support costs, improve customer service quality or meet regulatory requirements. The two most well-known standards - the IT Infrastructure Library (ITIL) and the Control Objectives for IT (COBIT) - have existed for at least 15 years, support a broad range of management services, and have been implemented by thousands of organizations of all sizes.

Keywords: IT service, IT Service Management, ITIL, COBIT.

1. Introduction

IT organizations are under increasing pressure to meet the business goals of their companies. Compliance requires strong corporate governance capabilities that are demonstrable to outside auditors. Because IT plays such a major role in business processes, the IT organization not only creates complexity for the business, but at the same time, provides the means to demonstrate this compliance.

Organizations rely on guidelines such as the IT Infrastructure Library (ITIL) and Control Objectives for Information and related Technology (COBIT) to help understand and address these challenges.

ITIL and COBIT can enable organizations to achieve three objectives:

- Establish proven best practice IT service management processes to manage IT from a business perspective and achieve business goals, including that of compliance.
- Put in place clear process goals, based on the organization's business goals, and provide a means of measuring progress against them.
- Ensure effective IT governance and control at the process level, and enable IT to demonstrate that it meets or exceeds the requirements set forth by government or external regulations. [3]

There is, however, confusion in IT organizations concerning these frameworks. Some think they are two alternate approaches to the same goal, and others think they are mutually exclusive. Actually, they are highly complementary, and together provide greater value than using just one or the other. **COBIT outlines what you need to do to meet these challenges and ITIL shows you how to get there.**

2. Approaches to IT Service Management

The INS (International Network Services) company conducted a web-based survey on approaches for IT Service Management, which was completed by 227 IT professionals around the globe in 2006. This survey was designed to help organizations manage their IT service processes, and the reasons driving its adoption. It is a follow-up to a survey on IT Service Management conducted by INS in September 2004.

Based on this research from year 2004 and 2006, *ITIL and COBIT are most widely used frames for managing IT services.*[7]

2.1. Comparison of ITIL and COBIT

IT control definition, testing and progress measurement are task categories that are natural COBIT strengths. The COBIT model is very specific in its definition of the processes and the auditable controls that need to be in place to ensure reliable and predictable IT processes.

The processes defined in COBIT are grouped into four separate domains that align with the IT implementation cycle. They are: Planning and Organization, Acquisition and Implementation, Delivery and Support, and Monitoring. [2]

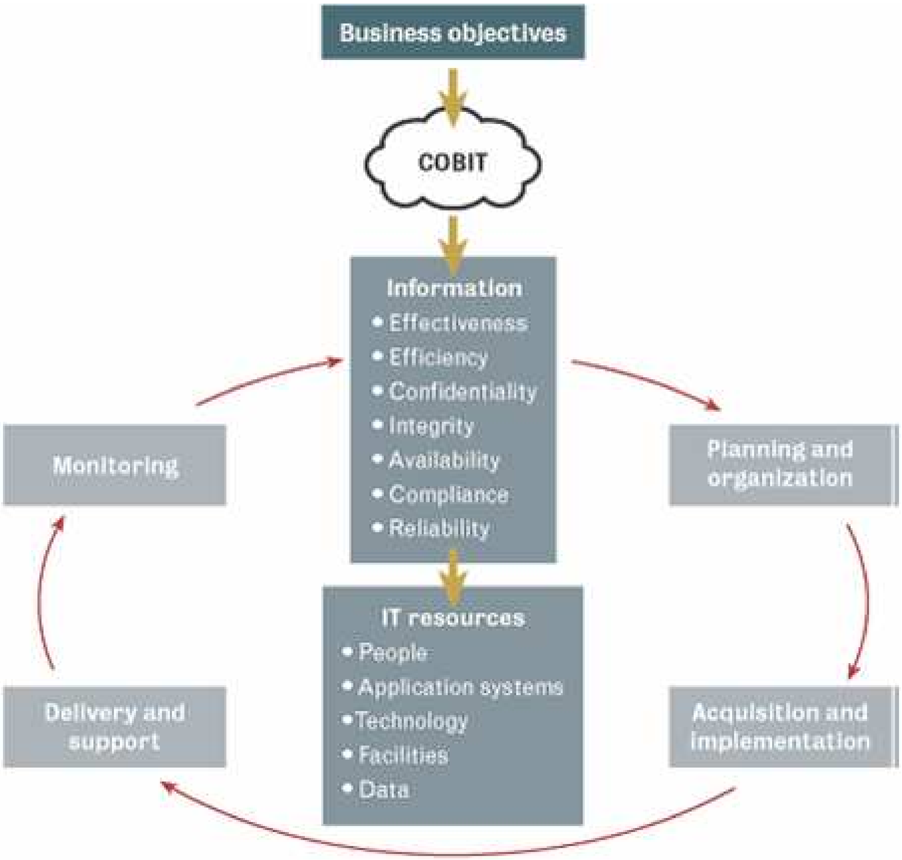


Fig. 1. The COBIT model. [2]

COBIT and ITIL are more complementary than they are competitive.

COBIT focuses on the definition, implementation, auditing, measurement and improvement of controls for specific processes that span the entire IT implementation life cycle. As such, it is an excellent reference model for IT governance across the entire implementation life cycle.

The primary focus of ITIL is to provide best practice definitions and criteria for operations management. More specifically, ITIL primarily focuses on defining the functional, operational and organizational attributes that need to be in place for operations management to be fully optimized in two key categories. These categories are called Service Support Management and Service Delivery Management, each of which has a number of supporting subcategories.

The management subcategories for Service Support Management include Service Desk, Incident, Problem, Configuration, Change and Release management, while those for Service Delivery Management include Service Level, Financial, Capacity, Service Continuity and Availability.

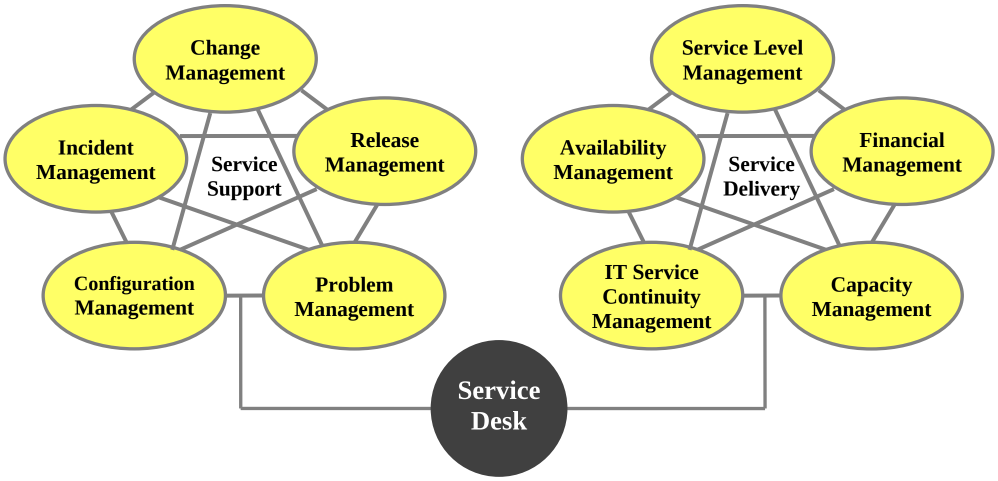


Fig. 2. ITIL Service management processes.

Each subcategory definition includes best practice criteria for many areas, including organizational support, cross management component integration, management reporting, product capability, implementation quality and customer service quality.

If the goal is improving the quality and measurability of IT governance across the entire networked application implementation life cycle or implementing a control system for improved regulatory compliance, COBIT would be a more effective choice.

On the other hand, if the objective is to continuously improve IT operations efficiency and IT customer service quality, ITIL would be the better bet. [5]

Comparison between these two approaches is in the following table 1.

	COBIT	ITIL
Function	Mapping IT process	Mapping IT service level management
Area	4 process and 34 domain	9 process
Issuer	ISACA, ITGI	OGC
Implementation	Information system audit	Manage service level
Consultant	Accounting firm, IT consulting firm	IT consulting firm

Tab. 1. Comparison between ITIL and COBIT approach for managing IT services.

As we can see from the table, ITIL is registered trade mark of the OGC (The Office of government Commerce), which is the government organization of the United Kingdom.

COBIT is registered trade mark of the ISACA (Information System Control Standard) non-profit organization and ITGI (The IT Governance Institute). [3]

3. Conclusion

The main aim of this article was to introduce two main approaches for IT service management, which is ITIL and COBIT.

ITIL is a stand for Information Technology Infrastructure Library. It is a set of framework for managing IT services, issued by Office of Government Commerce.

On the other side, COBIT is a stand for Control Objectives for Information and related Technology. The COBIT main function is to help the company, mapping their IT process.

Although ITIL is quite similar with COBIT in many ways, but the basic difference is, that COBIT set the standard by seeing the process based at risk, and in the other hand ITIL set the standard from basic IT service.

COBIT is usually chosen by the company who performing information system audit, whether related to financial audit or general audit IT audit.

The article also contains the comparison of these two IT service management approaches.

Acknowledgement

I would like to thank assoc. prof. Jana Šujanová, CSc. from Institute of Industrial Engineering, Management and Quality, MTF STU Trnava, who support our research with providing necessary materials for studying.

References

[1] BROOKS, P. *Metrics for Service Management.* Norwich (UK): Van Haren Publishing, 2006.

[2] IT GOVERNANCE INSTITUTE. *COBIT: 4.1 Excerpt.* Rolling Meadows (Ilinois, USA): IT governance institute, 2007.

[3] MARQUIS, H. *ITIL and COBIT.* ITSM solutions DITY newsletter reprint [online] May 2006, Volume 2.20 [cited 21.november 2007]. Available on: <http://www.itsmsolutions.com/newsletters/DITYvol1iss1.htm>

[4] OFFICE OF GOVERNMENT COMMERCE. *The official introduction to the ITIL service lifecycle.* 1st ed. London: The Stationery Office, 2007.

[5] OFFICE OF GOVERNMENT COMMERCE. *Service Delivery.* London: The Stationery Office, 2000.

[6] OFFICE OF GOVERNMENT COMMERCE. *Service Support.* London: The Stationery Office, 2000.

[7] BLUM, R. *IT Infrastructure Library (ITIL).* INS IT industry survey [online] July 2006, [cited 22.august 2008]. Available on: <http://www.ins.com/knowledge/surveys.asp>

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Comparison of Industrial Communication Buses

*Peter Holečko
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 1, 010 26 Žilina, Slovakia, holecko@fel.uniza.sk

Abstract. The paper deals with comparison of chosen most common industrial communication buses used in automation area. Three groups of characteristics are formulated, such as physical, transport mechanisms and performance characteristics.

Keywords: Industrial communication buses, comparison, physical characteristics, transport mechanisms, performance.

1. Introduction

Generally, communication represents information exchange between participants, which may be technical devices, too. The information transmission is realised via a so called communication cycles which include activities associated with connection establishment (subscribers interconnection), information transmission (often in a delimited time, for example in real-time) and connection termination.

Because of a more explicit interpretation of important digital communication terms it is useful to point out some general definitions [1].

Interface is a common boundary between two communicating units, which is defined by a characteristics of physical, signal and procedural interconnection (a so called interface aspects). The basic general interface parameters are the width of transmission paths (number of parallel bits), transfer rate (B/s), error resistance and extensibility (modularity).
Bus is an interface with three or more units connected. The bus is defined by a set of physical connections including the relevant conventions, i.e. protocols for information transmission within communication network. The bus transmits addresses, data, commands and states, therefore we distinguish the address, data, control and state buses.
Protocol is a set of rules how to initialise, realise and terminate the communication process. The control algorithms of data transmission within a protocol represent a group of guidelines agreed between communication participants. These algorithms radically don't refer to a specific devices communicating. In automation equipment within serial communication the byte-oriented and bit-oriented control practices are implemented.

2. Industrial communication buses comparison

Tables 1 to 3 represent the comparison of individual industrial communication buses in chosen areas based on specified criteria [2],[3],[4]. These data are usable during statement of initial or limiting requirements for given application and potentially for setting a narrower group of technologies suitable for subsequent testing and deployment.

Table 1 summarises the characteristics of individual buses based on capabilities of usable topologies, physical transmission media, maximum number of nodes connected and maximum coverage. In addition, references to individual bus technology project or vendor sites are inscribed.

	Topology	Physical media	Max. number of nodes	Max. coverage
PROFIBUS [11]	line, star, ring	twisted-pair (TP), fibre (F)	127 nodes + 3 master units	100m between segments (TP), 24km (F)
INTERBUS-S [9]	segmented with „T“ drops	TP, F, slip-ring	256 nodes	400m/segment, 12,8km overall
DeviceNet [13]	trunkline/dropline with branches	TP for signal and power	64 nodes	500m, 6km with repeaters
AS-I [5]	bus, ring, tree, star	2-wire cable	31 slave units	100m, 300m with repeater
Foundation Fieldbus H1 [7]	star or bus	TP, F, radio (R)	240 nodes/segment, 65000 segments	1900m
Foundation Fieldbus High Speed Ethernet (HSE) [7]	star	TP, F (R)	practically unlimited	100m (TP), 2000m (F)
WorldFIP [8]	bus	TP, F, R	256 nodes	40km
LONWorks [10]	bus, ring, loop, star	TP, F, R, power line	32000 nodes/domain	2000m
ControlNet [12]	line, tree, star or combination	coax (C)	99 nodes	1000m 2 nodes (C), 250m 48 nodes, 3km (F), 30km (F) with repeaters
CANopen [6]	trunkline/dropline	TP	127 nodes	25-1000m
Industrial Ethernet [14]	bus, star, daisy-chain	thin coax, thick coax, TP, F	1024 nodes, expandable via routers	185m (C), 100m (TP), 2,5-50km (F)

Tab. 1 Physical characteristics

Table 2 characterizes the technologies in term of applied transport mechanisms and reached parameters, such as communication methods, transmission properties (rates), data unit size, arbitration media, error checking mechanism and diagnostics.

	Comm. methods	Transmission properties	Data unit size	Arbitration method	Error checking	Diagnostics
PROFIBUS	master/slave peer to peer	9,6-500kb/s	0-244 B	token passing	HD4 CRC	station, module, channel
INTERBUS-S	master/slave	500kb/s	1-64 B data, 246 B parameters	-	16b CRC	segment localisation, cable
DeviceNet	master/slave, multi-master, peer to peer	125-500kb/s	8 B	CSMA	CRC	bus
AS-I	master/slave	data and power, EMI resistant	31 slave with 4 in. and 4 out.	master/slave with cyclic polling	Manchester, Hamming	slave, device
Foundation Fieldbus H1	client/server, publisher/sub scriber, event notification	31,25kb/s	128 B	scheduler, multiple backup	16b CRC	remote diag., network, parameter state
Foundation Fieldbus High	client/server, publisher/sub	100Mb/s	variable (TCP/IP)	CSMA/CD	CRC	-

Speed Ethernet (HSE)	scriber, event notification					
WorldFIP	peer to peer	31,25kb/s – 6Mb/s (F)	unlimited	central	16b CRC	device message expiration
LONWorks	master/slave peer to peer	1,25Mb/s	228 B	CSMA	16b CRC	CRC errors database
ControlNet	producer/ consumer Device Object Model	5Mb/s	0-510 B	CTDMA Time Slice Multiple Access	modif. CCITT	duplicate node ID, device failures
CANopen	master/slave, peer to peer, multi-cast, multi-master	10kb/s – 1Mb/s	8 B	CSMA	15b CRC	error and emergency messages control
Industrial Ethernet	peer to peer	10, 100Mb/s	46-1500 B	CSMA/CD	CRC 32	-

Tab. 2 Transport mechanisms

The performance characteristics of chosen set of industrial bus technologies, such as duration of digital and analogue cycles, duration of 128-byte data block transmission, are presented in Table 3.

	Cycle time: 256 digit. 16 nodes with 16 I/O	**Cycle time: 128 analog. 16 nodes with 8 I/O**	**Block transmission of 128B 1 node**
PROFIBUS	configuration dependent < 2ms	configuration dependent < 2ms	-
INTERBUS-S	1,8ms	7,4ms	140ms
DeviceNet	2ms	10ms	4,2ms
AS-I	4,7ms	-	-
Foundation Fieldbus H1	< 100ms	< 600ms	36ms at 31,25kb/s
Foundation Fieldbus High Speed Ethernet (HSE)	- ; latency < 5ms	- ; latency < 5ms	< 1ms
WorldFIP	2ms at 1Mb/s	5ms at 1Mb/s	5ms at 1Mb/s
LONWorks	20ms	5ms at 1Mb/s	5ms at 1Mb/s
ControlNet	< 0,5ms	< 0,5ms	< 0,5ms
CANopen	< 1ms	5ms at 1Mb/s	< 2,5ms
Industrial Ethernet	application layer dependent	application layer dependent	application layer dependent

Tab. 3 Performance characteristics

3. Conclusion

The paper maps the most common industrial communication buses from the perspective of physical, transport, application and other characteristics. The result is a transparent tabular form enabling the choice of a suitable technology with respect to weights set for the individual characteristics.

Further work will be focused on formulation of requirements laid on properties and parameters of analysed protocols and buses for the purposes of implementation in distributed control systems and subsequent analysis, modelling and simulation of chosen set of protocols and buses.

References

[1] ŠTURCEL, J.: *Komunikácia v distribuovaných riadiacich systémoch*. Bratislava: FEI STU. http://www.kar.elf.stuba.sk/predmety/iss/pdf/komunik.doc (2008-05-02)

[2] *The Grid Connect Fieldbus Comparison Chart*. http://128.121.135.22/compare.htm (2009-01-26)

[3] *An 'open' FIELDBUS comparison*. http://194.203.250.243/Download.nsf/d69036b7142a102a002568ac0041c221/4c816632a428def20025675a005a26fa/$FILE/8_97fw-1.pdf (2008-09-11)

[4] *FIELDBUS*: *Which Fieldbus to go for*? http://www.weighing-systems.com/TechnologyCentre/fieldbus1.html (2008-11-29)

[5] *AS-Interface: The System*. http://www.as-interface.net/System (2009-02-13)

[6] *CAN in Automation (CiA): Technology*. http://www.can-cia.de/index.php?id=13 (2009-02-13)

[7] *Fieldbus Foundation – Overview*. http://www.fieldbus.org/index.php?option=com_content&task=view&id=23&Itemid=308 (2009-02-14)

[8] *WorldFIP overview*. http://www.worldfip.org/SHWebClass.ASP?WCI=ShowCat&CatID=1 (2009-02-13)

[9] *Interbusclub: Technology*. http://www.interbusclub.com/en/technologie/main.html (2009-02-13)

[10] *LonWorks*. http://www.echelon.com/products/lonworks/ (2009-02-15)

[11] *PROFIBUS Technology and Application: System Description*. http://www.profibus.com/celummdb/doc/PROFIBUS/Downloads/Technical_Description/PROFIBUS-PA-system-descr_e_Aug07.pdf (2009-02-12)

[12] *ControlNet Technology Overview*. http://www.odva.org/Home/ODVATECHNOLOGIES/ControlNet/ControlNetTechnologyOverview/tabid/245/Default.aspx (2009-02-13)

[13] *DeviceNet Technology Overview*. http://www.odva.org/Home/ODVATECHNOLOGIES/DeviceNet/DeviceNetTechnologyOverview/tabid/72/Default.aspx (2009-02-12)

[14] BERGE, J.: *Čo je priemyselný ethernet?* In: AT&P Journal, Vol. 9, 2002, No. 9, p. 17-20, ISSN 2335-2237.

Parallelization and Optimization of Prime Numbers Algorithm

*Andrej Holúbek
*University of Žilina, Faculty of Management Science & Informatics, Department of Technical Cybernetics, Univerzitná 2, 01026 Žilina, Slovakia, {andrej.holubek}@fri.uniza.sk

Abstract. This work is about parallelization of prime number algorithms. It is focused on algorithm called Sieve of Eratosthenes. This article describes algorithm it self, also advantages of its parallelization and additional options.

Keywords: parallel algorithms, sieve of Eratosthenes, parallelization, optimization

1. Introduction

Sieve of Eratosthenes is a simple algorithm for finding all prime numbers up to a specified number n [1].

1.1. Algorithm definition

Sieve of Eratosthenes works on gradational “sieving” list of natural numbers. At the beginning of the process of sieving, the list contains all natural numbers from 2 to n. Gradually all multiples are reject from the list of natural numbers, begins with multiples of k=2. Next number, which is not rejected, is also prime number and there for all multiples of it are reject as well. This process is repeated until $k>\sqrt{n}$. After algorithm finish the list contains all primes until n [6].

2. Serial algorithm

2.1. Algorithm process

1. List of all natural numbers up to n is created <2..n>
2. Multiples of number k are rejected from the list. (begin with prime number k=2)
3. Next not rejected number is next prime number k.
4. Step 2 and 3 are repeated until $k<\sqrt{n}$.

2.2. Serial complexity

Algorithm works on list of n numbers. In every step it has to make n/p-1 operation, where p is prime number, which multiples are rejecting. Algorithm is repeated x times, where x is number of primes. There for:

For multiples of k=2: (n/2 –1)
For multiples of k=3: (n/3 –1)
…
For multiples of k=p: (n/p –1)

$$t_{ser.} = (n/2 - 1) + (n/3 - 1) + (n/5 - 1) + (n/7 - 1) + ... + (n/\sqrt{n} - 1) \quad (1)$$

Computing complexity of serial algorithm is $t_{ser} = O(n^2)$

Where time complexity of serial algorithm is O(N*log(log N)) , where N is top limit of list [7].

2.3. Optimization of serial algorithm

Basic algorithm sieve of Eratosthenes is defined very clearly and simply, because of that it is not optimized. Some operations are unnecessarily repeated several times, therefore the processor is not used effectively.

Some of defaults in algorithm can be reform easily. For example, as is stated for basic algorithm, all multiples of defined prime numbers must be rejected. Based on commutation, it is simple to reduce number of unnecessarily repeated steps of rejecting multiples of define prime. Instead of rejecting all multiples of define prime, it can be reduce just to rejecting multiples higher then p^2, where p is defined prime. With this partial optimization algorithm can be speed up of p^2 steps.

Even after this partial optimization, some of numbers will be rejected more then once. This happens because of multi conjunction, where one number can be a conjunction of more numbers (17*30 = 17*2*3*5= 2*255=3*170=5*42 = 510). In this case number 510 will be rejected from list of numbers 4 times instead of once. A new condition can be applied for this algorithm, where will be rejected just numbers and their multiples which weren't rejected before.

This partial optimization can speed up the algorithm even more, since after rejecting of all the k=2 multiples, we can reduce the amount of the next rejection on half, with k=3 on third etc [8].

3. Parallelization of sieve of Eratosthenes algorithm

Eratosthenes sieve algorithm is simple algorithm, but his parallelization isn't so simple. Main problem is dependency of steps following consecutively. That means next step can't be started until previous step finds next prime number (Without rejecting multiples of prime 2, it is not known which is next prime, in this case prime 3). Next problem is that algorithm is working on whole field of natural numbers <2..N>. Various methods exist to solve this problems and despite of them parallelize Eratosthenes sieve algorithm. [2, 3]

Two basic models are known to parallelize serial algorithms. First model is parallelization algorithm on shared memory. Second one is parallelization on distributed memory. [4, 5]

3.1. Parallelization of algorithm on shared memory

In case of shared memory model, every process of algorithm works on one field <2..N>. However every process works with another prime and is responsible for rejecting its prime multiples. In shared memory list of boolean values is saved, where rejected multiples are marked. And also last found prime p is saved here.

Steps to parallelize algorithm on share memory:

1. Field of natural numbers <2..n> is created
2. M processes are created
3. Process(i) begins for prime $k=p_i$ (i=1,k=2)

4. While rejecting its multiples, it sets next prime (p_{i+1}).
5. Afterward Process(i+1) begins for k=p_{i+1}(i=2, k=3)
6. Steps 3 – 5 are repeated until k<√n. Processes runs in cascade.
7. If one of processes finished, next prime number will be assigned to it.
8. Algorithm finished when all processes are finished

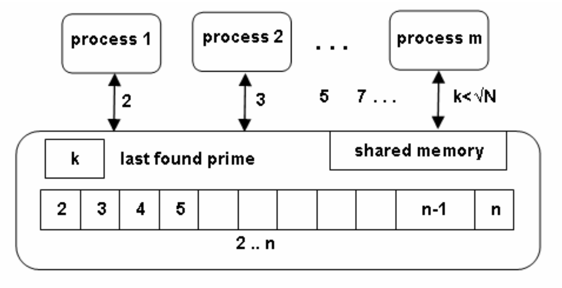


Fig. 1. Schema of parallel algorithm on shared memory

3.2. Parallelization of algorithm on distributed memory

For model with distributed memory, is necessary to divide the filed of natural numbers <2..N> to m intervals, where m is number of processes. Intervals are divided equally, therefore each segment will be n/m wide. On each segment processes run rejection of multiples for all primes p, where p are of <2..√N>. All processes work with all primes, where every process is responsible for rejecting multiples on its segment.

Steps to parallelize algorithm on share memory:

1. M processes are created
2. Field of natural numbers <2..N> is created
3. This field is split to a m equal segments
4. One segment of field is assigned to each process
5. All processes run for prime k=p_i (i=1,k=2) and reject multiples on their segments
6. Process(x) set next prime k=3 (x=1) (x is increased every time when k<=x*n/m)
7. Steps 5 and 6 are repeated until k>√n
8. Algorithm finish when all processes are finished

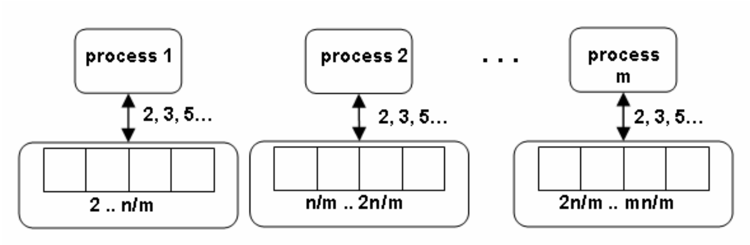


Fig. 2. Schema of parallel algorithm on distributed memory

3.3. Parallel speed up with communication requirements

As shown on figure 1, 2, computing time of parallel algorithms is much shorter compare to serial algorithm. But the strong dependency of processes can't be forgotten, therefore higher communication resources are needed. Those communication resources increase the time necessary for computing. On figure 3 is shown dependency computing time to number of processes including communication resources [9].

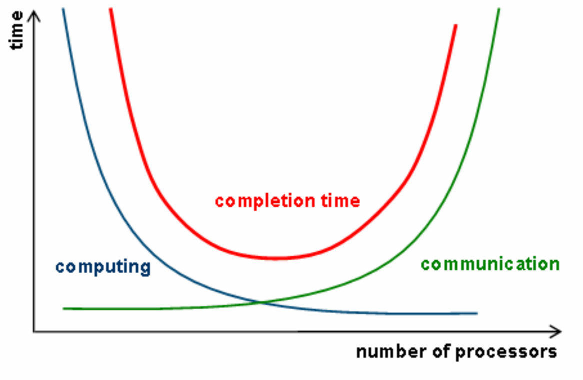


Fig. 3. Schema of parallel algorithm on distributed memory

4. Conclusion

Sieve of Eratosthenes is one of basic algorithms for finding prime numbers. It is proper algorithm for parallelization, therefore it is possible to watch influence of parallel environment factors. Sieve of Eratosthenes algorithm has high communication demands, therefore it is proper algorithm for observation influence of communication demands on overall computation time.

Acknowledgement

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I want to thank the Department of Technical Cybernetics, especially to professor Hanuliak, for giving me permission to commence this thesis in the first instance, to do the necessary research work and to use departmental data.

References

[1] O'NEILL, M.E., *The Genuine Sieve of Eratosthenes*, Association for Computing Machinery, 2006

[2] BOHNSACK M., ERHARDT E. B.,*S 442 Introduction to Parallel Processing Project 1: MPI Sieve of Eratosthenes*, 2006

[3] WILKINSON B., *Parallel Programming*, ISBN 9780131405639, Prentice Hall, 2004

[4] SOONWOOK H., KYUSIK CH., DONGSEUNG K., *Load Balanced Parallel Prime Number Generator with Sieve of Eratosthenes on Cluster Computers*, Korea University, Soongsil University, 2007

[5] BOKHARI S.H., *Multiprocessing the sieve of Eratosthenes*, IEEE Computer 20(4): April 1984, pp.50-58

[6] HARDY G., WRIGHT E., *An introduction to the theory of numbers*, Clarendon Press, 1979

[7] X. LUO, *A practical sieve algorithm for finding prime numbers*, Communications of the ACM, Vol. 32 No. 3, pp. 344-346, March

[8] BENGELLOUN S. A., *An incremental primal sieve*, Acta informatica, 23, 119–125.,1986

[9] HOLÚBEK A., *Modelovanie a optimalizácia paralelných algoritmov rozkladu veľkých čísiel na prvočísla*, ŽU, 2008

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Error Concealment Techniques for Reconstruction of Damaged Video Streams

*Róbert Hudec, *Miroslav Benčo, *Zuzana Dubcová, *Peter Lukáč
*University of Zilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitna 1, 01026 Žilina, Slovakia,
{robert.hudec, miroslav.benco, zuzana.dubcova, peter.lukac}@fel.uniza.sk

Abstract. In present communication systems, the role of higher QoS (Quality of Service) providing for video data transmission become an essential task. This paper reviews the techniques that have been developed for error concealment in the past 20 years. In this document, the various classification approaches oriented on methods for reconstructing the corrupted or lost video data are mentioned. An overview of three dominant video concealment categories, namely, Forward Error Concealment Methods, Post-processing Error Concealment Methods and Interactive Error Concealment Methods are analyzed.

Keywords: Video, error concealment, MPEG, forward error correction (FEC), video quality metrics.

1. Introduction

In present narrowband or noise sensitive communication systems, the problem with a lot of transmission error can cause the lower quality of video stream. There are two dominant types of errors. First of one marked as *random bit error* is mainly caused by transmission via imperfection channel. The other one called as *erasure error* is mainly represented by packet lost in the time, when the transmission system is failure for a short time or in the case of incomplete data receiving. There are three basic principles of error concealment techniques whether the algorithms are oriented on the encoder, decoder or both sides, namely, FECM (Forward Error Concealment Methods), PECM (Post-processing Error Concealment Methods) and finally IECM (Interactive Error Concealment Methods).

The outline of the paper is as follows. In the next three sections, an overview about FECM, PECM and IECM are introduced. The video evaluation metrics for the video quality measurement by objective and subjective criteria are presented in section no. 5. Finally, the brief summary about present error concealment techniques and future tasks are discussed in conclusion.

2. Forward Error Concealment Methods

To the class of FECM belong these error concealment algorithms where the encoder or transmitting side of communication system play primary role. The error suppression is carry out by source, channel, layered or transport coding. Moreover, in these algorithms the ECC (Error Control Coding) and ARQ (Automatic Retransmission Request) as part of FEC (Forward Error Correction) is widely used. On the other hand, these algorithms are not using the error concealment techniques at the decoder side. There are a lot of basic principles of FECM, like transport level coding, entropy or waveform coding, multiple description coding, joint source and channel coding and layered coding with prioritized transport (mainly oriented on MPEG-2 transport streams).

FECM techniques are developed for suppression of two kinds of distortion observed at the decoder side, namely, the quantization noise and the transmission errors. Most popular technique from previous mentioned is joint source and channel coding, where the source coder is producing the layered stream assuming that the channel coder will guarantee the delivery of the most important source layer [1]. Other most popular technique is FEC coding that is used not for error correction only but for error detection too. Besides, FEC is increasing the amount of transmission data what is its disadvantage mainly for video transmission via narrowband communication channels (shown in Fig. 1).

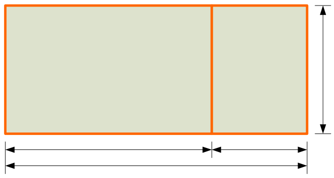


Fig. 1. An example of FEC structure.

3. Post-processing Error Concealment Methods

The PECM are strictly oriented on the error correction at the decoder side and by the disposition of the transmitted data they can be divided into generations. First one are heuristic based algorithms and second one context based modeling.

3.1. First PECM generation

The first generation of PECM are developed for reconstructing the lost video data based on heuristic approaches, which assume smoothness or continuity of the video data in different domains as spatial, spectral and temporal domain. They are also called heuristic methods consist of two categories: spatial/spectral and temporal. Lost image information by the spatial/spectral error concealment methods are reconstructed by interpolating from neighboring pixels arranged in spatial, temporal or spatial-temporal interpolation windows specially oriented on MPEG-2 and MJPEG streams. Moreover, using the temporal error concealment methods, lost data are computed by motion vector estimation. There are other methods based on POCS (Projection Onto Convex Sets), Hough transformation, MRF (Markov Random Fields), BMA (Boundary Matching Algorithm), MFI (Motion Field Interpolation), Hybrid Spatial/Temporal algorithm, etc. can be included to the first PECM generation

3.2. Second PECM generation

The algorithms of second generation of PECM were developed after the knowledge achieved by transmission the content-based or object-based video data like MPEG-4 streams. In this approaches, the video stream use already contains ROI (Region Of Interest) or ORI (Object Of Interest) information. PCA (Principal Component Analysis) or MPC (Mixture of Principal Components) has been widely used to model objects or ROI. The principle of PCA modeling is shown in Fig. 2, where the lost DCT (Discrete Cosine Transform) macroblocks of object-face are reconstructed by the same face parts acquire from other face models.

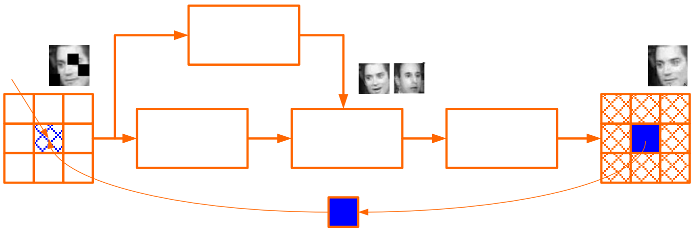


Fig. 2. The adaptive POCS error concealment.

4. Interactive Error Concealment Methods

The IECM are based on the both side interaction between the encoder and decoder. Previous mentioned FECM and PECM techniques have no created backward channel, thus they are not available to specify a level or nature of the video errors, to predict the changes in transmission channel facilities or to adapt the transmission to the error concealment needs. Thanks to backward channel, the coding parameters at the coder side can be adapted by feedback from the decoder side. Moreover, the same information can be used for change the useful video data to FEC data ratio too. There are a lot of IECM algorithms, where the principles based on the multicopy retransmission, selective encoding, retransmission without waiting, adaptive source coding, prioritized retransmission, etc.

5. Video evaluation criteria

From the point of view of measurement of video quality, there are two major approaches, namely, subjective and objective ones. In general, the correct evaluation of visual information can be offered by subjective criteria only. Being directly based on human eyes perception, they are more expensive. On the other hand, the objective criteria like SSIM or VQM try to substitute the subjective approaches. Moreover, they are widely used worldwide as they offer results closer to the response of Human Visual System (HVS).

5.1. Subjective criteria

There are a lot of subjective methodologies standardized by ITU-R BT.500 [6] and ITU-T P.910. Among these criteria have one for example: Double Stimulus Impairment Scale (DSIS), Double Stimulus Continuous Quality Scale (DSCQS), Single Stimulus Continuous Evaluation (SSCQE), and Simultaneous Double Stimulus for Continuous Evaluation (SDSCE) or Subjective Assessment Methodology for Video Quality (SAMVIQ) methodology. Following the reasons given above, the subjective criteria will be replaced by the highly correlated objective criteria.

5.2. Objective criteria

The main problem of the objective approaches regarding correct evaluation of the visual information is modeling the human eye, which is a difficult task. The objective criteria are expected to unify and repeatability of the evaluation process. These criteria can by divided into two classes.

In the first class have one simple criteria like Mean Absolute Error (MAE), Mean Absolute Error Reduction (MAER), Mean Square Error (MSE), Noise Reduction (NR), Signal-to-Noise Reduction (SNR), Color Difference (CD), Motion Sum of Absolute Differences (MSAD), PSNR, Blurring measure, Blocking measure, etc., which are focused on a special image parameter. For example, MAE measures preservation of image details and edges, MSE measures suppression of image artifacts, CD measures changes in image chromaticity, etc. As to color image quality evaluation with the objective criteria closer to HVS, a combination of more than one image parameter like the image contrast, luminance, color, structure, texture etc. must be included.

In the second class have one enhanced criteria like New Quality Index (NQI), VQM or SSIM. They combine more than one image parameter enabling them to be closer to response of HVS than the first criteria category. At the present time, the second class criteria are preferred.

6. Conclusion

This paper offer an overview of available error concealment methods and evaluation criteria oriented on the processing of damaged or lost video data transferred over the error prone communication channels are introduced. A various error concealment algorithms are discussed and classified into three main categories of error concealment methods, namely, FECM, PECM and IECM. Present developing algorithms are focused on the hybrid PECM architecture joining the spatial and temporal characteristics of video stream. Besides, the expansion of HDTV broadcasting in MPEG-4 standard brings increased using the error concealment techniques based on the modeling the video objects. Hence, the hybrid second generation of PECM algorithms could be dominant.

Acknowledgement

This paper has been supported by the VEGA grant agency no. 1/4066/07, “New systems and principles of semantic description and retrieval of multimedia content”.

References

[1] WANG, Y., ZHU, Q.-F., *Control and Concealment for Video Communication: A Review.* PROCEEDINGS OF THE IEEE, VOL. 86, NO. 5, PP.974-997 , MAY 1998, ISSN: 0018-9219

[2] HSU, CH.-T., CHEN, M.-J., LIAO, W.-W., LO, S.-Y., *High-Performance Spatial and Temporal Error-Concealment Algorithms for Block-Based Video Coding Techniques.* ETRI Journal, Vol. 27, No. 1, pp. 53-63, February 2005

[3] CHEN, T. P.-CH., CHEN, T. *Second-generation Error Concealment for Video Transport over Error Phone Channels*. Proceedings on International Conference on Image Processing 2002. Volume: 1, pp. I-25- I-28 vol.1, ISSN: 1522-4880, ISBN: 0-7803-7622-6

[4] FURHT, B. MARQUES, O. *The Handbook of Video Databases: Design and Applications*, Austin, Texas, USA: CRC Press, 2003, ISBN-13 978-0849370069

[5] MARCHEVSKY, S. MOCHNAC, J. *Hybrid concealment mechanism*, Acta Electrotechnica et Informatica Vol. 8, No. 1, 2008, 11–15, ISSN 1335-8243

[6] ITU-R RECOMMENDATION BT.500-11, *Methodology for the subjective assessment of the quality of television pictures*, Geneva, 2002.

[7] WANG, Z., BOVIK, A.C., SHEIKH, H.R., SIMONCELLI, E.P. *Image quality assessment: From error visibility to structural similarity*, In IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600-612, Apr. 2004

[8] HADAR, O., HUBER, M. HUBER, R., GREENBERG, S. *New Hybrid Error Concealment for Digital Compressed Video*, EURASIP Journal on Applied Signal Processing Volume 2005 (2005), Issue 12, Pages 1821-1833

[9] XIAO, F. *DCT-based Video Quality Evaluation*, MSU Graphics and Media Lab (Video Group), 2000

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Automatic Audio Signal Classification. A Survey on Automatic Audio Content Classification

*Michal Chmulík
*University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia, michal.chmulik@fel.uniza.sk

Abstract. In this paper is presented the topic of general audio signal classification and its abilities. The sound classification chain and its elements, audio features that can be used for classification process and classification methods itself including their advantages and disadvantages, is discussed. The previous research related to the audio signal classification and its results is also discussed.

Keywords: content-based audio description, audio features, sound classification and segmentation, MPEG-7

1. Introduction

In the last decade, the mass explosion of digital audio happened. Mp3 audio players, CDs, DVDs, internet and many more became part of our lives. With increasing amount of data requirements to index and manage this data in databases also increases. Various new applications and services became possible with the arrival of a content-based audio description, for example applications allowing user to search, identify, filter and browse audio data in databases. The text-based description of the audio used in the past has significant drawbacks: in addition to being very laborious in large volume data it is not enough rich to describe data for some application. The content-based description represents audio data by its content (spectral energy distribution, fundamental frequency, harmonic ratio, spectral flux…) rather than text. Efficiency of the description (feature vector or vectors) used for comparison and classification depends greatly on the particular application, extraction process and the richness of the description itself.

2. General sound classification

Typical tasks of audio analysis exploiting sound similarity and sound classification approach are:

- segmentation of audio into basic elements - speech, music, sound or silent segment…
- segmentation of speech into segments with speakers of different gender, age, identity…
- identification of speakers and sound events - explosion, applause, laugh…
- classification of music into genres - rock, pop, classic…
- classification of musical instruments into classes - percussive, string, brass…

For the cases of sound similarity, spectral features of sound which allow to distinguish one sound from another, are very useful. Majority of sound classification systems has a structure like depicted in Fig. 1 [1]. Segmentation process of the input audio signal may include separation of input audio signal into basic elements but also isolation of relevant sound segments from the background (from instance instrument from background noise or other sounds). Feature extraction stage obtains from segmented audio meaningful information about its content and selection of used features depends on the particular application. A dimensionality reduction of the feature vector (i.e. mapping of the feature vector onto another vector with lower dimension) is often performed after this stage. The classification stage performs separation itself into classes according to the chosen application (for example instrument classification). The sound classifiers are often based on statistical models. We emphasize that the choice of the feature vector and the choice of the classifier are critical in the sound classification design.

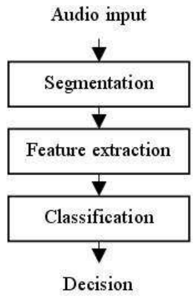


Fig. 1. General sound classification system

3. Audio features and dimensionality reduction

Audio classification systems generally use 3 categories of features that describe the content of audio signal - temporal, spectral and statistical [2]. One of the most commonly used audio features is MFCC. The MFCC (Mel-frequency cepstrum coefficients) is conventional feature vector for the representation of human voice and partially musical signals because it is a compact representation of the spectral envelope of audio signals which respects the nonlinear human perception of pitch. Therefore it is the feature very often used in speech and speaker recognition tasks and it is also able to analyze and represent the musical signals. In [3] MFCCs are used for music genre classification. Other interesting use of the MFCCs is in gender recognition [4] and speaker age and sex identification [5]. Along MFCC, there is variety of the time-frequency features that can be used for sound classification, which have been introduced in literature - loudness, pitch frequency, brightness, total spectrum power, bandwidth… [13], [14], [15] - so an effort to select the most significant ones (and/or to define new useful features) has led to the acceptance of standard MPEG-7 (that includes audio as well as visual content description tools and schemes).

The MPEG-7 audio feature are a set of time-frequency parameters standardized by MPEG ISO group, and covers 17 low-level features, which can describe and distinguish any type of sound hence they are intended for an application in various sound classification tasks. The choice of particular features (e.g. audio power, audio harmonicity, audio spectrum envelope…) depends on the application. In [6] Casey introduces a using MPEG-7 standard for an automatic classification of environmental sounds, musical instruments, music genre and human speakers. He argued the classification accuracy of the environmental sounds and music genre more than 90 % (80 in the worst case) that is comparable with the accuracy of the systems using different time-frequency features. Features and tools of MPEG-7 standard seem to be promising for future work. A comparison of MFCC coefficients and Audio Spectrum Projection of MPEG-7 is presented in [1]. In the past, there were also published approaches using some other features (for example spectral flux for the separation of the speech from music [7], LPC [8]).

After the feature extraction and selection has been performed it is often very practice to reduce the size of feature vectors that way to retain maximal amount of important perceptual information. This will reduce the computational demands and hence decrease the time of computation. Typical way how to do this is using one of following techniques: Singular Value Decomposition (SVD), Principal Component Analysis (PCA), Independent Component Analysis (ICA) or Non-Negative Factorization (NMF) [1], [6], [9]. Other methods that could be used for optimization of the feature vector dimension follow evolution-inspired optimization approaches such as Genetic Algorithms (GA) [13] or Simulated Annealing (SA).

4. Classification

For the purposes of the sound classification have been most often used the model-based classifiers that include:

- Gaussian Mixture Model (GMM)
- Hidden Markov Model (HMM)
- Neural Network (NN)
- Support Vector Machine (SVM)

GMMs have been widely used in the field of speech processing, mostly for speech recognition, speaker identification and voice conversion [1]. GMM models are very popular because of their capability to model arbitrary probability densities and to represent general spectral features. The greatest advantage of GMMs is their computational inexpensiveness. The disadvantage is that only low-level information about the sound is conveyed in the temporal audio signal. Applying of GMMs is presented in [11]. HMM is a statistic method widely used in the field of pattern and time-series classification - originally for speech recognition and speaker verification. HMMs can model processes with the time varying characteristics and after training it is possible to analyze the models more closely. The advantages of HMMs include efficient applying in the large systems, easy integration of multiple knowledge sources and capabilities of effective sound similarity. The most significant drawback is a poor discrimination (only correct models receive the training information). Applying of HMMs is presented in [2], [6], [9]. Various NNs, such as Kohonen self-organizing maps (SOM), multi-layer perceptron (MLP), time-delayed neural network (TDNN) and hidden control neural network (HCNN) have been used in the field of speech recognition and sound classification (in the meaning of segmentation into basic elements). The advantages of NNs are excellent performance and few parameters. The disadvantages are slow training procedure and the fact when a new sound class is added to the classification system the network must be retrained. Applying of NNs is presented in [2], [13]. SVMs have been used in the tasks such as speaker identification, vowel classification and have been also introduced for the classification and segmentation of the audio stream and clips. SVMs have the advantages such as unique solution for given classification, insensitivity to small changes of the parameters and often improved performance in the comparison with other classification algorithms. Applying of SVMs is presented in [11], [14], [15].

Beside these techniques there are more possibilities how to perform the classification. K-nearest neighbor (k-NN) classifier was applied in [2] for a singer detection and a string instrument detection and in [12] for a segmentation into basic elements. [10] is an example where authors used linear discriminant analysis (LDA) for music genre classification or in [8] is proposed using Bayesian classifier for the segmentation into basic elements.

Accuracy of individual classifiers differs from case to case according to a particular application and features but it generally reaches the values of 90 % and more. GMM and SVM classifier is compared for the same classification tasks in [11] (GMM is better) and a comparison of k-NN, HMM and artificial NN classifier is presented in [2] (NN is the best).

5. Conclusion

In this paper is shortly described the topic of audio signal classification, its capabilities, audio features and different types of classifiers that can be used in a process of general sound classification. As is written, for the proposal of good working classifier are the choice of audio features vector and the choice of classifier the most important. In a work previously published, audio features have been selected (or newly designed) by the authors "a priori" and mostly intuitively with accordance to the desired application i.e. they have chosen features that they considered to be the most significant for the given application. In my future work I aim to substitute man by machine so the process of feature selection will be managed by a computer. Next the feature vector dimension will be optimized by genetic algorithms and the audio signal classification itself will be based on new types of supervised learning algorithms. My primal scope of interest will be the segmentation of the audio signal into basic semantically significant elements and I will concentrate on the set of audio features of MPEG-7.

References

[1] Kim H-G., Moreau N., Sikora T. *MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval*. Wiley, 2005

[2] Mingsain R. Bai, Meng-Chun Chen. *Intelligent Preprocessing and Classification of Audio Signals*. Journal of Audio Engineering Society, Vol. 55, May 2007

[3] Lee, Chang-Hsing; Shih, Jau-Ling; Yu, Kun-Ming; Lin, Hwai-San; Wei, Ming-Hui. *Fusion of Static and Transitional Information of Cepstral and Spectral Features for Music Genre Classification*. Asia-Pacific Services Computing Conference, 2008, IEEE, Dec 2008

[4] Ting, H.; Yingchun, Y.; Zhaohui, W. *Combining MFCC and Pitch to Enhance the Performance of the Gender Recognition*. 8th International Conference on Signal Processing, 2006

[5] Naini, A.S.; Homayounpour, M.M. *Speaker age interval and sex identification based on Jitters, Shimmers and Mean MFCC using supervised and unsupervised discriminative classification methods*. 8th International Conference on Signal Processing, IEEE, 2006

[6] Casey M. *General Sound Classification and Similarity in MPEG-7*. Journal of Organized Sound, Cambridge, Vol.6, Aug 2001

[7] Burred J. J., Lerch A. *Hierarchical Automatic Audio Signal Classification*. Journal of Audio Engineering Society, Vol. 52, No. 7/8, July/August 2004

[8] Li D., Sethi I.K., Dimitrova N., McGee T. *Classification of General Audio Data for Content-based Retrieval*. Pattern Recognition Letters, Elsevier, Vol. 22, 2001

[9] Kim H-G., Moreau N., Sikora T. *Audio Classification Based on MPEG-7 Spectral Basis Representations*. Transactions on Circuits and Systems for Video Technology, IEEE, Vol. 14, No. 5, May 2004

[10] Umapathy K., Krishnan S., Jimaa S. *Multigroup Classification of Audio Signals Using Time–Frequency Parameters*. Transactions on Multimedia, IEEE, Vol. 7, No. 2, April 2005

[11] Arias J.A., Pinquier J., André-Obrecht R. *Evaluation of Classification Techniques for Audio Indexing*. Proc. European Signal Processing Conference, 2005

[12] Lu L., Zhang H.J., Jiang H. *Content Analysis for Audio Classification and Segmentation.* Transactions on Speech and Audio Processing, Vol. 10, No. 7, Oct 2002

[13] Xi Shao, Changsheng Xu, Kankanhalli M. S. *Applying Neural Network on the Content-Based Audio Classification*. Proc. of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, Dec 2003

[14] Guodong Guo, Stan Z. Li. *Content-Based Audio Classification and Retrieval by Support Vector Machines*. Transaction on Neural Networks, IEEE, Vol. 14, No. 1, Jan 2003

[15] Bai Liang, Hu Yanli, Lao Songyang, Chen Jianyun, Wu Lingda. *Feature Analysis and Extraction for Audio Automatic Classification*. Systems, IEEE International Conference on Man and Cybernetics, Oct 2005

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Fuzzy Optimization Models for Online Advertising Control Systems

[*] Jaroslaw Jankowski
[*] West Pomeranian University of Technology, Computer Science Faculty, Information Systems Institute, ul. Zolnierska 49, 71-210 Szczecin, Poland, {jjankowski}@wi.ps.pl

Abstract. Electronic marketing and advertisement emission and control systems require adequate representation of input data, processing algorithms and decision support. This article presents the applications of fuzzy and multi-objective optimization with uncertain parameters in media planning, which helps to build a more realistic representation of audience characteristics in a dynamic interactive environment, as well as improving efficiency and resource usage.

Keywords: multi-objective optimization, Internet, interactive marketing, fuzzy modeling.

1. Introduction

The increasing importance of interactive technologies has been brought about by the development of a new media and is affected by information technology. Kotler Ph. and Postma P. emphasize that marketing is entering a new era with a growing number of applications and new interactive media within which to operate [6]. Interactive media allow the development of new forms of communication with applications in advertising operations defined by Hoffman D. and Novak T. [5]. Traditional mass media use a *one to many* communication model, where it is impossible to analyze direct effects and influence on customer behavior [11]. The main distinguishing feature of electronic media (and the Internet) is two-way communication, which was the main reason for creating a new model. The *one to one* model allows information to be sent and feedback to be received. In this model sender N_i prepares message k_j addressed to receiver O_n and sends it to coding/decoding unit DC_i, where the message is translated into the form k^*j that is required by the medium (Fig. 1).

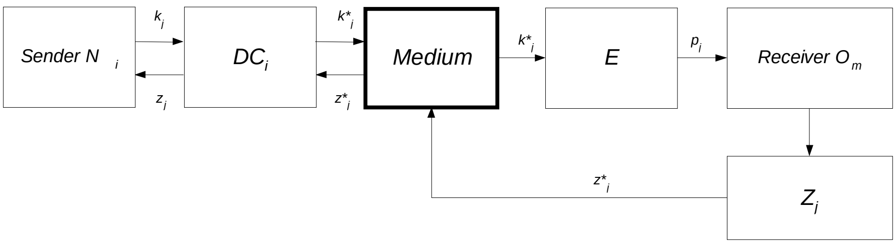


Fig. 1. Interactive communication model

In the next step the message is integrated with an advertising medium and, as a result of the emission conducted by unit E, personalized message p_j is received by its target O_m. The receiver has the option of generating feedback information Z_j, which after decoding is received as z_j. Control systems responsible for emission planning and request services contain dedicated modules responsible for different actions and require optimizations at operational layers that can be developed towards real time optimizations [12]. This article presents fuzzy approach to interactive media planning and decision support models for resources optimization.

2. Audience representations in uncertain environments

Decision support in online advertising emission systems has an impact on results and effectiveness. Most of the decision support models presented use deterministic parameters based on average values and single optimization goals [3]. In control systems for advertising networks uncertainty arises from changing effectiveness and changing audiences [13]. On the basis of historical data we can estimate that in the next time period *t+1* “about” *n* advertising units may be emitted, which is a more real statement than the precise statement that n units will be emitted with representation based on fuzzy numbers defined by Zadeh L.A. [14]. Example data representing audience measurement where calculated from online data [7] using time series addressed methods proposed in [4]. Defining resources for the next period *t+1* can be based on average values and deterministic approaches, so that parameters like page views (PV=473107) and unique visitors (UU=36886) do not use characteristics for a changing environment. Fuzzy representations of input data for *PV* (page views) , *UU* (unique users) and emission costs *EC* that are presented in Fig 2 - 4 can be used as parameters for decision support models during media planning.

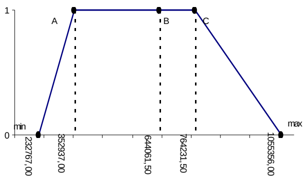


Fig. 2. Emission resources *PV*

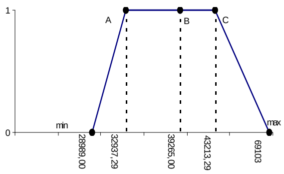


Fig. 3. Unique users *UU*

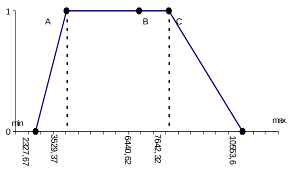


Fig. 4. Activity costs *EC*

The number of contacts with advertising content can be used for evaluations of effectiveness and user engagements [9]. The decedents approach will affect the final values to be used for planning purposes, for example the optimistic approach can be used for the period *t+1* values UU=49591, and the pessimistic approach UU=28989. The number of users for servicing publishers can be defined as fuzzy numbers with linguistic representations. Several methods can be used for membership function creations [10] and an example is presented in Fig. 5. A similar approach can be used with effectiveness factors like conversion ratio or audience transfer indicators. In some cases, i.e. pulsing emissions [1], a fuzzy campaign budget can be allocated and that representation makes such parameters more reliable than categorical values.

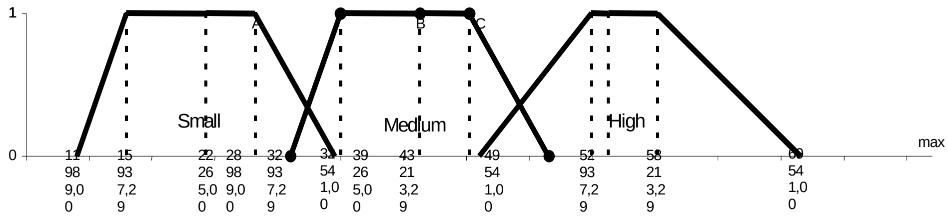


Fig. 5. Membership function for the linguistic term „number of users"

There is an unknown value of emission plan $x_j(i,g,m,d)$, which is the defining number of units of type d to be emitted by advertiser j in target group g, publishing service i for the period time t. For the deterministic version of that model method DIDAS-L was used [8].

3. Emission planning and optimization in advertising control systems

Resources usage and allocation model can be defined as multi criteria decision support model with fuzzy parameters. Maximization of function (1) increases the overall financial performance of emission balance and the number of interactions is important for effectiveness and aggregated results, as described by function (1):

$$f_g^t(x) = \sum_{Z \in \Omega} \left(x_j^{t+1}(i,g,m,d) * \left(k_j^{t+1}(i,g,m,d) - z_j^{t+1}(i,g,m,d) \right) \right) \rightarrow \max \tag{1}$$

where: x_j – number of advertising units emissions, z_i cost of emission ad unit $p(j,m,d)$ for target group g using advertising space in publisher i paid for by the emission system management organization, k_i - cost of emission of ad unit $p(j,m,d)$ for target group g using advertising space in publisher i paid for by an advertiser, E_{CTR} - fuzzy effectiveness factor representing the number of interactions (conversion) for advertiser j in target group g and publisher i. Other defined functions $f_1(x)$, $f_2(x)$, $f_3(x)$ represent financial results, aggregated results of publishers and aggregated advertiser results. Solutions were generated using the ARBAN method proposed in [2]. Fig. 6 presents the results for each function with a different dissident approach to risk factors (optimistic, neutral and pessimistic).

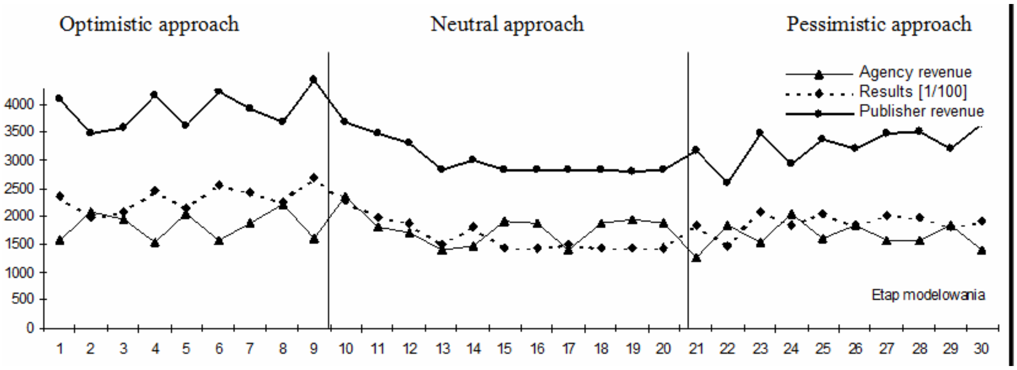


Fig. 6. Decision results modeling using fuzzy parameters and constraints (Source: own research)

The interactive decision generating process was divided into several stages. In stages 1-10 for emission resources and budget constraints realization levels τ where set at different levels from 1 to 0.1 with zero values for parameters $w(\tau)$. For constraints realization at level

1 (step 1) goal functions were above aspiration function values. Solution number 1 is the first compromise solution received with an 'optimistic' approach. In stages 11-20 for *τ=1* different membership factor *w(τ)* was used: in stage 11 it was 0.25, in 12 it was 0.5, and in 13 it was 0.75 for all constraints, representing the 'neutral approach'. At other stages different values for *w(τ)* where set. The solution from step 10 with *τ=0.1* delivers satisfactory levels of fuzzy aspiration $f_1(x)=2369$, $f_2(x) =3678$, $f_3(x)=226964$ and is between the optimistic and neutral level, which makes it more reliable for future usage. For solutions with a low τ characteristic there is a high risk that emission resources will not be available. For level 12 it is lower for each function where calculated $(fg_1(x)=1707, fg_2(x)=3293, fg_3(x)=186954)$, but factors *τ=1* and w(τ)=0.5 for resource constraints and *τ=1* and *w(τ)=0.5* for budgets causes a higher level of expected constraints fulfillment. Using this method and experimental data it was possible to increase the number of interactions and resources usage and achieve new approaches for electronic media planning.

4. Conclusion

The complex nature of interactive online stems requires an interdisciplinary approach which integrates various academic disciplines e.g. psychology, economics and computer science. In electronic marketing an important role is played by methods of data analysis used as a methodological basis for advertising servers and methods of extracting knowledge from data bases designed for use in the Internet (eg. web mining). Along with the development of marketing techniques there is a corresponding development in the technology used in their application, whilst new problems and areas of research continue to emerge. The optimization technique presented in this paper is an example of the use of a fuzzy model in the planning of an advertising emission. One of these new areas is the standardization and creation of techniques for the personalization of advertising messages and behavioral approaches.

References

[1] Beliczynski J. Planowanie mediów w zarządzaniu reklamą, Antykwa, Kraków, 1999.

[2] Banas J. Adaptacja algorytmów wielokryterialnych do optymalizacji zadań rozmytych, Praca doktorska, Wydział Informatyki Politechniki Szczecińskiej, 2002.

[3] Chickering D., Heckerman D. Targeted advertising with inventory management, ACM Special Interest Group on E-Commerce (EC00), Minneapolis, 2000.

[4] Frantti T. Timing of fuzzy membership funcions from data, University of Oulu, 2001.

[5] Hoffman D.L.: Novak T.P. Marketing in Hipermedia Computer-Mediated Environment: Conceptual Foundations, Vanderbilt University, Nashville, 1995.

[6] Kotler Ph., Postma P. The New Marketing Era, Marketing to the Imagination in a Technology World, McGraw-Hill, New York, 1999.

[7] Military Times Media Group, Media Kit, militarycyty.com. 2005.

[8] Makowski M., Wierzbicki A.P., J. Wessels J. Model-based Decision Support Methodology with Environmental Applications, Kluwer, Dordrecht, 2000.

[9] Nowak W. Media Planning, MDW, Kraków, 2001.

[10] Piegat A. Modelowanie i sterowanie rozmyte, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 1999.

[11] Rust T. Advertising Media Models, Lexington Books, 1989.

[12] Wadhwa H. Real-time Optimization, Avantmarketer.com, 2002.

[13] Wong J., Wu G., Leckenby J. Publication Interval and Web Audience Estimation, The University of Texas, Working paper, 1999.

[14] Zadeh L.A. Fuzzy sets, Information and Control 8, 1965.

Multidimensional Adaptive Modeling of Interactive Objects in Web Applications

[*] Jaroslaw Jankowski
[*] West Pomeranian University of Technology, Computer Science Faculty, Information Systems Institute, ul. Zolnierska 49, 71-210 Szczecin, Poland, {jjankowski}@wi.ps.pl

Abstract. In the article a procedure is presented of adapting objects used in the interactive communication process. It assumes object decomposition and identification of the system reaction functions for different selected collections of input variables with simultaneous minimization of the number of variants.

Keywords: Internet, web applications, usability, data modeling.

1. Introduction

The development of information systems and internet applications takes place in the technological and social aspect that more and more often constitutes the basis for many projects. The development of broadband Internet has made it possible to introduce new areas of applications, the quantity of multimedia content has grown and decentralization of publishing functions has occurred. In many applications focused on interactive environment, there are problems with creation of data access interfaces and structure modeling to ensure usability and generate specific effects. In the article the concept is presented of decomposition of interactive objects into component elements under procedure of searching for extreme values of the response function, which identifies system characteristics and gives a possibility to maximize expected effects in the adaptive process. Proposed method is based on adaptive parameters selection with quantitative and qualitative input sets.

2. Evolution of data structures in web applications

Along with the development of websites focused on online communities, new structures of internet applications and different approaches for creating and running them are used, i.a. in the form of ontology and semantic networks whose directions for development are indicated by J. Davies and R. Studer [6]. This phase in Internet evolution integrates elements: creation of content by users, collective intelligence, data repositories of so far unprecedented scale, social networks, scale effect, openness [1]. As are indicated by M. Madden and S. Fox the present stage of development is a result of natural evolution taking place by way of verification of various concepts known before [9]. These studies identify related areas (e.g.: folksonomies [10], studies of social networks, mashup type web aggregators) in the period of the greatest interest. Integrated are methods of building structures of interfaces oriented on exploring content of electronic commerce systems in the form of collaborative filtration systems whose integration unified methods were presented by J. Wang and P. Arjen [13]. For different applications one can identify typical data structures used in communication with the user. Measures of their usability, broadly discussed by A. Granica and V. Glavinica [8], enable identification of objects being a mediating layer between the user and the application part, and are subject to modification and adjustment in time. S. DeLoach and E Matson specify more accurately for the Internet environment adaptive and agent approaches applied in design of IT systems and interfaces where, as the ideal solution, regarded is implementation of a system that adjusts to the changing conditions and user needs [5]. Quantitative parameters $C_1, C_2,..., C_m$ represent measurable characteristics of the object and its constituent parts which include parameters of graphic elements and are represented by continuous or discrete variables. Discrete variables are included in the form of collections of alternative values $ZD = \{e_1, e_2..., e_n\}$. When generating a version of interactive communication, selection is made of the elements from the given set, based on the given selection function $f_s(x, z, n)$, responsible for selection of element x from set z to cause communication n, which is intended to maximize effects (for example advertising campaigns) or object usability. Selection of decisions support methods and algorithmization of the data analysis process are important for effective attainment of the objectives of calculation and generation of structures.

3. Parameterization of interactive objects

Implementation of an adaptive system for generating interactive objects requires selection of analytic data processing methods and acquisition of decision-making solutions. Literature provides models of generating the so-called fractional plans (with the use of gradual increment in the value of particular parameters) deemed as the basis for this research stream published i.a. in the works by C. Bayen and I. Rubin [2], D. Montgomery [12] and S. Deming and S. Morgan [6]. The development of methods in this area was initiated by reaction surface methodology RSM suggested by Box and Wilson [14], which encouraged experiment planning to set the extreme of a function of many variables: $x= f(u_1,u_2,\dots,u_n)$. In this method it is assumed that function $f(u)$ is unknown, continuous and has one extreme. Variables u_k $(k=1,2,\dots,s)$ can accept values on two main levels $u_{s0}+\Delta u_s$, $u_{s0}-\Delta u_s$. The method assumes that the unknown characteristic can be approximated in the neighborhood of a given point $u_{1,0}, u_{2,0}, \dots, u_{s,0}$ with the use of hyper plane presented by means of equation (1).

$$\hat{x} = b_0 + b_1 u_1 + \cdots + b_s u_s + b_{11} u_1^2 + \ldots + b_{ss} u_s^2 + + b_{12} u_1 u_2 + \ldots + b_{s-1,s} u_{s-1} u_s \qquad (1)$$

where $b_0, b_1, \dots, b_k, b_{11}, \dots, b_{ss}, b_{12}, \dots, b_{s-1,s}$ are coefficients that need to be determined. In the case of presence of a large number of variables experiment realization plans are sought that will permit examination of the main factors determining function behavior. In such situations Plackett and Burman plans shall apply, defined as saturated due to the use of all elements for identification the main effects without degrees of freedom left [7].

4. Experimental results with adaptive objects

Examples of the procedure based on fractional plans are illustrated for object with six input parameters having continuous values, representing components of variables responsible for visualization and selection of the color scale of an interactive element. For different parameters acceptable ranges of values have been defined from RGB scale. In the first phase initial plan selection method was used from Placket and Burman experiment the experiment system was generated. Initial central point was defined for value $c_{1,1}$ = 127.5. Incremental matrix for each variable determines the direction of changes in different steps of the experiment. Based on this in the neighborhood of the central point the initial experiment plan is generated. The experiment plan is presented by hyper plane on Fig. 1. For different input parameters evaluation was made of the object structure and ranks were obtained for different variants of the plan. Fig. 2. presents the effect of changes in value of parameter $c_{1,1}$ to response of the system at values $c_{1,3}$ = 0, $c_{2,1}$ = 178. Analysis of the object takes place in iterative process in which data from measuring systems and initial values of the parameters from input systems are used. The function extreme value is determined iteratively in several steps. In the first one a small number of experiments are performed that will find a locally orientated mathematical description that determines the surface given by the following equation:

$$\hat{x} = b_0 + b_1 u_1 + b_2 u_2 + \cdots + b_s u_s \tag{2}$$

Coefficients $b_1, \ldots, b_s$ define the direction of gradient that permits searches of the extreme point. The method of searching for the direction of the fastest possible function growth (method of steepest ascent) assumes movement of the direction of experiment along the route of object reaction function value growth.

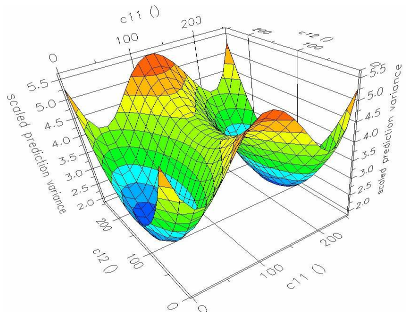


Fig. 1. Starting plan for experimental stages (Source: own calculations)

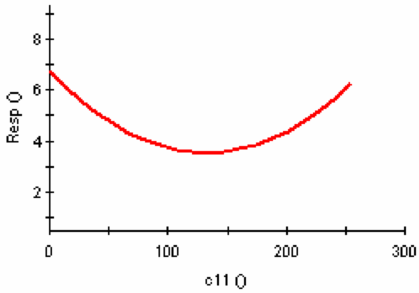


Fig. 2. Response allocation for $c_{1,1}$ $c_{1,3}$=0, c_{21}=178 (Source: own calculations)

After the first processing stage, model regression coefficients are used to identify the path of searches. Changes in value x_j when moving along with function growth are proportional to significance of the regression coefficient. In the work [3] the model is presented in the form of:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k \tag{3}$$

Fastest possible function growth path is running from the central point of the project in the direction of the greatest function value growth with spherical limitation. Function maximization procedure uses Lagrange multiplier:

$$Q(x_1, x_1, \ldots, x_k) = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_j - \lambda\left(\sum_{j=1}^{k} x_j^2 - r^2\right) \tag{4}$$

where for multiplier λ stepwise Δ parameter change values are defined and values of other variables are determined. Detailed methods of fixing the solutions are discussed in the works by R. H. Myers and D. C Montgomery [12] and A. I. Khuri and J.A.Cornell [11]. The obtained effects indicate a possibility for broader application of planning methods for structures of interactive objects and their versification involving fractional plans. Depending on the set target it is possible to identity the system reaction function and look for combinations of input parameters in order to maximize it.

5. Summary

Growth in the complexity of internet applications, variability of data structures, use of multimedia objects make it difficult to model information structures and design access interfaces. The application potential is not always used and the obtained results maximized. The presented procedure of identifying an interactive object and selecting parameters may constitute extension of existing adaptive systems in the Internet. The approach used gives a possibility to verify various systems and minimize the space of the searches of the design variant in web applications, where combinatorial nature of a job makes it difficult to use other methods. In many cases it is useful to reflect the system, and a mathematical analytical model is determined on the basis of its behavior. Further stages of research assume construction of a simulation environment and integration with Internet systems for the needs of further experiments.

References

[1] ANDERSON P. What is Web 2.0? Ideas, technologies and implications for education, JISC, Department of Information Science, Loughborough University 2006.

[2] BAYNE C. K., RUBIN, I. B. Practical experimental designs and optimization methods for chemists, VCH Publisher, 1986.

[3] BORROR C., KOWALSKI S., MONTGOMERY D.C. A Modified Path of Steepest As-cent for Split-Plot Experiments, Journal of Quality Technology, Vol. 37, No. 1, 2005.

[4] DAVIES J., STUDER R., WARREN P. Semantic Web Technologies: Trends and Re-search in Ontology-based Systems, Wiley Publishing, 2006.

[5] DELOACH S. A., MATSON E. An Organizational Model for Designing Adaptive Multi-agent Systems, The AAAI-04 Workshop on Agent Organizations: Theory and Practice, 2004.

[6] DEMING S. N., MORGAN, S. L. Experimental design: A chemometric approach, Am-sterdam, The Netherlands: Elsevier Science Publishers B.V., 1993.

[7] GIESBRECHT F., GUMPERTZ P. Planning, Construction, and Statistical Analysis of Comparative xperiments, Wiley Series in Probability and Statistics, 2005.

[8] GRANIC A., GLAVINIC V. Functionality specification for adaptive user interfaces Electrotechnical Conference, Melecon, 2000.

[9] MADDEN M., FOX S. Riding the Waves of Web 2.0, Pew Internet Project, 2006.

[10] STURTZ D. Communal Categorization, The Folksonomy.INFO622, Content Representa-tion, 2004.

[11] KAURI A.I., CORNELL J.A. Response surfaces: Design and analysis, DEkker, 1996, New York.

[12] MONTGOMERY D.C., MYERS R.H. Response surfaces methodology, Willey, New York, 1995.

[13] WANG J., ARJEN P., MARCEL J.T. Unifying User-based and Item-based Collaborative Filtering Approaches by Similarity Fusion, SIGIR, 2006.

[14] WILSON K.B., BOX G.E.P. On the experimental attainment of optimum conditions, Journal of the Rogal Statistical Society, 1951.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Online Education Tools a State of Art

*Adam Jaroš
*University of Žilina, Faculty of Management and Informatics Engineering, Department of Technical Cybernetics, Univerzitna 2, 01026 Žilina, Slovakia, ad@frtk.fri.uniza.sk

Abstract. In last fifteen months there is a lot of news in the area of an online publishing and design of new online tools. Some are a businesslike and some are focused on online presentation and many have a user interactivity support. This are a new trends against an old approaches with only basic interactivity when most of online tools offered only static information like HTML, images, audio and video. In this paper I present a state of art. A practical application of design of tools for education is presented here.

Keywords: Education, Adobe Flash & AIR, Flex Builder, SUN JavaFX, Microsoft Silverlight, AutomatFlash viewer.

1. Introduction

This paper gives an overview of the latest tools for design of the online simulation models which exists on Microsoft Windows platform. All of them exist on Mac and Linux platforms too. I shall not focus on too many details, which can be found on homepages of mentioned tools. As all of them are rapidly developed, we have to make a decision witch tools are convenient for a particular area of our interest.

In face of a wide use of computers and rapid development of hardware and software tools there are only few areas where interactive multimedia for education is used. One of them is tool for design of simulation models from Mathworks Inc. (Matlab and Simulink) and similar tools. Mostly all of them need installation of additional software on a user computer.

Now we look how internet browsers like Microsoft Internet Explorer, Mozilla Firefox and Opera with installed related plug-in can be used as simulation environment.

2. Platform for Interactive Applications from Adobe Company

This platform represents at the present time a major platform for online interactive application development. Adobe Company offers for so many years a stable solutions for online publishing, online interactivity. One of advantage is wide spread formats of their production tools, for example PDF (Adobe Reader), SWF (Adobe Flash), PSD (Adobe Photoshop), AI (Adobe Illustrator) and many more.

There are some important requirements for new and perspective tools for the future. Adobe Company every eighteen months comes with major updates of their all production tools, even at this time come with important updates.

2.1. Adobe Flash & AIR Runtime Browser Plug-in

The Flash Player is the cornerstone of the Adobe Flash & AIR platform [1]; it offers the developer a single, cross-platform runtime with all the capabilities of the Internet today. It supports more than just vector animation—it’s also a multimedia player. In fact, the majority of all online video today is delivered via Flash, with the capacity to offer true High Definition video.

A lot of effort has been invested into making sure that Flash-based RIAs work successfully for the wide variety of devices and operating systems.

Adobe has successfully matured the Flash runtime in the short time that they've had it as a product. With each iteration, particularly Players 9 and 10, significant improvements in performance have occurred; Player 10 now offers hardware acceleration for advanced graphical effects. Flash Player 10 also introduces new functionality, including 3D effects, custom filters and effects, advanced text support, and dynamic streaming for improved video performance. Flash Lite 3 is a scaled-down version of the runtime for use on mobile and handheld devices and soon in High Definitions TV.

And then there's Adobe AIR—a new category of runtime that combines Flash Player 10 and WebKit (the web page rendering engine behind Safari and Chrome) into a stand-alone runtime capable of running offline applications. On Windows, OS X, and Linux.

Developers can use their web application building skills to create apps for the desktop and deploy them via AIR. The AIR runtime incorporates functionality missing from the browser (for security reasons, runtime represents Sandbox—a secure and isolated environment), such as access to the file system to create, delete, and maintain files and folders.

3. Platform for Interactive Applications from Microsoft

This platform represents at the present time a second major platform for online interactive application development. Microsoft came relatively later with their solution—Silverlight [2]. Therefore in less than fifteen months it release three major versions Silverlight 1, 2 and soon this year it will be released version 3. This gives an advantage to Adobe Company. There are more advantages of Microsoft. One of greatest is that in coming new operating system Windows 7, Silverlight will be a part of it. Microsoft has a mature development tools—Microsoft Visual Studio 2008, Microsoft Expression Studio 2 and well established .NET platform tools. There is no major difference between Silverlight and Flash Player plus AIR support. Linux platform has implementation of Silverlight known as Moonlight [3]. There is a possibility that Moonlight version 3 will be released at the same time as Silverlight 3.

Microsoft Silverlight is a cross-browser, cross-platform, and cross-device plug-in for delivering the next generation of .NET based media experiences and rich interactive applications for the Web. Silverlight offers a flexible programming model that supports AJAX, VB, C#, IronPython, and IronRuby, and integrates with existing Web applications. By using Expression Studio and Visual Studio, designers and developers can collaborate more effectively using the skills they have today. Silverlight will support all major browsers on both Mac OS X, Linux and on Windows.

4. Platform for Interactive Applications from SUN Company

SUN JavaFX is a rich client platform for building cross-device applications and content. Designed to enable creation and deployment of rich internet applications (RIAs) and behave consistently across diverse form factors and devices is on **Fig. 1**. JavaFX 1.0 is from all major platforms newest, but it marks the first step in Sun's strategy to enable rich immersive media and content across all screens.

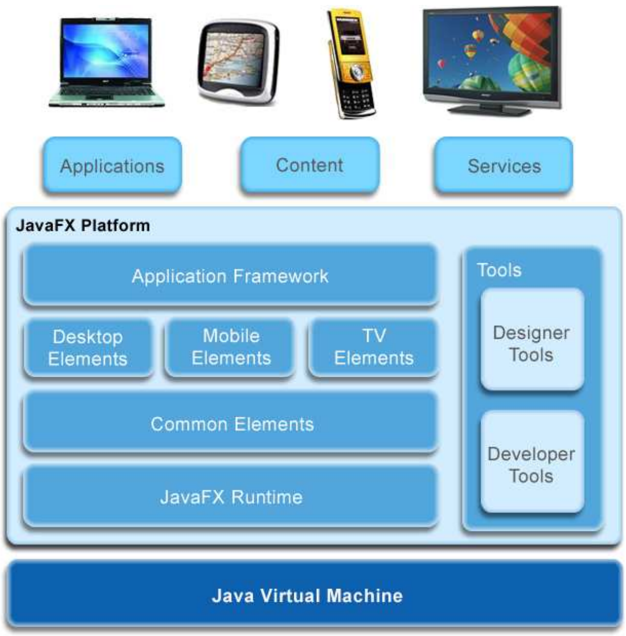


Fig. 1. An overview of the SUN JavaFX 1.0 architecture.

The SUN JavaFX 1.0 platform [4] includes the following components: JavaFX 1.0 SDK which includes the JavaFX compiler and runtime tools, graphics, media, web services, and rich text libraries to create RIAs for the desktop, browser and mobile platforms. NetBeans IDE 6.5 for JavaFX 1.0 which provides a sophisticated integrated development environment for building, previewing, and debugging JavaFX applications.

5. Online Education Tools for Simulation and Control

To design a series of useful tools for education I start with Adobe Flash CS4 and Microsoft Expression Studio 2 SP1. After four months I choose a Flash platform to continue because it is well established. I started with Macromedia Flash 8 Pro, but two months later I switched to Adobe Flash CS4 Pro and latest version of Action Script 3.0. There is a lot of learning tutorials of a first class quality [5], many books [6] and examples. All major operating systems are well supported. Next we have a good perspective in the future because of well established Adobe Company on the market.

I designed and started create a series of an online applications for education of subject Logical systems—presentation of Moore and Mealy automats with build in interactivity and reusability in more complex applications. Designed flash application AutomatFlash Viewer is presented in [7].

Main task is to create simple and reusable flash applications with extensible characteristic usually XML based. The goal is to create a complex e-learning education course for comprehensive study of mentioned subject. With Adobe AIR technology e-learning course can be simply studied online at web, but with a simply drag and drop do desktop on Windows, Linux or Macs we get a full-featured application. Some e-books of a new style already exist even in Czech and Slovak republic. Adobe Flash player & Adobe AIR technology were recognized as powerful platform. Designed flash application AutomatFlash Viewer [8] is used in more complex interactive application as shown on **Fig. 2**.

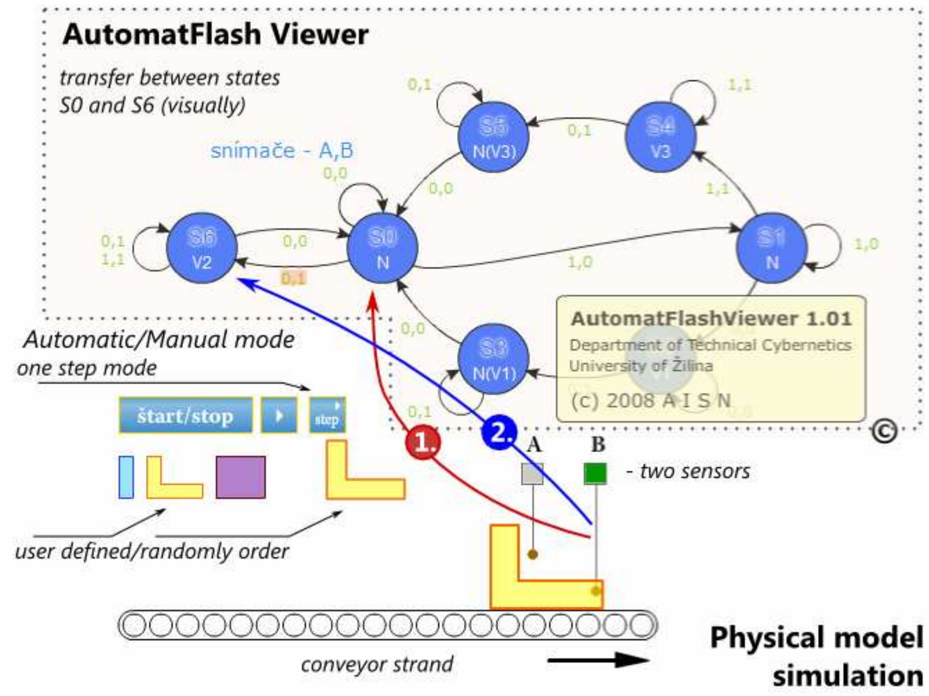


Fig. 2. Interactive flash application for three different products recognition based on AutomatFlash Viewer.

6. Conclusion

Each of three mentioned platform is under heavy development. Each one wants to become a leader on the field of online publishing. Last few years it was Adobe platform, but there are more players in game, Microsoft, SUN, and even Google.

As all hardware is evolving to be more and more powerful, all platforms are growing to a full multimedia platform. We can notice that hardware support for acceleration is growing very fast too, for example hardware decoder for H2.264 codec and massive parallelism.

Acknowledgement

This work has been supported by VEGA 1/0808/08. Self Tuning and Learning Algorithm for Automatic Control.

References

[1] Adobe Company and Adobe Flash & AIR Homepage: *http://www.adobe.com*

[2] Microsoft Silverlight Homepage: *http://www.microsoft.com/silverlight/default.aspx*

[3] Linux Implementation of Silverlight Homepage: *http://www.mono-project.com/Moonlight*

[4] SUN JavaFX 1.0 Platform Homepage: *http://www.sun.com/software/javafx/index.xml*

[5] Lynda.com E-learning Education Company Homepage: *http://www.lynda.com*

[6] KEEFE, M. *Flash and PHP Bible*. Wiley Publishing, Inc., USA, 2008. ISBN: 978-0-470-25824-8.

[7] JAROŠ, A. *New Tools for Design of the Online Simulation Models.* 43rd International Conference. MOSIS'09. Modelling and Simulation of Systems. Rožnov pod Radhoštěm, Czech Republic. April 7 - 9, 2009.

[8] AutomatFlash Viewer Applications Homepage: *http://frtk.fri.utc.sk/~ad/logicke-systemy.html*

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Application of WMS and Pick by Light in Logistics Center

*Rudolf Kampf, **Alexander Chlaň, ***Miroslav Gottfried, ****Martin Dvořák
* University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics, Studentská 95, 532 10 Pardubice, The Czech Republic, kampf@upce.cz
** University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics, Studentská 95, 532 10 Pardubice, The Czech Republic,Alexander.Chlan@upce.cz,
Technical University of Lisboa, Instituto Superior Técnico, Department of Civil Engineering and Architecture, Av. Rovisco Pais, 1049-001 Lisbon, Portugal, alexander.chlan@civil.ist.utl.pt
*** University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics, Studentská 95, 532 10 Pardubice, The Czech Republic
**** University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics, Studentská 95, 532 10 Pardubice, The Czech Republic

Abstract. The article deals with description of major benefits that the WMS used during consolidation of material and finished goods in logistics centre can provide to user. Author focuses also on a tool that helps to minimize mistakes and mismatches during picking operations with the increase of speed of picking and its efficiency at the same time. Application is verified by feasibility study that includes payback time calculation.

Keywords: pick by light, pallet, box, rack, warehouse management system = WMS, logistics unit, MOQ, RFID tag.

1. Introduction

Many logistics factories are just now working on implementation of modern system and technologies focused on warehouse management that brings many benefits like easier and faster warehouse operations, fast warehouse inventory check and information about necessity for next material delivery into warehouse or from the warehouse to production. In such kind of situations warehouseman usually orders MOQ – minimum order qty that was initially presented by pallet qty and now MOQ is usually presented by box qty as “new” logistics unit. Change from pallet to box brings many benefits for example less necessity to own warehouses of big capacity and heavy machinery in, less waste, less handling and less money dropped in material in warehouse as factories order only material that they really need based on demand for their products and its order qty is rounded up to box qty. All risks caused by price drops or material losses are handed over to suppliers.

To achieve this position in supply chain factories have to change many processes and to make change faster and more successful they can use earlier mentioned systems like Warehouse Management System designed by many IT companies and that can be developed exactly according to customers requirements for example with additional systems for faster picking (Pick By Light) and order’s consolidations done without mistakes.

2. Warehouse Management System WMS – Controlled Warehouse

WMS = Warehouse Management System is the complex solution for atomization of processes and material movements in controlled warehouse. WMS is part of the supply chain and primarily aims to control the movement and storage of materials within a warehouse and process the associated transactions, including shipping, receiving, put away and picking and enabling users to deliver superior customer service by improving order accuracy, increasing efficiency, streamlining materials handling, meeting compliance requirements and refining inventory control. The systems also direct and optimize stock put away based on real-time information about the status of bin utilization.

Core of the system is presented by software that controls all operations mentioned above. This software is usually customized based on requirement coming from customers who want to successfully answer on upcoming requests like increase of frequency of supplies, minimization of mistakes and mismatches during consolidation and increase on efficiency and productivity.

Additions to WMS

- **WIFI technology and SmartStock.WMS** – it allows on-line connection of mobile terminals to company information system – it is the key part of paper less system based on information displayed on terminal able to read barcodes or RFID tags and without any necessity to handle papers with notes during operations; all movements in the system are done in the real time by simple scanning of labels with barcodes that are sticked on racking system and all picked material and confirmations are directly uploaded to the inventory system and used for other purposes – planning etc. = core of paperless system. In such kind of system of racks, bins and material must be unambiguously marked and labeled to make identification easier and to allow to all benefits coming from this system to come.

WMS helps in all kind of operations in the whole factory from order receiving, via material preparation, consolidation to expedition as is described on picture 1 below:

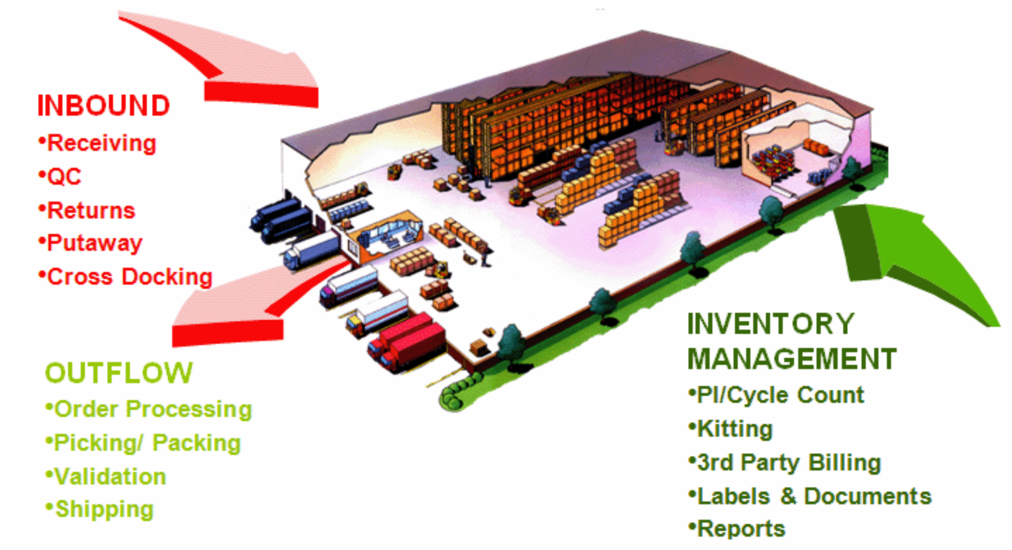


Fig 1: Controlled WH (Source: www.google.com)

By application of WMS in warehouse companies can achieve following goals:

- Easier inventory check and sure material movements within company/warehouses.
- Increase of productivity by automatization and optimization of warehouse operations – providing better services to customers, operation in waves, optimal place utilization
- Decrease on level mistakes and mismatches, minimization of human factor affects.

All these benefits usually help to have better overview on material in WH, to satisfy customers by faster deliveries right on time, of correct qty and at the right place. It also helps to increase the work productivity, decrease transportation costs and increase quality of to customers provided services.

3. Efficient Paperless Consolidation of Mixed Goods

Based on current situation companies have to find an efficient tool for consolidation of shipments from many different materials at one place where an operator is the key-part of the process and his ability to do all operations quickly and safely involves its performance. Usually it is the operator who affects speed, reliability, accuracy and error rate of operations like material preparation, packing and expedition to a customer. Operators are more and more pushed to do all stressful operations faster and faster. The operator usually handles printed version of material lists and goes thru WH from rack to rack, from bin to bin, picking material to the basket directly from boxes or pallets, without any system of online pick confirmation and without any confirmation if the picked material is correct one for example by scanning of PN from label (PN is unique code of material). The operator marks picks to the paper by a pen and when he is sure that he picked all requested stuff the operator handles it to the next team who pack it and send for the expedition. This way of done operation doesn't bring any reliability, causes many mistakes and moreover stock qty and movements are not reported in the real time.

With a view to maximize reliability and minimize the fail rate, many developers started with development of supporting tools connected to the information system to achieve main goals – higher speed, better reliability and low fail rate. The system will be connecting visual tools without necessity to use papers and interface to WSM has been developed.

This system consists of SW racks, rails mounted on and modules with diodes of more colors, buttons and displays plugged in. One module is dedicated to one bin in the warehouse. One bin is dedicated to one PN = material. Qty is reported to the WMS. The system is called Pick by Light.

Basic module is presented at the picture 2 below.

Fig. 2. Pick By Light Modul (Source: http://www.lightningpick.com/docs/TandyBrandsCaseStudy_Pick.pdf)

Application of whole system is at the picture 3 below:

Fig. 2. Application of Pick By Light (Source: www.google.com, 2008.)

The main difference between the paper and the paperless system is that the operator is order to pick the material and its qty by light. When the order is uploaded to the system, PBL controller switches on lights installed on modules/bins of related material that is requested to be picked up. When OP goes through the Warehouse and picks the material from the bins OP confirms these picks by pushing on button and light automatically switches of as confirmation of pick and successful system transaction. The operator can also set picked QTY or inform about missing material by pushing special button. In the same time these picks are reported to the inventory system for next purposes. This system can be updated for material PN scanning that helps to increase currently high reliability. In case of implementation of more color diodes, 7 operators can work in the WH in the same time because of the fact that one color is assigned to one operator. On the other side, more operators can work on one order, if supervisor decides to assign colors to orders instead of assigning to operators.

Benefits of this technology are structured in table 1 below:

1	Speed	System allows up to 70% increase on speed of picking and consolidation of mixed material orders.
2	Fail rate	By using of this technology can be assured that material is not missing.
3	Directivity	All displays and diods are easily visible to avoid overlook
4	Flexibility	Due to 7 colors more operations can be done in the same time
5	Expandability	System can be expanded without any additional IT and SW investments
6	Fast Inventory Check	System connected to WMS can report qctual qty in bin at any time
7	Effeciency	By current application saves on manpower can be acchieved

Tab. 1. Features of the system (Source: Paperless picking and supply of material, KBS-GMBH.)

4. Feasibility study

The logistics centre had to face problems with consolidation of small kits from various material of low quantity that are then supplied to customers. Regarding to results of analysis that discovered high fail rate in picking and low speed company has decided to implement **Pick By Light** system to one of repacking lines used for kitting of kits that consist of documents, books, CDs and others.

Assumption based on analysis results is below:

- 190 PN located in rack,
- 120 kits must be produced in 1 hour = 120 UPH,
- Current man power is 5 operators.

The department responsible for improvements prepares an option that counts with implementation of 4 racks, each rack with 48 bins where each bin is equipped with module with one diode. First simulations of solution initiated information that is possible to increase **UPH of 40%** or decrease necessary manpower from **5 to 3**.

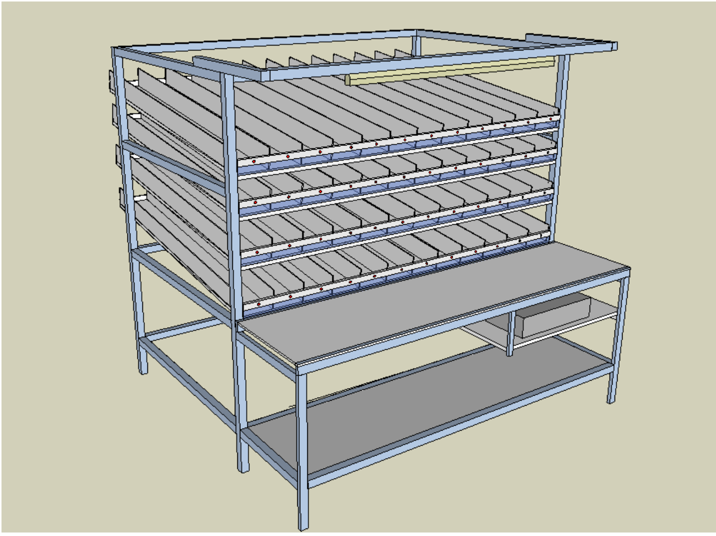


Fig. 3. D drawing of Pick By Light

Regarding to the simulation's result simple Payback Time in months was calculated.

Payback Time in months (PiM) indicator presents "How long will it take to get my money back" and is used for its simplicity. An investment's payback period in months is equal to the net investment amount divided by the average month cash flow from the investment.

$$PT = \frac{IN}{Z_M} \tag{1}$$

where IN presents costs of system's implementation and ZM presents income of the project – in our case income is equal to safe on costs for manpower.

INPUTS

Price of the system	450 000 CZK
Saved manpower per month	225 000 CZK

CALCULATION

$$PT = \frac{IN}{Z_M} = \frac{450000}{225000} = 2 \tag{2}$$

EVALUATION

Result PT = 2 presents that all invested money will return within 2 month after investment is done. This return is caused by saves on manpower of 2 people.

5. Conclusion

Companies focused on logistics processes are just now facing to not satisfying development on international and national markets and to economical situation of their customers. Companies are pushed to save money and to decrease prices of their services and products. One way how to increase the speed and frequency, decrease fail rates and optimize place utilization what leads to money saving is to implement new technologies as Pick By Light connected thru interface to Warehouse management System is. Regarding to the feasibility study and its results it is possible to say that this technology can bring measurable benefits to investors very soon.

Acknowledgement

Příspěvek vznikl v rámci řešení projektu MD ČR – CG731-118-520 „Prostorová lokalizace veřejných logistických center.“

References

[1] Řízený sklad [online]. 2006 [2008-11-01]. WWW: <http://www.rizeny-sklad.cz/>.

[2] Řízený sklad v novém logistickém centru společnosti Intersnack v plném provozu [online]. 2008 [2008-11-01]. WWW: <http://www.pressweb.cz/zprava/7995-rizeny-sklad-novem-logistickem-centru-spolecnosti-intersnack-plnem-provozu>.

[3] Implementace systému SmartStock.WMS ve společnosti Nohel Garden [online]. 2008 [2008-11-01]. WWW: <http://www.smartstock.cz/case-studies/industry/Obchod-a-distribuce-11>.

[4] Řízený sklad v dodavatelském řetězci [online]. 2007 [2008-11-01]. WWW:<http://www.businessworld.cz/bw.nsf/id/Rizeny-sklad?OpenDocument&cast=1>.

[5] Paperless picking and supply of material [online]. 2008 [2008-11-01]. WWW:<http://www.kbs-gmbh.de/en/beleglose_kommissionierung/#pick_by_light>.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Notation for Computational Independent Model in Model Driven Architecture

[*]Martin Kardoš, [**]Zuzana Bizoňová, [***]Marek Tavač
[*]University of Žilina, Faculty of Management Science and Informatics, Department of InfoComm Networks, Univerzitná 8215/1, 01026 Žilina, Slovakia, {Martin.Kardos}@fri.uniza.sk
[**]University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitná 8215/1, 01026 Žilina, Slovakia, {Zuzana.Bizonova}@fri.uniza.sk
[***]University of Žilina, Faculty of Management Science and Informatics, Department of Software Technologies, Univerzitná 8215/1, 01026 Žilina, Slovakia, {Marek.Tavac}@fri.uniza.sk

Abstract. Model Driven Architecture (MDA) is modern software engineering methodology which models information systems at four levels: Computational Independent Model (CIM), Platform Independent Model (PIM), Platform Specific Model (PSM) and Code. Modeling of the information system on higher levels of abstraction brings advantages in transparency and comprehension in the development. Nowadays, development of information systems and software applications based on MDA concentrates on modeling lower layers as PIM and PSM but the highest layer of this architecture - CIM has insufficient attention. This paper deals with possible notations for modeling CIM in MDA and their comparison.

Keywords: Model Driven Architecture, MDA, Computational Independent Model, CIM, Computational Independent Modeling, Information System, Architecture of Information Systems, Business Analysis, Business Process, Business Modeling

1. Introduction

System development life cycle (SDLC) of information systems or software applications is a well known approach. Commonly, it has five phases in following order: Planning, Analysis, Design, Implementation and Maintenance [1]. However, new methodologies are still arizing which incorporate SDLC and provide a little bit different view on system modeling.

One of them is Model Driven Architecture (MDA) [2]. MDA has been created by established organization Object Management Group (OMG) which is publishing standards in the field of Software Engineering and development of information systems. MDA is a model driven framework for development of information systems and software applications. The main idea of this concept is in creation of information models on different levels of abstraction and their following transformation to final executive code. According to the definition of MDA the process of creation of software starts with information models on computational independent level (CIM) and transforms these into models on platform independent level (PIM). These models are enriched and then transformed into platform specific models (PIM) which after the last transformation result in executable source code [3]. This concept is displayed on fig.1.

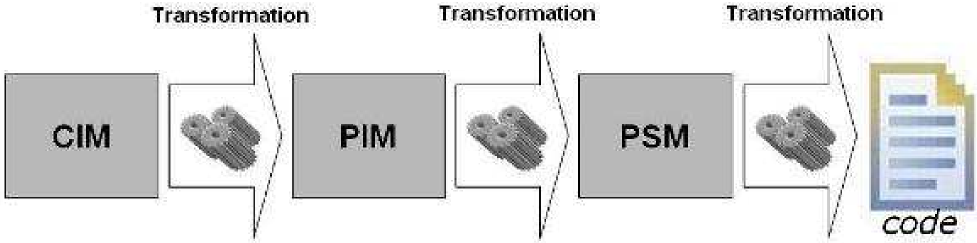


Fig 1. Concept of Model Driven Architecture. Development of an information system on various levels of abstraction.

1.1. Computational Independent Model

The CIM represents highest abstract level of modeling system. It covers the first two phases of SDLC.

The business process maps, workflows and organization structure are modeled on this level of abstraction. The CIM is describing the situation in which the system will be used. Such a model is sometimes called a domain model or a business model. It may hide much or all information about the use of automated data processing systems. Typically such a model is independent of how the system is implemented.

A CIM is a model of a system that shows the system in the environment in which it will operate, and thus it helps in presenting exactly what the system is expected to do. It is useful, not only as an aid to understanding a problem, but also as a source of a shared vocabulary for use in other models. In an MDA specification of a system CIM requirements should be traceable to the PIM and PSM constructs that implement them, and vice versa.

However, nowadays there is still a lack of the methods to transform the requirements identified on this high level of abstraction to the lower level models – PIM – that can be described in UML.

The CIM models are the models closest to the reality. They should be understood even by people with no computer knowledge and therefore used for mapping of the functionality of the real system.

Before the information system is designed, we need to get the whole picture of how the real system works – we need to recognize processes in the system. When such a model (CIM) of the system is created we can see more clearly which processes are redundant and which of them may be computer supported. These can further be transformed into lower level models (PIM, PSM) and the information system is implemented out of them.

1.2. Platform Independent Model

The PIM describes how the system will be created without any specification of technologies used for the implementation of the system. It does not express details specific for a certain platform or technology (Java, MySQL, Corba ...). The notation mostly used for description of PIM is UML.

1.3. Platform Specific Model

The PSM describes the platform specific details of the PIM model. It serves as the mediator between the UML model describing the functionality of the system and the final implementation code.

2. Notations for CIM

The notation used by IT architects for description of CIM is often very different from the notation used by business analysts. A business analyst can be seen as a person with basic computer knowledge but at the same time he is a professional for the domain which will be supported by the information system. Thus, it is very important that business analysts and IT architects co-work in one team. Next important issue is that a notation has to be easy understandable for both sides because the CIM level should be easily transformed to the PIM level.

Many business analysts and IT architects use and understand Data Flow Diagram notation (DFD) or Business Process Modeling Notation (BPMN) for CIM level modeling.

2.1. DFD notation

General idea of Data Flow Diagram notation is in decomposing of the modeled system into lower-level subsystems. Each subsystem represents a process or activity and illustrates how data are processed in the form of inputs and outputs. At the last level processes or activities cannot be more decomposed. On the top of the hierarchy is the context diagram. It represents simplified view on the modeled system (example is on fig.2).

Fig 2. A simple example of the context diagram in DFD notation. It represents a process how a customer can order some goods.

The DFD [1] notation is used mainly in System Analysis and Design. It is easy to learn and understandable for people with no computer knowledge too. But it has one important disadvantage. It is an old notation from late 1970^{th} and probably there does not exist a standardized metamodel of DFD notation based on XMI which could be used for the exchange of DFD diagram layout information. The other problem is that many modern software engineering methodologies claim that DFD is not recent anymore and so it is hard to find a software support for it in MDA tools. It would be then very difficult to make automatic transformation from CIM to PIM level with this notation.

2.2. BPMN notation

The Business Process Modeling Notation (BPMN) [4] was created by OMG in 2004 and is used for modeling business processes. The primary goal of BPMN is to provide a notation that is readily understandable by all business users, from the business analysts that create the initial drafts of the processes, to the technical developers responsible for implementing the technology that will perform those processes, and finally, to the business people who will manage and monitor those processes. Thus, BPMN creates a standardized bridge for the gap between the business process design and process implementation [4]. We could say that is relatively new notation. It provides graphical notation for business processes by Business Process Diagram (BPD). The BPMN emphasizes data flows and represents a joint between activities and data. It is very similar to activity diagram in UML.

In the figure 3 we have displayed the same situation as on figure 2, this time described by BPMN.

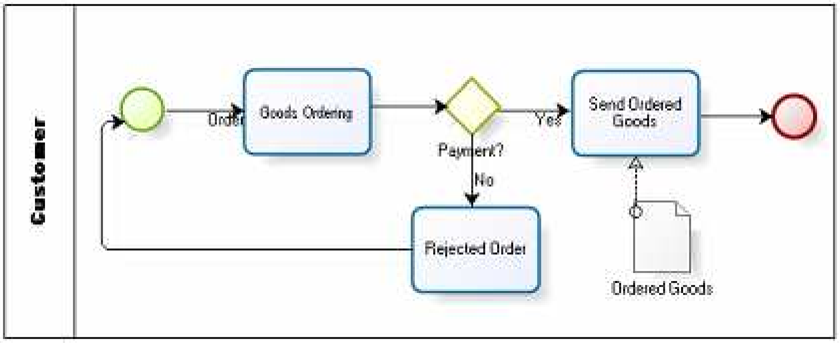


Fig. 3. Variation of figure 2 in BPMN notation.

In this case the diagram on figure 3 looks very similar to activity diagram in UML. However pure UML is not sufficient for description of CIM because of its lack of some expressions which can be found in BPMN. For example a “milestone” that describes a subpartition within a process and this way it is possible to describe much more complex systems then with an activity diagram.

Unlike in DFD, there exists the BPMN metamodel described by XMI which enables exchange of BPMN diagram semantic information. As we mentioned before, DFD is based on dataflow modeling between processes and external entities. Therefore the DFD notation also lacks many possible expressions, for example it is not possible to clearly express decision in the process. The reader can see the difference when comparing the process of rejection on the fig. 2 and fig 3.

3. Conclusion

In this article we have described and compared possible notations for description of a CIM in MDA. As the result of this we have chosen the BPMN notation as the most expressive possibility for description of a system on the highest level of abstraction. It is easily understandable by both business analysts and IT architects and is further exchangeable thanks to its diagram semantic information described in the XMI metamodel.

Our research will further continue in search for an automatic transformation of CIM model based on BPMN into PIM model of the MDA.

References

[1] DROZDOVÁ, M. *Study materials of Architectures of Information Systems*. Faculty of Management Science and Informatics, University of Žilina, 2008.

[2] OBJECT MANAGEMENT GROUP *Model Driven Architecture (MDA).* http://www.omg.org/docs/ormsc/01-07-01.pdf, 2001.

[3] ANDREAS, M., LOOS, P. *Software support for the Computation Independent Modelling in the MDA context.* http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-376/paper5.pdf, ISSN1613-0073, 2008.

[4] OBJECT MANAGEMENT GROUP *Business Process Modeling Notation Specification*. http://www.bpmn.org/Documents/OMG%20Final%20Adopted%20BPMN%201-0%20Spec%2006-02-01.pdf, 2009.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Access Network and Technology Cable Modem, xDSL, FTTx

*Dagmar Kmeťová
* University of Žilina, Faculty of electrical engineering, Department of telecommunications and multimedia, Univerzitná 1, 01026 Žilina, Slovakia, dagmar.pajdusakova@fel.uniza.sk

Abstract. This paper describes access networks, which use technologies as FTTx (Fiber to the home), HFC (Hybrid Fiber-Coax) and xDSL (Digital Subscriber Line). Millions of subscribers around the world are today connected via broadband (cable, xDSL, FTTx) access. DSL, FTTx and cable modem are the dominant broadband technologies in the world. DSL is the most common technology with a 65% share of the global broadband market. Cable modem comes in second place and than comes technology FTTx. Utilizing this technologies independent from countries where they are offer.

Keywords: HFC, FTTx, xDSL, OLT, ONU.

1. Introduction

Most broadband access networks still rely on either twisted pair copper wires (DSL) or coaxial cable (HFC networks). Currently, are adapting their networks for triple play, the combination of Internet, TV and telephone services distributed over one network. HFC networks evolved out of the original TV distribution networks and combine optical fiber with coaxial cable. An optical connection arrives in an optical node that feeds one service area (SA) with a shared coaxial network [1]. Digital Subscriber Line is a relatively advanced method of transmitting information over ordinary (copper) telephone lines. FTTH is simply the 100 percent deployment of optical fiber in the access network.

2. Access Network

As noted previously technologies DSL, FTTx and Cable modem are the dominant broadband technologies in the world. These technologies are utilized in access networks. Technology DSL in access network use device DSLAM, cable technology use device CMTS (Cable Modem Termination System) and technology FTTx use OLT (Optical Line Termination). Figure 1 shows architecture access network with technologies xDSL, HFC and FTTx.

2.1. Technology xDSL

DSL is a relatively advanced method of transmitting information over ordinary (copper) telephone lines. The applications of DSL involve the transport of high-speed data, voice and video, to residential and business subscribers. DSL in general describes the technology, while xDSL represents individual varieties of DSL technology. There are the following versions of DSL technology:

- HDSL (High data rate Digital Subscriber Line) – this is a symmetric technique needing two (currently) or 3 (in the future) twisted pairs. The maximum transmission speed is 2 Mbps and maximum range – up to 5 kilometers (for 0.5 mm cable);

- SDSL (Single line Digital Subscriber Line) – this is a symmetric technique needing single pair of wires. The maximum transmission speed is 2.3 Mbps and maximum range – up to 2 kilometers (for 0.4 mm cable);
- ADSL (Asymmetric Digital Subscriber Line) – this is an asymmetric technique, which means that data transmission speed to the subscriber is higher than data transmission from the subscriber. This technique needs single pair of wires. The maximum data transmission speed is 8 Mbps to the subscriber and 640 kbps from the subscriber (for 0.5 mm cable);
- SHDSL (Single-pair High-speed Digital Subscriber Line) – this is a symmetric technique with maximum transmission speed of 2.3 Mbps and maximum range of 3 kilometers. This technique needs single-pair wire. If two pairs are used, the transmission speed doubles.
- VDSL (Very high data rate Digital Subscriber Line) – in this technique the transmission can be performed in symmetric or asymmetric way. It needs single pair of copper wire. The maximum transmission speed is 52 Mbps and maximum range – 300 meters [2].

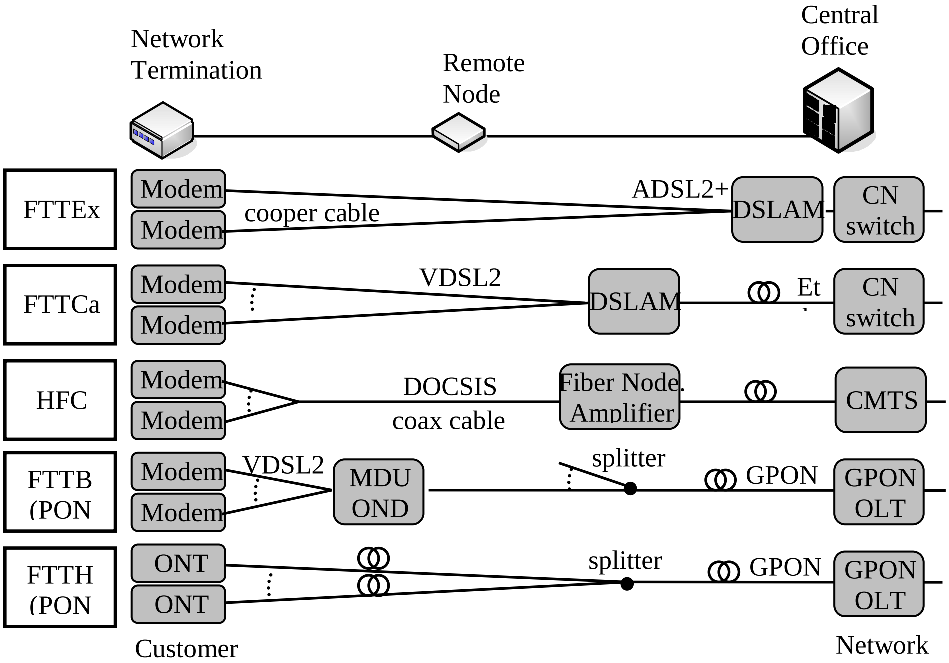


Fig. 1. Access Network, technologies xDSL, HFC, FTTx.

2.2. Technology FTTx

FTTH technology represents an attractive solution for providing high bandwidth from the central office to residences and to small and medium sized businesses. Active and passive are two commonly used FTTH architectures for FTTH deployment. Active Architecture is also called as Pont 2 Point (P2P) and Passive Optical Network (PON) architecture is called Point to Multi Point (P2M) [3].

Active Optical Network

Active optical networks rely on some sort of electrically powered equipment to distribute the signal, such as a switch, router, or multiplexer. Each signal leaving the central office is directed only to the customer for whom it is intended.

Active Star Ethernet (ASE) architecture is a point-to-point architecture in which multiple premises share one feeder fiber through a remote node located between the Central Office (CO) and the served premises. The remote node can be shared between four to a thousand homes via dedicated distribution links from the remote node (Fig. 2).

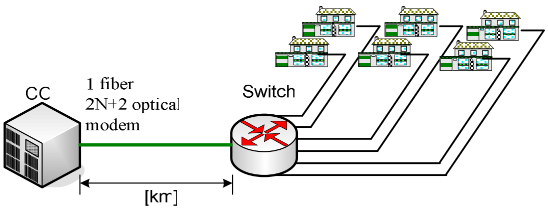


Fig. 2. Active Optical Network

Passive Optical Network

Passive Optical Networks are shared media, or point-to-multipoint, networks in which multiple users share the same bandwidth. In this network architecture, passive optical splitters are used to divide the bandwidth from a single fiber among up to 64 users over a maximum distance of 10-20km. In a PON, a CO-located OLT connects to customer-premise-located ONTs to terminate the fiber. Both the OLT and the ONT are powered. The architecture is called passive because all splitters and intermediate equipment located between the CO and the ONT is passive (Fig. 3).

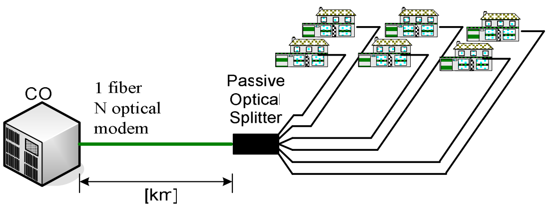


Fig. 3. Passive Optical Network

2.3. HFC network

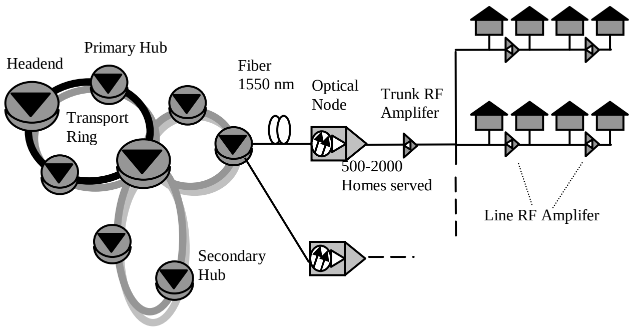
Although Cable TV was known to be a one way system for broadcasting TV programs. One way transmission of information was sufficient for these purposes. Customer's demands and competition had stimulated operators of cable networks to build this network not only for one-way transmission, but also for the two-way one. The one-way components were replaced by the two-way ones to ensure the communication in both ways.

Nowadays a two way communication across the Cable networks is possible, and this step helps in delivering new services to the users across the cable networks, for example DOCSIS, Interactive TV, Pay Per View (PPV) etc.

The structure of the cable network was following: Antenna system, Master station, Distributional network [4].

New generation of cable systems (Fig.4) are the hybrid fiber-coax distribution systems that consist of optical rings with additional hubs included along the rings. The signals are conveyed from the hubs to the nodes over the optical fibers. In the nodes the optical signals are converted into electrical. After that the signals are distributed to the subscribers by coaxial distribution system that has capacity for 500 to 2000 (typically 500) subscribers. Hence, previously built coaxial distribution systems were combined through the usage of optical rings and in this way the subscriber service was localized in one headend [5], [6].

Data-Over-Cable Service Interface Specifications (DOCSIS) is the standard technology used by most cable operators to transmit data over a HFC network [7].



Standard	Bit rate [Mbps] (shared)	
	DS	US
DOCSIS 1.0/1.1	42-56	10
DOCSIS 2.0	42-56	31
DOCSIS 3.0	180	120

Fig. 4. Architecture HFC network

3. Conclusion

These technologies support telephony and data services, analog and digital video services, which are very interesting for customers. In the present time and also in the future it will be more and more interest of customers in Triple Play service.

Acknowledgement

This project was supported by the VEGA grant No. 1/0375/08 of Ministry of Education of the Slovak Republic.

References

[1] LANNOO, B. - VERBRUGGE, S. - OOTEFHEM, J.: *Evolution of Access Networks: FTTH and WiMAX*.

[2] POLAK, T. - RYPINA, T.: *Last mile - Internet access technology*, http://www.dipol.com.pl/bib80e.htm

[3] AB-RAHMAN,M., S. - CHE AZIZ, S.,A. - JUMARI, K.: *Protection route Mechanism for Survivability in FTTH-PON*, December 2008.

[4] SHAH, N. - KOUVATSOS, D. - MATRIN, J. - MOSER, S.: *A Tutorial on DOCSIS: Protocol and Performance Models*.

[5] HOTTMAR, V. - SCHWARTZ, L. - TRSTENSKÝ D.: *End User Box for Interactive Cable Television*, IEEE Transaction on Consumer Electronics, Vol. 53, No. 2, p. 412 – 416, ISSN 0098-3063, May 2007.

[6] ADAMEC, B.: *Access Devices for Multimedia Home Networks*, X International PhD Workshop OWD – Wisla, Poland, ISBN 83-922242-4-8, 2008.

[7] Data Over Cable Service Interface Specifications DOCSIS 3.0, *Physical Layer Specification*; Cable Television Laboratories, Inc.; 04.08.2006.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Fixed Charges in Re-building of the Marshaling Yard System

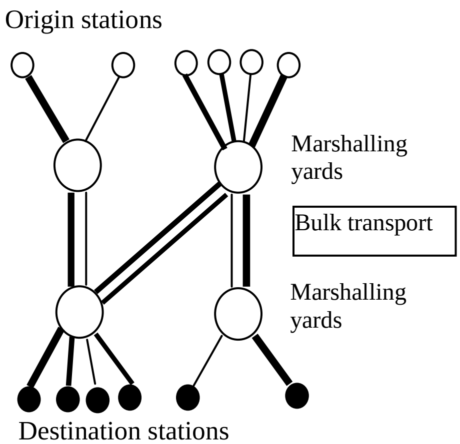
*Michal Koháni
*University of Žilina, Faculty of Management Science and Informatics, Department of Transportation Networks, Univerzitná 1, 01026 Žilina, Slovakia, {kohani}@frdsa.fri.uniza.sk

Abstract. The marshaling yard system is a special case of transportation system, where flows of goods from primary sources to customers are concentrated in terminals to create a bigger flow between terminals. The problems with fixed cost in this system can have strong influence on the solution. In this paper we deal with two different ways of fixed charge estimating to re-build of the current system to more effective one.

Keywords: marshaling yard, fixed cost, many-to-many distribution system, cargo railway system

1. Introduction

A cargo railway system is a transportation system, which has approximately the same number of primary sources as number of customers and provides transport of carriages between railway stations. Such a system is called "Many-to-Many Distribution System". In this case, demands of customers form a matrix of yearly flows of carriages from source railway stations to destination railway stations. We denote this matrix as $\boldsymbol{B} = \{b_{sj}\}$, for $s \in S$ and $j \in J$, where S is a set of source railway stations and J denotes the set of destination railway stations. The fact that unit cost of transportation is smaller when bigger bulks of items are transported, approves concentration of flows between different pairs of source and customer to stronger flows at least on a part of their way. This flow concentration needs marshalling yards, in which transhipment of transported items is performed and bigger bulks are formed or, on the other side, where bulks (direct trains) are split into smaller groups (manipulating trains) designated to different destination railway stations.



On the contrary to the classical distribution systems, in which big bulks leave primary source, another situation emerges in this many-to-many distribution system. Primary sources send relative small bulks of items and it is useful to concentrate them to bigger bulks in the marshalling yards located near the sources and then to send these bigger bulks to remote marshalling yards and to split them there (see Fig. 1).

We focused on the system, in which a railway station is assigned to only one marshalling yard and an exchange of the consignments between the this station and other stations is done via this assigned marshalling yard, as it is shown in Fig. 1. Furthermore, we consider the general case, in which any origin station is a destination station simultaneously.

We do not make any difference between an origin station and a destination station hereafter and we introduce the set *J'* of railway stations in general.

Matrix ***B*** gives the yearly volumes of the carriages denoted as coefficients b_{sj}, which are sent from the object *s* to the object *j* and it gives the total yearly volume sent from the object *j* to the object *s* by coefficients b_{js}. In the next sections we try to model a symmetrical many-to-many distribution system with unique assignment of customers to terminals.

2. Model of many-to-many distribution system design problem

Let us consider a case with a linear cost estimation function with unit cost e_0 for transport of one item along unit distance on the way from an origin railway station to a marshalling yard or from a marshalling yard to a destination railway station. Next, let us consider unit cost e_1 for transport of one item along unit distance on the way from one to other marshalling yards. Furthermore we denote a set of possible terminal locations by symbol *I*, where each place $i \in I$ is associated with yearly fixed charge f_i for building and performance of a terminal at the location *i* and with unit cost g_i for transhipment of one unit in the terminal. In accordance to the previous definition, we denote by *J'* the set of objects which mutually exchange the carriages of the yearly total amounts b_{sj} from $s \in J'$ to $j \in J'$. Symbol d_{ij} denotes the distance between locations *i* and *j*. Our objective is to assign each sending or receipting object (railway station) to exactly one terminal so that the total yearly cost of the designed system is minimal. If we denote by $y_i \in \{0, 1\}$ for $i \in I$ the bivalent variable, which corresponds with the decision whether a marshalling yard will or will not be built at the place *i* and if we introduce the variable $z_{ij} \in \{0, 1\}$ for $i \in I$ a $j \in J'$, which says whether the station *j* will or will not be assigned to the place *i*, than we can formulate the following mathematical programming model of problem:

$$\text{Minimise} \sum_{i=1}^{m} f_i y_i + \sum_{i=1}^{m} \sum_{j=1}^{n} (e_0 d_{ij} + g_i)(\sum_{s=1}^{n} b_{js} + \sum_{s=1}^{n} b_{sj}) z_{ij} +$$

$$+ \sum_{i=1}^{m} \sum_{k=1}^{m} e_1 d_{ik} \sum_{j=1}^{n} \sum_{s=1}^{n} b_{sj} z_{ij} z_{ks} \;. \qquad (1)$$

$$\text{Subject to} \quad \sum_{i=1}^{m} z_{ij} = 1 \quad \text{for } j=1, \dots, n, \qquad (2)$$

$$z_{ij} \le y_i \quad \text{for } i=1, \dots, m, j=1, \dots, n, \qquad (3)$$

$$y_i \in \{0, 1\} \quad \text{for } i=1, \dots, m, \qquad (4)$$

$$z_{ij} \in \{0, 1\} \quad \text{for } i=1, \dots, m, j=1, \dots, n. \qquad (5)$$

This model belongs to the discrete quadratic programmes due the third term of *(1)*.

3. Fixed Charges Calculation for the Re-building System

In the case, when we want to re-build the distribution system, we need to calculate new fixed charges. We have few possibilities, how to manage it.

First alternative, we have, is to use the original fixed charges. This possibility brings several complications. The new solution can be deformed, since the model doesn't take into account the better equipment in the marshaling yard with bigger fixed charges.

$$\underline{f_{ij}} = f_{ij} \tag{6}$$

The second alternative is based on the calculation of the fixed charges using

$$\underline{f_{ij}} = \max\{f_{ij}, \forall i = 1..m, \forall j = 1..n\} - f_{ij} \tag{7}$$

In this case we estimate charges, which are necessary for fully equipement of the marshaling yard fully equipped as the yard with maximal fixed charges.

4. Solving Technique

To be able to use a specific solver for location problems, the model of the problem must be reformulated to the linear form of the uncapacitated facility location problem. Let us focus on the particular ways of quadratic term linearization, which consists in replacing the sum in brackets by some term t_{is}, which doesn't depend on index k and which could be good approximation of the sum.

$$\sum_{i \in I} \sum_{k \in I} e_1 d_{ik} \sum_{j \in J'} \sum_{s \in J'} b_{sj} z_{ij} z_{ks} = e_1 \sum_{i \in I} \sum_{j \in J'} \sum_{s \in J'} b_{sj} (\sum_{k \in I} d_{ik} z_{ks}) z_{ij} \tag{8}$$

This replacing is done subject to constraint $\sum_{k \in I} z_{ks} = 1$.

The way of estimation makes use of so-called set $K(i, s)$ of relevant locations to which object (customer-source) s is allowed to be assigned. It should be noted here that prime cost of one unit transport from terminal i via terminal k to object s is $e_1 d_{ik} + e_0 d_{ks}$. This transportation chain is economically sensible only if this prime cost is less than cost of the direct transport after from i to s, what is $e_0\, d_{is}$. This idea enables to establish the set of relevant locations $K(i, s) = \{k \in I: e_1 d_{ik} + e_0 d_{ks} < e_0\, d_{is}\}$ and to take into consideration only k's from this set. We can define $t_{is} = \beta_{is} d_{is}$, where

$$\beta_{is} = (\sum_{k \in K(i,s)} d_{ik} / d_{is}) / |K(i,s)|, \tag{9}$$

what is the average of relevant ratios.

Obtaining estimation t_{is} the term *(9)* can be rewritten as follows: $e_1 \sum_{i \in I} \sum_{j \in J'} \sum_{s \in J'} b_{sj} (\sum_{k \in I} d_{ik} z_{ks}) z_{ij} = e_1 \sum_{i \in I} \sum_{j \in J'} (\sum_{s \in J'} b_{sj} t_{is}) z_{ij}$, what is a linear expression. Then the problem *(1)*, *(2)*-*(5)* takes form of a linear uncapacitated location problem, which is easy to solve using BBDual procedure [1].

5. Experiments

The set of test problems was created from a real network and from a real many-to-many problem, which was formulated in the frame of the project [1], in which railway infrastructure and yearly carriage flows were analyzed. This original problem, in which 53 possible marshaling yards and almost five hundred railway stations-customers were considered, was partitioned into sequence of smaller subproblems. The original fixed and variable charges were adjusted for a particular subproblem proportionally to its total flow size to obtain nontrivial results. This way, 10 instances came into being for each problem size. A goal of this investigation was to compare two ways of the fixed charge estimations. In the Table 1 there are solutions of both alternatives of fixed charges and the 50 possible marshaling yards in the system. In the column *Marshaling Yard* is the name of the marshaling yard candidate (in the order of importance), in the column *Fix1* is the frequency of appearance the marshaling yard in the solution of 100 solved subproblems with fixed costs (6), in the column *Fix2* is the frequency of the marshaling yard appearance in the solution of 100 solved subproblems with fixed costs (7).

Marshaling Yard	Fix1	Fix2	Marshaling Yard	Fix1	Fix2
Bratislava UNS	35	100	Kúty	47	9
Košice	39	100	Michaľany	26	0
Žilina	0	100	Vrútky	4	0
Zvolen	13	78	Zohor	16	0
Nové Zámky	35	52	Fiľakovo	0	0
Leopoldov	31	18	Strážske	35	28
Čierna nad Tisou	10	4	Púchov	71	0
Komárno	0	4	Prievidza	36	27
Plešivec	21	18	Devínska Nová Ves	65	0
Banská Bystrica	70	0	Galanta	45	0
Trenčianska Teplá	25	22	Levice	0	0
Prešov	18	0	Poprad - Tatry	65	0
Trnava	9	0	Topolčany	83	1
Štúrovo	96	16	Nové Mesto nad Váhom	11	0
Spišská Nová Ves	15	14	Kraľovany	81	14

Tab. 1. Results for 50 possible locations of marshaling yards

6. Conclusion

In the comparison of the two different ways of the fixed charge estimation for this type of problem we can obtain better solution with adjusting fixed cost as cost, which we repair it in accordance to the previous suggestion. In the solution with original fixed costs the biggest marshaling yards were not selected due to the bigger fixed cost and mostly in the solution were selected less equipped marshaling yards with lower fixed costs.

Acknowledgement

This work was supported by the grant VEGA 1/0591/09.

References

[1] JANÁČEK A KOL.: *Správa o riešení pilotného projektu úlohy 5.3 „Programové riešenie pre optimalizáciu riadenia prepravnej práce“*. Žilina, Dec. 2001, Fakulta riadenia a informatiky, ŽU-Žilina, 11 p

[2] KOHÁNI, M.: *„Umiestňovanie vlakotvorných staníc v železničnej sieti“*, zborník seminára „Úlohy diskrétní optimalizace v dopravní praxi – Lokace středisek s negativními vlivy na okolí“, Katedra informatiky v dopravě, Dopravní fakulta Jana Pernera, Univerzita Pardubice, Pardubice, november 2008, str.43-51

[3] KOHÁNI, M.: *„Metódy umiestňovania vlakotvorných staníc“*, elektronický časopis „Perner's Contact“, Univerzita Pardubice, Pardubice, december 2008

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Orientation and Placement of RFID Tags as a Factor in The Successful Postal Items Reading

[*]Peter Kolarovszki, [*]Juraj Vaculík
[*]University of Žilina, Faculty of Operation and Economy of Transport and Communications, Department of Communications, Univerzitná 1, 010 26 Žilina, Slovakia,
{Peter.Kolarovszki, Juraj.Vaculik}@fpedas.uniza.sk

Abstract. This publication primarily deals with orientation and placement of RFID tags into postal containers, which is one of the factors affecting the successful postal items reading. Publication contains several tests, which has been accomplished in Lyngsoe systems testing centre located in Denmark. These tests include combination of active and passive technology of radio frequency identification (RFID), which is a form of automatic identification and data capture (AIDC). Technology uses electric or magnetic fields at radio frequencies to transmit information.

Keywords: RFID, Orientation, Test, Technology, Tag, Reader.

1. Introduction

RFID is a highly useful and exciting technology. It seems that everywhere one looks there is some article about RFID and the huge benefits the technology promises. Moreover, there are many examples that demonstrate how this technology is fulfilling its potential. RFID reduces costs and dangers, improves processes and enhances entertainment. RFID tags can be read through many materials (includes metal or water) affording identification without line of sight. One question is - if is there any interdependency between active and passive tag in regard of its readability. This question will be solved later in this article. Next three sections will provide answer of the following question:

- What is the architecture of RFID technology and what type of the RIFD tags exist?
- How depending tag orientation on its readability?
- How is interdependency between active and passive tag in regard of its readability?

2. What is RFID?

Radio frequency identification is a wireless data collection technology that uses electronic tags which store data, and tag readers which remotely retrieve data. It is a method of identifying objects and of transferring information about the object’s status via radio frequency waves to a host database. RFID is not necessarily a direct replacement for bar codes, but as the costs of RFID systems continue to decrease, the functional utility of RFID will greatly surpass that of bar codes. [2]

2.1. Components of an RFID System

An RFID system is a set of components that work together to capture, integrate, and utilize data and information. This section describes some of them. The components are as follows:

- Sensors, Tags, Antennas, Readers.
- Connectors, Cables, Networks, Controllers.
- Data, Software, Information Services.

2.2. RFID tags and frequencies

An RFID tag is a small device that can be attached to an item, case, container, or pallet so it can be identified and tracked. It is also called transponder. The tag is compodes of a microchip and an antenna. These elements are attached to a material called substrate to create an inlay. [4]

Tags are categorized into three types based on the power source for communication and other functionality.

- Active.
- Passive.
- Semi-passive.

Carrier frequencies

Today, there are four carrier frequencies implemented for RFID that are popular globally: 125 KHz, 13.56 MHz, UHF ranging from 866 to 950 MHz depending on local country radio regulations and microwave frequencies of 2.45 GHz and 5.8 GHz. There is also frequency range 430-440 MHz, allocated to amateur radio usage around the world. The ISM band 433.05-434.790 MHz is located near the middle of the amateur radio band. The amateur radio band has emerged as an RFID channel in a number of applications. The frequency range has been called the "optimal frequency for global use of Active RFID". [3]

Tag orientation (polarization)

The read range depends on tags antenna orientation. How tags are placed with respect to the polarization of the reader's field can have a significant effect on the communication distance for noth HF and UHF tags, resulting in a reduced operating range of up to 50%, and in the case of the tag being displaced by 90° and not being able to read the tag at all. The optimal orientation for the two antenna coils (reader and tag) to be parallel to each other (Figure 1). UHF tags are even more sensitive to polarization due to the directional nature of the dipole fields.

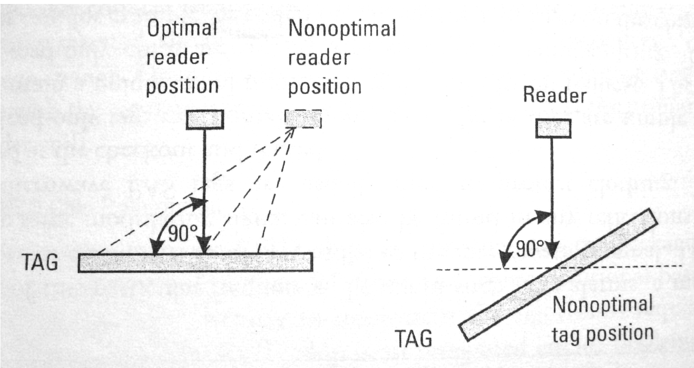


Fig 1. Optimal and nonoptimal tag and reader position [1]

3. Interdependency between active and passive tag

Like I mentioned above, several tests has been made in Lyngsoe systems testing center located in Denmark. The following tests examined the interaction and interference of passive and active UHF tags, which were placed in a horizontal position.Table 1 clearly displays tags that have been tested.

Type of tag	Tag ID	In short version
Container tag	25300000042	CT42
passive UHF tag	300000002	P02
passive UHF tag	300000007	P07
passive UHF tag	300000019	P19
passive UHF tag	300000018	P18
passive UHF tag	300000021	P21
passive UHF tag	300000022	P22
active UHF tag	25302300044	A44
active UHF tag	25302300046	A46
active UHF tag	25302300047	A47
active UHF tag	25302300051	A51
active UHF tag	25302300055	A55

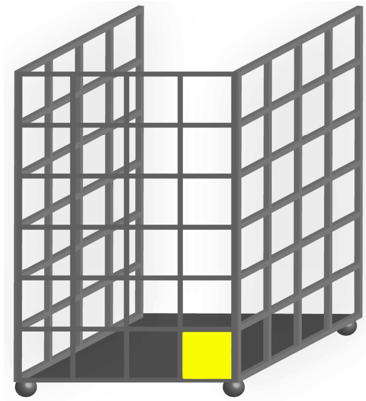


Fig. 2 Tags location

Tab 1 List of testing UHF tags

Test No.1

In test No.1, RFID tags were placed horizontally in a single crate that was in the sixth level shifting boxes (figure 2). Sequence tags with the appropriate "space" as follows: A55-2cm-P19-2cm-P18-5cm-A51-1cm-P21-1cm-A46-0cm-P07-0cm-A44-5cm-P02-0cm-P22-5cm-A49

The test was performed ten times while achieving the same accuracy scanning (table 2).

Reader	Exciter / Antenna	Tag ID	Signal	Time
11	2	300000022	0	14:36:49
11	1	300000022	0	14:36:49
11	1	300000019	0	14:36:49
11	1	300000018	0	14:36:49
11	1	300000002	0	14:36:50
11	2	300000002	0	14:36:50
1	10	25302300049	-70	14:36:50
1	10	25302300046	-69	14:36:50
1	10	25302300055	-70	14:36:51
1	10	25302300051	-72	14:36:51
1	10	25302300044	-69	14:36:51
1	10	25300000042	-76	14:36:51

Tab 2 Result from the test 1

Table 2 shows that were scanned four passive tags, which gives us the reading accuracy at 66.66%. As we have seen passive tags with numbers 07 and 21 were scanned. The reason may be their location or directly in close proximity to active tags No. 51,46,44. Accuracy of reading active tags placed in the container, as well as container label, reaching 100%.

Test No. 2

Due to the uncertainties resulting from the capture in the test No. 1, we have tried to locate the unread tags so that they were more distance. Individual labels are similar to the test. 1 placed in a horizontal position and also in the same order. The only change was the distance between the passive tag, with the number (21, 7) and labeled with the number of active (51, 46, 44). Space between the tags, which were implemented through the empty envelopes were as follows:

A55-2cm-P19-2cm-P18-5cm-A51-2cm-P21-2cm-A46-2cm-P07-2cm-A44-5cm-P02-0cm-P22-5cm-A49

The test was performed ten times while achieving the same accuracy scanning (table 3).

Reader	Exciter / Antenna	Tag ID	Signal	Time
11	2	300000018	0	14:43:35
11	1	300000007	0	14:43:36
11	1	300000019	0	14:43:36
11	1	300000022	0	14:43:36
11	2	300000021	0	14:43:35
11	2	300000002	0	14:43:37
1	11	25302300044	-70	14:43:39
1	11	25302300051	-69	14:43:39
1	11	25302300049	-69	14:43:39
1	11	25302300046	-69	14:43:39
1	11	25302300055	-72	14:43:39
1	11	25300000042	-78	14:43:39

Tab 3 Result from test 2

Table 3 shows that by increasing space between passive and active tags, which were not read in the test 1 we have achieved 100% accuracy of sensing. We found that changing the distance between the tags would have an impact on the accuracy of sensing.

4. Conclusion

Testing has a key role to play in every RFID installation. RFID technology is reliable and predictable, but it is not “plug –and-play” technology. Even the simplest tag-and-ship application will require tests. This publication shows you how RFID system works and provide answer on question about interdependency between active and passive tag in regard of its readability.

References

[1] LEHPALMER H. *RFID design principles.* Norwood (USA):Artech House Publishing, 2008.

[2] ANDRECHAK, WILSON, ZIMMARDI, RFID Item Level Management, Mullaney publishing group, 2007

[3] SHEPARD S. *RFID: Radio frequency Identification. New York (USA),* McGraw-Hill Publishing, 2005

[4] DENNIS E. BROWN. RFID Implementation. *New York (USA),* McGraw-Hill Publishing, 2007

[5] POIRIER CHARLES, RFID *Strategic Implementation and ROI*: A Practical Roadmap to Success, J. Ross Publishing, Inc., 2006.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Implementation and Measurement of a Failure Inferencing Fast Re Route Prototype

*András Balázs Lajtha, *Gábor Enyedi
*Budapest University of Technology and Economics, Faculty of Electrical Engineering and Informatics, Department of Telecommunications and Media Informatics, Műegyetem rkp. 3-9. H-1111 Budapest, Hungary, {lajtha.balazs, enyedi}@tmit.bme.hu

Abstract. Nowadays, the growing services of the Internet have become an integral part of the economy, science and everyday life as well. The need for real-time applications (e.g. IP telephony, video telephony and IPTV) appeared in the past few years and is increasing, presenting new requirements to IP networks. Transport layer service quality and hence failure-handling is one of the top priorities. Several new methods were developed to lower network response time to component failures. Most of these methods promise fast failure handling, loop-free paths and total coverage while making restrictions on the number of failures and the network topology. In this paper we are introducing an Interface Based Failure Inferencing protocol for handling single node failures in two-connected networks; we will present metrics that compare this solution to traditional Shortest Path First protocols.

Keywords: IP routing, QoS, proactive failure handling, interface based failure inferencing.

1. Introduction

With the ever increasing importance of the Internet IP networks are facing new Quality of Service requirements they were never designed for. Although these networks have become more and more successful in meeting the new QoS demands during the last few years, an important piece of the puzzle still lags behind. Real time traffic such as VoIP, IPTV or online gaming needs not only low latency, low packet lose and high bandwidth but also fast IP based resilience.

Although resilience was always one of the most important attributes of IP networks, current techniques are not able to fulfill the modern requirements. Routing protocols (e.g. Open Shortest Path First (OSPF), Intermediate system to intermediate system (IS-IS))[2] are handling occurring failures by first discovering the new topology of the network and then computing the forwarding paths, in this way making a global and reactive response, not even close to handle the 50-100ms requirements of real time traffic.

To suite real time needs, Internet Engineering Task Force has launched the IP Fast Reroute (IPFRR) framework [1] for creating fast resilience techniques with local and proactive response. Locality means that only the nodes adjacent to the failed resource change their states, and – contrarily to the flooding mechanism of routing protocols – no message of the failure is propagated. These solutions are reactive because alternative paths are computed and stored for the possible failure cases. By now, several local and proactive fast reroute solutions have been developed [3][4][5][7][8][9]. One of them is the Failure Inferencing based Fast Rerouting (FIFR)[10]. Variants of this algorithm offer fast resilience after link or node failures.

During our work we studied the performance of this algorithm. Using a test network, we measured the Key Performance Indicators (KPI) of recovery after a node failure. We compared the results with the performance of traditional OSPF protocol and with an enhanced version of OSPF using Netlink for failure detection. We also used extensive simulations for computing the increasing of path lengths when fast rerouting is applied.

The rest of this paper is organized as follows. First we discuss algorithm FIFR-N in Section 2. In Section 3 we turn to discuss the designing of prototype system we used for measurements, sharing the results in Section 4. Finally, in Section 5 we conclude our paper.

2. Failure Inferencing based Fast Rerouting for Handling Transient Link and Node Failures (FIFR-N)

In this section we briefly discuss FIFR-N algorithm. More detailed description can be found in [10].

As we mentioned earlier FIFR-N is a local and proactive solution for dealing with node and link failures. Without going into the details of the algorithm, we would like to introduce the basic concept. FIFR-N uses interface based forwarding tables. As opposed to traditional IP routing where the next hop is determined from the destination of the package, FIFR-N considers not only the destination but also the incoming interface of the package. This permits the system to store alternative paths each marked by destination-interface pairs that have different next hops than the shortest paths. When a node detects the failure of one of its neighbors, packages are transmitted via these alternative paths.

This solution is proactive because interface based routing tables are calculated in advance. FIFR-N claims to give a local response to a failure. There is no proven limit of the extent of the response, but it is believed that these paths only affect a smaller segment of the network. In our simulations we were trying to find empiric answer to the question of locality and optimality.

3. Prototype and simulation environment

To test the FIFR-N, we used a testbed deployed on PC based routers with GNU/Linux. A prototype network (see Figure X) was set up to prove the concept. The framework hasn’t implemented the automatic network discovery and route precalculation since these issues are not important for our measurements, network topology was an input parameter and the calculation of detours was made offline. Based on these calculations three types of tables were set up in each router: forwarding tables containing entries for the shortest paths; backwarding tables for shortest paths with a failed neighbor; FIFR tables for each interface containing the detours. Incoming packages were marked with the id of the interface, and rules determined which of the above tables should be used for the package with the given mark.

Fast failure detection was handled by Netlink. Netlink provided events when a network connection was removed or added. The testbed consisted of scripts listening to these events. If the kernel up-call signaled a failure or recovery of a link or node, the rules were changed. This method proved to be faster than cleaning and rewriting an entire routing table.

We also studied FIFR-N by extensive simulations organized in a Java framework. This framework provided a virtual network which could calculate actual paths from each source to each destination with a set of “failed” nodes or edges based on the same configuration file that the prototype consumed. The framework matched the result against the real shortest path in the given scenario, logged the results and kept statistics.

4. Results

To measure the KPIs we used the Distributed Benchmark System (DBS). We created a single stream on a path that when compromised would be redirected to a path marked by interface dependent routing rules. The network topology was designed in a way that the failure of a simple link triggers the same effect as the failure of a node. This facilitated not only the failure emulation but more importantly that of the recovery.

The DBS provided log files containing the time stamp of the departure of the packet with the time stamp of its receipt. These logs were processed to obtain different KPIs that characterized the solution. We made 20-20 measurements using OSPF, Netlink aided OSPF, and FIFR-N, each measurement starting in an intact network and constituting of a failure and a recovery of a network node. In average the FIFR-N solution responded to the failure in 120 ms while for the Netlink aided OSPF it took 250 ms after the failure to ensure the data flow (when using newer hardware and an improved framework response time for FIFR-N lowered to 27 ms).

For the simulations we used four real life network topologies [11][12], each satisfying the requirements for the proper functioning of FIFR-N with 16 to 33 nodes. In order to get enough samples, we applied 1000 different random weight sets associating a cost of 1.0 to 3.0 to each link. In the typical hierarchies we used, more then 95% of the paths were not affected by the node failures. We focused on the remaining 5%. Fig 1 shows the increase of hops in the networks compared to the shortest path when using FIFR-N.

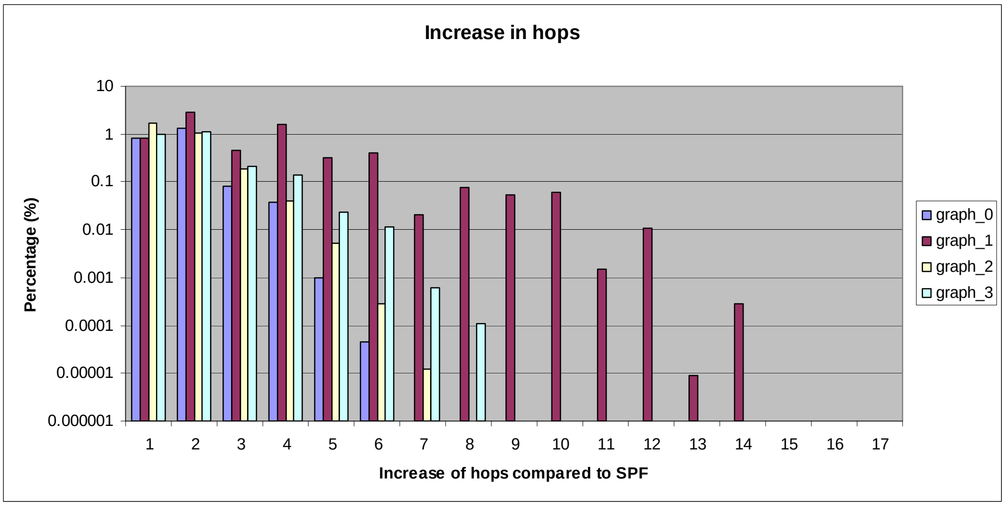


Figure 1: Increase in the number of hops compared to the shortest path when using FIFR-N algorithm (percentages on logarithmic scale)

This proves that FIFR-N gives local response to the failures. Less then 3% of the paths used more nodes to deliver the packages. The total increase of the length of all paths was less then 6%. In 90% of these cases the detour was less then 2 hops longer than the optimal path. This suggests that using FIFR-N to handle transient failures does not increase considerably the load of concerned network components. We must point out that in Fig 1 we can see that in networks based on graph_1 some failures called for more than 8 hop increase of the optimal path. This may result in an unacceptable increase in network delay. However this only occurred in less than 0.05% of the cases, and could be avoided with network design taking FIFR-N in consideration.

5. Conclusion

IP Fast ReRoute (IP FRR) is one of the last missing technological steps for IP to become a full-fledged carrier grade protocol. In this paper we dealt with FIFR-N, one of the promising IPFRR algorithms. We have studied this method by both measurements and extensive simulations. Based on our experience with the fast rerouting framework and the FIFR-N algorithm we are confident that embedding the Failure Inferencing Rerouting mechanism in today's IP routing protocols provides better quality of service and should make IP network suitable for real time applications.

Acknowledgements

We would like to thank our tutor Gábor Rétvári for supporting our work with advice and directions. We also thank Zoltán Tóth and Péter Szilágyi for their valuable work in helping in the development of the fast rerouting testbed and the measurements.

References

[1] SHAND, M., BRYANT, S. *IP Fast Reroute Framework*. IETF Network Working Group, Internet Draft, June 2007.

[2] MOY, J. *RFC 2328 - OSPF Version 2*. http://www.faqs.org/rfcs/rfc2328.html.

[3] MINAS GJOKA, VINAYAK RAM, XIAOWEI YANG. *Evaluation of ip fast reroute proposals*. IEEE Communication Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference, 7-12. January 2007.

[4] ATLAS, A., ZININ, A. *Basic Specification for IP Fast-Reroute: Loop-free Alternates*. IETF Network Working Group, Internet Draft, September 2007.

[5] PREVIDI, S. *IP Fast ReRoute technologies*. Asia Pacific Regional Internet Conference on Operational Technologies (APRICOT), 2. March 2006.

[6] FARKAS, J., ANTAL, C., WESTBERG, L. *Fast Failure Handling in Ethernet Networks*. IEEE International Conference on Communications, 11-15 June 2006.

[7] BRYANT, S., FILSFILS, C., PREVIDI, S., SHAND, M. *IP Fast-reroute Using Tunnels*, IETF Network Working Group, Internet Draft, April 2005.

[8] BRYANT, S., SHAND, M., PREVIDI, S. *IP Fast Reroute Using Not-via Addresses*. IETF Network Working Group, Internet Draft, July 2007.

[9] AMUND KVALBEIN, AUDUN FOSSELIE HANSEN, TARIK CICIC, STEIN GJESSING, OLAV LYSNE *Fast IP Recovery using Multiple Routing Configurations*, IEEE INFOCOM, 2006.

[10] ZIFEI ZHONG, SRIHARI NELAKUDITI, YINZHE YU, SANGHWAN LEE, JUNLING WANG, CHEN NEE CHUAH *Failure Inferencing based Fast Rerouting for Handling Transient Link and Node Failures*, IEEE INFOCOM, 2005.

[11] CHINOY, B., BRAUN, H., W. *The national science foundation network*. Tech. Rep., CAIDA, available online: http://www.caida.org/outreach/papers/1992/nsfn/nsfnet-t1-technology.pdf, Sep 1992.

[12] GARCIA-OSMA, M., L. *TID scenarios for advanced resilience*. Tech. Rep., The NOBEL Project, Work Package 2, Activity A.2.1, Sep 2005.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Color Image Segmentation for Retrieval in Large Image-databases

*Peter Lukáč, *Miroslav Benčo, *Róbert Hudec, *Zuzana Dubcová
*University of Žilina, Faculty of Computer Engineering, Department of Telecommunication and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia,
{peter.lukac, miroslav.benco, robert.hudec, zuzana.dubcova}@fel.uniza.sk

Abstract. Image segmentation is an important part of image processing. The goal of segmentation is to acquire segments that collectively cover the entire image. The goal is finding robust and universal image segmentation technique, which will segment common images in large databases. In this paper, comparison of two good known image segmentation techniques, Perceptual color image segmentation and Graph based segmentation, are presented.

Keywords: Image, segmentation, graph based, *k*-means, algorithm

1. Introduction

In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into more meaningful and easier form to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics [1].

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or features, such a color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics [1].

Image segmentation is possible to classify into full segmentation and partial segmentation. Full image segmentation is used for specific applications because segments image into discrete regions. Partial image segmentation finds parts of the regions in the image. These parts are homogeneous of attribute aspect. For us is interesting the partial image segmentation because we are oriented on the common images. In this paper, two good known image segmentations, Perceptual color image segmentation and Graph based segmentation are presented.

2. Perceptual color image segmentation

The algorithm for color image segmentation is defined in few steps. In the first step is the creation of a perceptual tower from the original image. From [2] is known that human eye reaction is highly dependent on spatial frequencies and high frequencies show the picture more blurred than low frequencies. An attempt to imitate this dependency is the perceptual tower. Using the functions from S-CIELAB toolkit [5], a group of images that are gradually blurred is created. That represents human eye view from increasing distances. First, human is looking on image from long distance (from tower) then closer to it to look for details.

The second step is *k*-means clustering. It is very popular technique for partitioning large data sets with numerical attributes. *K*-means clustering is process of the partitioning or grouping and given set of patterns into disjoint *k* clusters. *K* is positive number and indicates number of searched objects in image [5]. Grouping is created by minimization the sum of squares distance among data and their cancroid in one segment. For the next processing is image converted from matrix into vector form. This vector is characterized by LAB value as well as x, y coordinates of each pixel. The final form of vector is (x, y, L, A, B), where L (Luminance), A (Red-Green) and B (Blue-Yellow). Each pixel is during iterations associated with the cluster nearest to it. Distance is defined as the norm of the vector, found by subtraction of the pixels coordinates (x, y, L, A, B) from the mean coordinates. After one classifying all the pixels the mean of the cluster is recalculated. It is repeated until the mean values are stabilized.

$$delta = \frac{mean(NewCentroid) - mean(Centroid)}{numberofpixels} \tag{1}$$

After *k*-means clustering the image is partitioned into large group of small clusters. The small clusters are in the next step used to form an initial set of larger clusters. They create master segments. It iterates through the group of clusters and for each small cluster finds all neighboring clusters. If the neighboring clusters have a mean in LAB space very closer to the value of current cluster, the clusters are merged into one bigger segment. The similarity of the means in LAB space is defined by using a threshold value. This merging process is repeating by the time, when it cannot find any similarity of the means.

The last step of the algorithm identifies insignificant clusters which their percentage of the total image is below another defined threshold and the corresponding cluster is assigned to the neighboring segments. After this step is the image segmentation finished and the results are master segments.

3. Graph based segmentation

This segmentation is based on Graph theory. Let G = (V, E) be an undirected graph with set of vertices v_i = V and set of edges (v_i, v_j) ϵ E. They present pair of neighboring vertices and each edge (v_i, v_j) ϵ E has a responding weight w (v_i, v_j), which is given as nonnegative dissimilarity measure neighboring vertices v_i and v_j. In the image segmentation, the elements belonging into V are single pixels of image and the weight of edges is represented like a dissimilarity measure between two pixels interconnected by the edge (e.g. the difference in color, motion, intensity, location etc) [3]. There are several techniques to measure quality of the segmentation, but the main idea is that elements in one segment to be similar and elements in the others segments to be dissimilar. So edges between two vertices in the same segment should have low weights and high weights for edges between two vertices in different segments.

The image segmentation algorithm described in [3] starts by the trivial segmentation, where each component contains only one pixel. The pairs of the components based on follow conditions are repeatedly merged:

$$Diff(C_1, C_2) \leq Int(C_1) + T(C_1), \tag{2}$$

$$Diff(C_1, C_2) \leq Int(C_2) + T(C_2), \tag{3}$$

where *Diff* (C_1, C_2) is difference between C_1 and C_2 components, Int(C_1) and Int(C_2) are internal differences of C_1 and C_2 components, T(C_1) and T(C_2) are threshold functions of C_1 and C_2 components.
Treshold function is defined as:

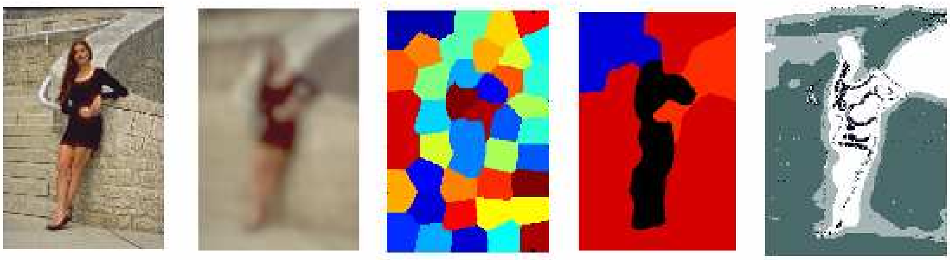
$$T(C) = k / |C|, \tag{4}$$

where $|C|$ presents the size of component, *k* parameter is constant, which manages size of the components. Better evidence for a boundary is required for small components. Larger *k* causes a preference for larger components but is not a minimum component size. Small components are allowed when there is a large difference between neighboring components.

4. Results and experiments

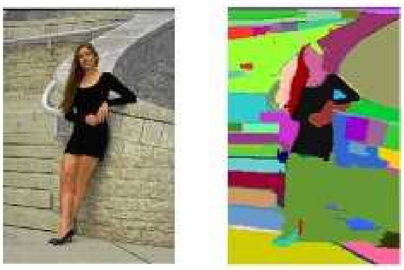
In this part, experiments of both segmentations are presented. For segmentation comparison, the results of perceptual color image segmentation algorithm, graph based segmentation algorithm and segmentation performed by the human subject are used.

The segmentations are applied to image database [4]. Example of experiments results are shown in Fig. 1 and Fig. 2. In Fig. 1(b)–(e) is showed proceeding of the perceptual image segmentation algorithm. It starts with perceptual tower (Fig. 1(b)). The next step, k-means clustering makes 41 smaller segments (Fig. 1(c)) that are gradually merged into larger 4 segments (Fig. 1(d)). In Fig. 1(e) are final 4 segments that represent result of perceptual color image segmentation algorithm. Result of the graph based segmentation algorithm is presented in Fig. 2(b), where *k* parameter is 300 and segmentation output is 83 segments.



(a) (b) (c) (d) (e)

Fig. 1. Progress of perceptual image segmentation algorithm: (a) original image; (b) result of perceptual tower; (c) result of k-means segments; (d) result of merged segments (e) result of final segments after segmentation



(a) (b)

Fig. 2. Result of Graph based segmentation: (a) original image; (b) result

Realized sets of experiments shown the main advantages and disadvantages of used algorithms. Perceptual image segmentation algorithm is not very fast, but it makes less large segments than graph based segmentation algorithm. On the other hand, the graph based segmentation algorithm is much faster, has more accurately edges so it could be used for larger databases. Comparison of results of the segmentation algorithms is in Fig. 3(b)–(d). There is also added one segmentation performed by the human subject (Fig. 3(d)). Experiments was realized in MATLAB.

(a) (b) (c) (d)

Fig. 3. Comparison of results of the segmentations: (a) original image; (b) result of perceptual image segmentation; (c) result of graph based segmentation; (d) result performed by the human subject

5. Conclusion

A comparison of two partial image segmentation techniques is presented in this paper. The perceptual color image segmentation algorithm consists of perceptual tower through *k*-means algorithm and merging segments till final segments. In graph based segmentation algorithm is important to regulate *k*-parameter, which manages size of the components. In the future work, a new algorithm with fusion both algorithms will be developed.

Following the results, advantages both of segmentations are provided. The perceptual color image segmentation algorithm is slow and has larger segments. On the other hand, the graph based segmentation algorithm is fast with small segments and has more accurately edges.

Acknowledgement

This paper has been supported by the VEGA grant agency no. 1/4066/07, “New systems and principles of semantic description and retrieval of multimedia content”.

References

[1] LINDA G. SHAPIRO and GEORGE C. STOCKMAN. *Computer Vision*. New Jersey, Prentice-Hall, 2001.

[2] X. ZHANG and B.A. WANDELL. *A Spatial Extension for CIELAB for Digital Color Image Reproduction*. Proc. Society for Information Display Symposium Technical Digest 27, 1996.

[3] P.F. FELZENSZWALB and D.P. HUTTENLOCHER. *Efficient Graph-Based Image Segmentation*. International Journal of Computer Vision, Vo.59, No.2, 2004.

[4] Corel 1000 image database. http://wang.ist.psu.edu/~jwang/test1.tar [11.03.2009] Available on the internet.

[5] ANGELA CHAU and JEFF WALTERS. *Perceptual color image segmentation*. 2003, http://scien.stanford.edu/class/psych221/projects/03/angichau/index.html [11.03.2009] Available on the internet.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Trends of Industrial Ethernet Appliance in Distributed Control Systems

*Juraj Ľupták
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 2, 01026 Žilina, Slovakia, juraj.luptak@fel.uniza.sk

Abstract. This paper deals with trends of using industrial Ethernet within distributed systems of control. It contains a summarized characterization of industrial Ethernet types and standardization in this area. Main part is oriented on description of Profinet network and security relevant communication with the help of ProfiSafe profile.

Keywords: industry Ethernet, network, ProfiNet, protocol, automation.

1. Introduction

A definite change in the architecture of the control system in the area of industrial automation arised. The progress in the sphere of industrial automation shifted from centralized architecture, characterized by control computers and minicomputers to distributed systems. The intelligence of the control system infiltrates straight into the process. Intelligent sensors and actuators, distributed intelligent terminals and input-output card are used for this purpose.

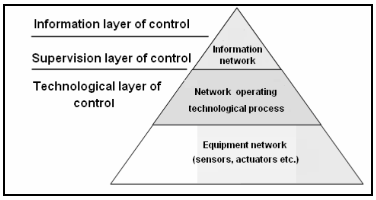


Fig. 1. Hierarchal model of the distributed system of control

Industrial networks also enable remote control and process components adjustment (change of sensor parameters, control and calibration of sensors). With the regard on the development of electronics there is no reason not to provide sensors and actuators with microcontrollers which make data pre-processing possible. The communication is recently one of the fastest developing area of the industrial automation. Communication buses and networks are becoming an important automation element. The distributed system of control can be described by three-level model as displayed in the picture Fig. 1. Layers differ in constitution, severity in communication and the type of communication nodes. The communication runs within all three layers, what is different are protocols used in particular layers. There are following layers:

- the layer on the level of information network
- the layer on the level of automation and control network
- the layer of the network components and sensors

2. Types of industrial Ethernet

According to the IEC 61 158 Standard Ethernet is recently divided into 10 types:

1. EPA (Ethernet for Plant Automaton) – uses TCP/UDP/IP protocols, real time flow control is integrated in data link layer.
2. EtherCAT (Ethernet for Control Automation Technology) – supports devices and application profiles CANopen. Standard CANopen defines communication profiles of devices and application profiles.
3. EtherNet/IP (Ethernet Industrial Protocol) – implements the algorithm of communicating devices clock synchronization in conformity with IEEE 1588 standard through the CIPsync (Control and Information Protocol).
4. Powerlink (Ethernet Powerlink) – is standard, in which the cyclic data transmission is realized by EPL protocol [3].
5. ModbusRTPS (Real Time Publish-Subscribe) – enables simple connection of the Modbus with the ethernet network [5].
6. P-net on IP – determines the interface method of the present P-NET bus with the networks from the UDP/IP protocol without any special requirements on the real time data transmission security by the ethernet network [2].
7. **Profinet** – works on the basis of producer-consumer principle. The producer is the transmittion node and the consumer is the receiving node.
8. SERCOS III (Serial Real-time Communication System) – is digital interface among control units and drives. The real time data transmission is protected by hardver implemented communication channel.
9. TCnet (Time-Critical Control Network) – uses DOMA access method (Deterministic Ordered Multiple Access) for real time data transmission
10. Vnet/IP (Virtual Network Protocol) – dispatches with cyclic mode for technological data transfer and with the transmission of acyclic data by the standard TCP/IP path [2].

2.1. Profinet

The most frequently used industrial network in the past and also at the present is Profibus. At the variant Profibus-DP (provides high speed data interchange among control unit (Master) and slave devices) it becomes the most extended industrial communication network in the world. The following expanded alternatives were PROFIBUS PA, PROFIBUS – FMS (dedicated for communication in higher level, service at this part of PROFIBUS enables the communication in different applications and provides high flexibility).

Moreover, the concept Profinet (part of IEC 61158 Standard) supported by the PNO (PROFIBUS User Organization) becomes one of the most extended industrial communication network [4]. Profibus has twenty application profiles which are gradually extended and implemented into Profinet.

Recently, Profinet represents open communication standard based on the Industrial Ethernet principles. Profinet is convenient for many automation applications. Profinet enables its integration with standard fieldbus systems like Profibus, As-Interface or Inteus without any change in existing devices or additional investments.

Profinet uses several means of communication:

- **non-time-critical data transmission –** transmission of parameters, configuration data and other is non-time-critical and uses standard protocols from computer networks as TCP or UDP and IP.
- **Time-critical data transmission** – uses Real-Time channel (RT) for time-critical data transmission. For partly severe tasks isochronous Real-Time channel (IRT) can be used [1].

With the help of Profinet it is possible to integrate simple distributed input/output equipments and time-critical applications into Ethernet communication as precisely as to integrate distributed automation system into general automated system. Profinet can be divided into two basic types Profinet IO, Profinet CBA.

2.1.1 Profinet IO (Distributed inputs/outputs)

Distributed inputs/outputs communicate together through Profinet IO (Input/Output). The principle of communication is similar to the one used in the Profibus network. Data transmitted from devices are periodically transferred into control device.

Profinet IO is intended for production process automation mostly of discrete type. Profinet IO uses primary cycle – update-time period, for attendance of deterministic access to communication within selected network segment. This cycle is set along with the configuration and is operated by the Profinet IO-Controller (master) [6].

Types of user points		
Profinet IO-Controller	**Profinet IO-Device**	**Profinet IO-Supervisor**
– has the control over the communication in entire Profinet IO network	– participants, subordinate point of network communicating with Profinet IO-Controller equipment	– is equipment with HMI interface or diagnostic network equipment.

Tab. 1. Types of user points

2.1.2 Profinet CBA (Distributed automation)

Profinet CBA (Component Based Automation) interconnects separated components of the entire system. Model of the system which is based on the components, defines components as independent executive units. Distributed automation system designed on the basis of technological modules simplifies modularity of particular devices and set of devices and from these derivated reusable appliance (also in another applications). The reference between is shown in Fig. 2. where all connection are represented by parts of the Ethernet network.

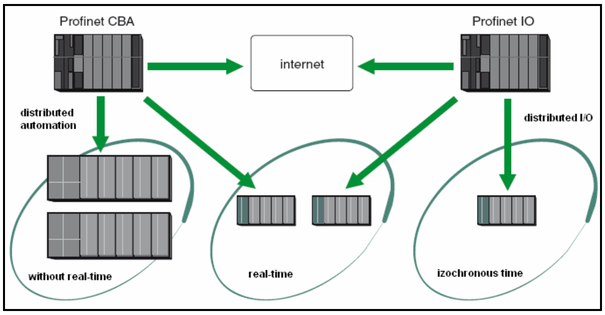


Fig. 2. The reference between Profinet CBA and Profinet IO

3. ProfiSafe

A communication profile ProfiSafe developed by the crew TC3/WG5 "Safety Technology was created to provide security relevant (SR) communication in ProfiNet I/O or Profibus DP/PA network among SR peripheries and control units. It is based on the pooling method among Master and Slave equipments and among Slave equipments and SR proceeding implementations for eliminating errors made at the process of transmission [7].

ProfiSafe has recently two operating modes:

- version V1 – for SR communication in Profibus DP/PA network types

- version V2 – for SR communication in ProfiNet I/O or Profibus DP/PA network types

4. Areas of Profinet networks application

Profinet CBA has a wide range of utilization within distributed automation systems. It is relevant for intelligent operating equipments with alternate programmable function, like operating units.

Profinet IO has a wide range exploration for kinetic systems, it is available in isochronous real time variant - IRT.

Within the automation of processes with 5 to 10 ms cycle time (update-time) and with in advance listed number of RT users a real time variant – RT is used.

5. Conclusion

In industrial networks Fieldbus standards are widespread. They are open and have wider range of possible appliance in the spheres like industrial production, operating of the process of production, control of technologies etc.

But the development in the area of industrial networks is still in process and new standard is finding his place. It is as open as the previous one, and so far, it has the wides range of possibilities of appliance in the industrial zones that means in distributed systems of control.

It is required in other various spheres than in the area of decentral periphery and it is expanding also in the areas of high-speed driver communication. Profinet offers a new communication platform which by available tools integrates not only hierarchal communication in concerns, but also different communication systems together on the basis of industrial Ethernet.

It is obvious, that Profinet is becoming a standard, easily available choice with continual development and his appliance in many divers projects is absolutely worth considering.

Acknowledgement

This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0040/08 “Mathematic-graphical modelling of safety attributes of safety-critical control systems”.

References

[1] ZEZULKA, F., HYNČICA, O. Průmyslový Ethernet IV, Automa, 2007, roč. 13, č. 10, s. 57–60.

[2] ZEZULKA F., HYNČICA, O. Průmyslový Ethernet VII, Automa, 2008, roč. 14, č. 2, s. 26–29.

[3] ZEZULKA F., HYNČICA, O. Průmyslový Ethernet VIII, Automa, 2008, roč. 14, č. 5, s. 62–66.

[4] BÉLAI, I. Komunikácia v priemyselnej automatizácii. *AT&P*, 2007, roč. 7, č 3, s. 64–66.

[5] BÉLAI, I. Komunikácia v priemyselnej automatizácii. *AT&P*, 2007, roč. 7, č 6, s. 81–82.

[6] DRAHOŠ, P., Gabriel, J.: Komunikačný systém Profinet IO. Automa, 2006, roč. 12, č. 7, s. 2–5.

[7] FRANEKOVÁ, M., KÁLLAY, F., PENIAK, P., VESTENICKÝ, P. Komunikačná bezpečnosť priemyselných sietí. Edis, ŽU Žilina, 2007. ISBN 978-80-8070-715-6

[8] ZELENKA, J., ROFÁR, J. The ratio criterion influence in safety and availability of communication system. Zvůle, 2008, s 289-292. ISBN 978-80-214-3709-8

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Beacons Position Influence of Node Localization

*Martin Matula, *Jozef Dubovan
*University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 010 26 Žilina, Slovakia, {matula, dubovan}@fel.uniza.sk

Abstract. This paper is deal in accurate position estimation in mobile ad-hoc networks by RSS method in dependency of beacons distribution. The aim was to set the position estimation for different types of environment with different fading's influences. The precision of position estimation has been changed in compliance with network conditions, coverage area, type of channel, the signal strength and the number of used beacons. However, mobile ad-hoc network are changing dynamically, so also situation (strength) of wireless channel, the number and position of nodes and beacons in network, so as the total coverage area are changing. We organized a graphical representation of estimation by RMSE map. RSS method is not the most precision method in consequence of these influences, but it is one of the evaluable from a hardware fastidiousness devices. It is necessary to use statistical means to tune up the position estimation in these very demanding conditions.

Keywords: Position estimation, RSS method, ad-hoc networks, wireless channel fading, RMSE maps, overlapping rings, histogram, CI, direct method.

1. Introduction

Mobile ad-hoc networks are specific in their dynamic change, what can be interpreted by all time changes of nodes number and their distribution in network. This change is demonstrative on wireless transmission channel by many types of fading's (multipath propagation, shadowing fading and etc.). We create an algorithm of position estimation based on RSS method for its simple application. However, this method is not the most precision, so some statistical methods are used in that algorithm to specify the estimation influenced by fading.

1.1. Ad-hoc network

We can say about ad-hoc networks, that they are autonomic mobile networks, dynamically changing and composed of equal mobile nodes without infrastructure. Nodes communicate between themselves by single-hop or multi-hop. This type of network has fast changing topology, because nodes could move. Communication standards in ad-hoc networks utilize an unlicensed frequency band as Bluetooth, ZigBee and the main thing is Wi-Fi (802.11x).

1.2. RSS localization method

When we are setting the localization, it is good to realize some factors like the assistance requirement of beacons for position estimation of nodes in ad-hoc network. These beacons are nodes, which contain GPS module and that's the way how they can set their localization. This value of position is considered to be absolute from the view of localization. We can obtain a relative position of node by assistance of this information.

For localization estimation is used direct method. This method is based on RSS received from nodes. DM required 3 distances from 3 nodes. If we use not DM but overlapping rings method it is not necessary 3 distances obtain.

2. Algorithm

The algorithm is based on information about loss in receiving signal strength from beacon to node, and on information about statistical methods.

2.1. Statistical tools

Three statistical methods (a histogram, a CI, and histogram from histogram) use the direct method (DM) of localization statement by three distances from three beacons.

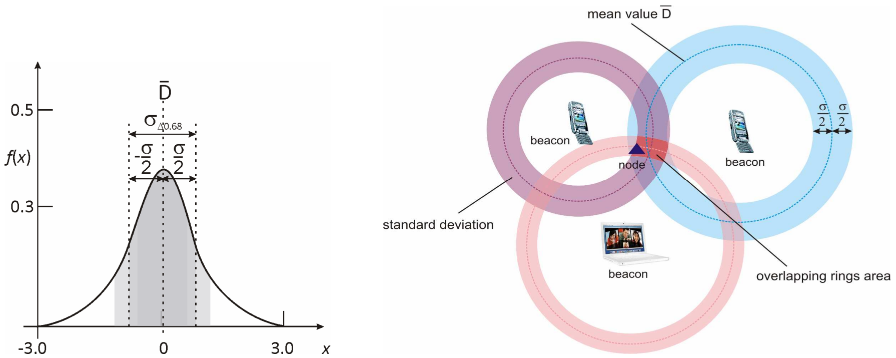
- Probability density function of all estimation was used for estimation via histogram maximum from all combinations of beacons.

- It was used linear approximation for confidence interval (CI) determination

$$h = p_1 x + p_2 , \qquad (1)$$

  and it requires at least two points for calculation. In our case we decided to set the confidence level 0.4. We use middle values of obtained CI (separately for *x*-coordinate and separately *y*-coordinate) for requirement in case of set probable localization. If that coordinates, find out by direct method, are out of distance (extent), it won´t be included to CI calculation. In case of possibility, when only one point left for that calculation, it is considered as possible point of position with confidential interval <0,0> and <0,0>.
  When all of obtained values are eliminated, we use point with coordinates [0,0] and also CI <0,0> and <0,0>.

- The third statistical method is histogram creation from all histograms (because we can realize PDF from one combinations of beacons – one realization) of position estimations composed from particular realization.



A) B)

Fig. 1. The figure A) the gauss curve expressed distribution of distance estimation between node an beacon and the possible mistake of measurement. The figure B) the field (area) of possible occurrence, where the node could be located. It is limited by beacon realization of beacon rings (overlapping).

Overlapping rings metrics (fig.1) is next statistical method applied on RSS methods. Localization estimation express distances between beacon and node by rings (overlapping) represented standard deviation.

We can get a field (area) with the highest value of rings by approximation of particular overlapping and then we can present the middle from that. This center expresses the estimation in this metrics.

3. Experimental

The position estimation depends not only on the network wide but, as the simulation show, on the beacons arrangement. For demonstration we use four networks beacons arrangement:

- in the quad,
- in the triangle,
- in the ring,
- random.

For simulation we set network wide 30m and transmitted power at 2400 MHz was -30dBm. For evaluation we use RMSE values arranged to the 2D map of the space.

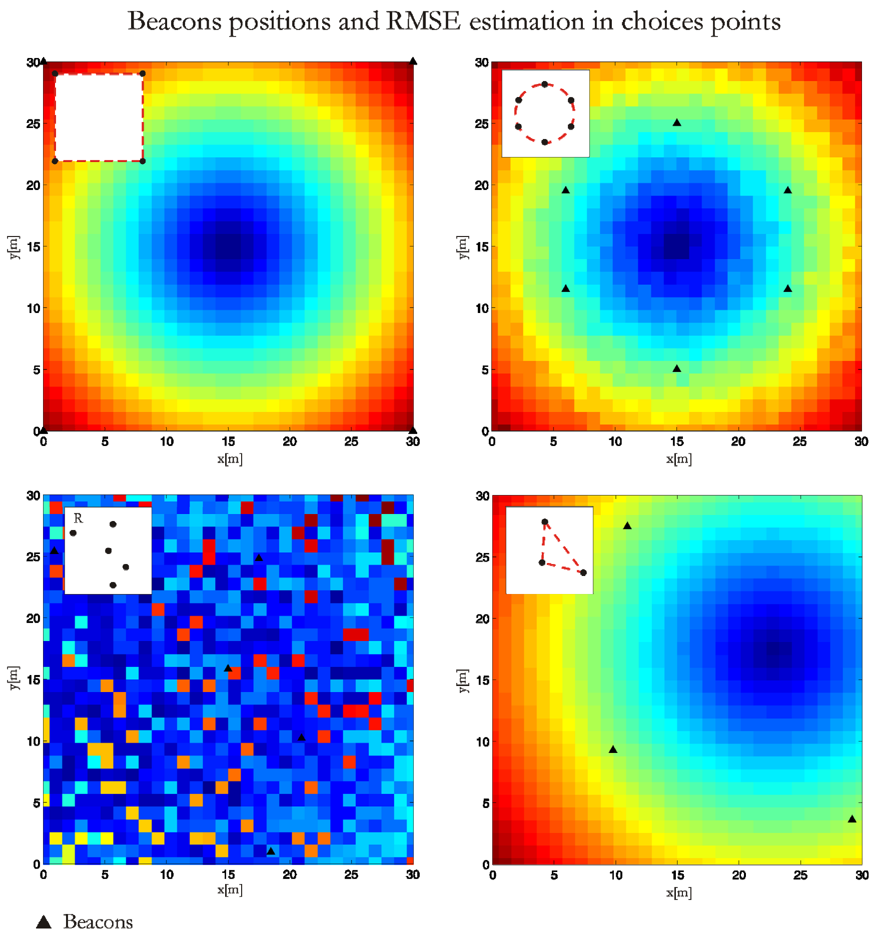


Fig. 2. Figures show RMSE dependency on the beacons topology for DM. The color tone of square element represents the RMSE value in choice point.

4. Conclusion

As we can see from these results the position estimation depends on network wideness as well as on topological arrangement of each beacons. Obtained results must be understudied as a map, which evaluate the accuracy of estimation in actually arrangement of beacons and its parameters (power, ...).

Acknowledgement

This paper was prepared with the support of scientific-agency grant, which belongs to Ministry of Education of Slovak Republic VEGA-1/4065/07.

References

[1] FONTAN, P. F., ESPINEIRA M. P. *Modelling the Wireless Propagation Channel: A Simulation Approach with MATLAB.* 1st ed. A John Wiley and Sons, Ltd., 2008

[2] BASAGNI, S., CONTI, M., GIORDANO, S. – STOJMENOVIC, I. *Mobile Ad Hoc Networking*. John Wiley & Sons Ltd., 2004

[3] HE, T. – HUANG, Ch. – BLUM, M. B. – STANKOVIC, A. J – ABDELZAHER, T. *Range-free localization schemes for large scale sensor networks*. International Conference on Mobile Computing and Networking, Proceedings of the 9th annual international conference on Mobile computing and networking**,** San Diego, CA, USA, 2003

[4] KUPPER, A. *Location-based Services*. John Wiley & Sons Ltd., 2005

[5] ILYAS, M., MAHGOUB, I. *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems.* CRC Press, 2005

[6] WALKE, H. B., MANGOLD, S., BERCMANN, L. *IEEE 802 Wireless Systems: Protocols, Multi-hop Mesh/Relaying, Performance and Spectrum Coexistence.* John Wiley & Sons Ltd., 2006

[7] SANTI, P.: *Topology control in wireless ad hoc and sensor networks,* John Wiley & Sons Ltd., 2005

[8] HEKMAT, R. *Ad-hoc networks. Fundamental properties and network topologies,* Springer, 2006

[9] MELOUN, M., MILITKY, J. *Statistická analýza experimentálních dat*. Academia, 2004

[10] DOBOS, L., DUHA, J., MARCHEVSKY, S., WIESER, V. *Mobilné rádiové siete*, EDIS, 2002

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Computation Methods for Travelling Salesman Problem

Miroslav Michalco, Nadežda Čuboňová
University of Žilina, Faculty of Mechanical Engineering, Department of Automation and Production Systems, Univerzitná 1, 01026 Žilina, Slovakia,
{miroslav.michalco, nadezda.cubonova}@fstroj.uniza.sk

Abstract. This paper deals about a very complex combinatorial optimization problem named “Travelling Salesman Problem”. In this article you can find simple description of the problem and several strategies used to solve it. These strategies are divided into exact and heuristics methods and they are used for different variations and sizes of the problem.

Keywords: Travelling Salesman Problem, algorithm, optimum, city

1. Introduction

The travelling salesman problem (TSP) is one of the most widely studied combinatorial optimization problems, with links to many fields of pure and applied mathematics like graph theory and integer programming (Lawler et al. 1985). Simply stated, a salesman has to visit TV cities, visiting each city exactly once and return back to the starting city with minimum cost. Search space is then permutation of n cities and every permutation is a possible solution and a permutation with shortest path is the optimal solution. Size of the search space is then $n!$.

TSP is relatively old problem: firstly described by Euler in 1759. Name “Travelling salesman” was first used in 1932. A lot of techniques were designed for TSP solving in last years, using different methods and strategies.

2. Computing a solution

The traditional lines of attacking for the NP-hard problems are the following:

- Devising algorithms for finding exact solutions (they will work reasonably fast only for relatively small problem sizes).
- Devising "suboptimal" or heuristic algorithms, i.e., algorithms that deliver either seemingly or probably good solutions, but which could not be proved to be optimal.
- Finding special cases for the problem ("subproblems") for which either better or exact heuristics are possible.

2.1. Exact algorithms

The most direct solution would be to try all permutations (ordered combinations) and see which one is cheapest (using brute force search). The running time for this approach lies within a polynomial factor of $O(n!)$, the factorial of the number of cities, so this solution becomes impractical even for only 20 cities. One of the earliest applications of dynamic programming is an algorithm that solves the problem in time $O(n^2 2^n)$[10]

The dynamic programming solution requires exponential space. Using inclusion–exclusion, the problem can be solved in time within a polynomial factor of 2^n and polynomial space.

Improving these time bounds seems to be difficult. For example, it is an open problem if there exists an exact algorithm for TSP that runs in time $O(1.9999^n)$

Other approaches include:

- Various branch-and-bound algorithms, which can be used to process TSPs containing 40-60 cities.
- Progressive improvement algorithms which use techniques reminiscent of linear programming. Works well for up to 200 cities.
- Implementations of branch-and-bound and problem-specific cut generation; this is the method of choice for solving large instances. This approach holds the current record, solving an instance with 85,900 cities.

An exact solution for 15,112 German towns from TSPLIB was found in 2001 using the cutting-plane method proposed by George Dantzig, Ray Fulkerson, and Selmer Johnson in 1954, based on linear programming. The computations were performed on a network of 110 processors located at Rice University and Princeton University. The total computation time was equivalent to 22.6 years on a single 500 MHz Alpha processor. In May 2004, the travelling salesman problem of visiting all 24,978 towns in Sweden was solved: a tour of length approximately 72,500 kilometers was found and it was proven that no shorter tour exists.

In March 2005, the travelling salesman problem of visiting all 33,810 points in a circuit board was solved using *Concorde TSP Solver*: a tour of length 66,048,945 units was found and it was proven that no shorter tour exists. The computation took approximately 15.7 CPU years. In April 2006 an instance with 85,900 points was solved using *Concorde TSP Solver*, taking over 136 CPU years.

2.2. Heuristic and approximation algorithms

Various *heuristics* and *approximation algorithms*, which quickly yield good solutions have been devised. Modern methods can find solutions for extremely large problems (millions of cities) within a reasonable time which are with a high probability just 2-3% away from the optimal solution. Several categories of heuristics are recognized:

2.2.1. Constructive heuristics

The nearest neighbour (NN) algorithm (or so-called greedy algorithm) lets the salesman choose the nearest unvisited city as his next move. This algorithm quickly yields an effectively short route. For N cities randomly distributed on a plane, the algorithm averagely yields length = 1.25 * exact_shortest_length.

However, there exist many specially arranged city distributions which make the NN algorithm gives the worst route. This is true for both asymmetric and symmetric TSPs.

Recently a new constructive heuristic, Match Twice and Stitch (MTS) has been proposed. MTS has been shown to empirically outperform all existing tour construction heuristics. MTS performs two sequential matchings, where the second matching is executed after deleting all the edges of the first matching, to yield a set of cycles. The cycles are then stitched to produce the final tour.

2.2.2. Iterative improvement

- **Pairwise exchange** or **Lin-Kernighan** heuristics.

The pairwise exchange or '2-opt' technique involves iteratively removing two edges and replacing these with two different edges that reconnect the fragments created by edge removal into a new and shorter tour. This is a special case of the k-opt method. Note that the label 'Lin-Kernighan' is an often heard misnomer for 2-opt. Lin-Kernighan is actually a more general method.

- ***k*-opt heuristic**

Take a given tour and delete k mutually disjoint edges. Reassemble the remaining fragments into a tour, leaving no disjoint subtours (that is, don't connect a fragment's endpoints together). This in effect simplifies the TSP under consideration into a much simpler problem. Each fragment endpoint can be connected to $2k - 2$ other possibilities: of $2k$ total fragment endpoints available, the two endpoints of the fragment under consideration are disallowed. Such a constrained $2k$-city TSP can then be solved with brute force methods to find the least-cost recombination of the original fragments. The k-opt technique is a special case of the V-opt or variable-opt technique. The most popular of the k-opt methods are 3-opt, and these were introduced by Shen Lin of Bell Labs in 1965. There is a special case of 3-opt where the edges are not disjoint (two of the edges are adjacent to one another). In practice, it is often possible to achieve substantial improvement over 2-opt without the combinatorial cost of the general 3-opt by restricting the 3-changes to this special subset where two of the removed edges are adjacent. This so-called two-and-a-half-opt typically falls roughly midway between 2-opt and 3-opt, both in terms of the quality of tours achieved and the time required to achieve those tours.

- ***V'*-opt heuristic**

The variable-opt method is related to, and a generalization of the k-opt method. Whereas the k-opt methods remove a fixed number (k) of edges from the original tour, the variable-opt methods do not fix the size of the edge set to remove. Instead they grow the set as the search process continues. The best known method in this family is the Lin-Kernighan method (mentioned above as a misnomer for 2-opt). Shen Lin and Brian Kernighan first published their method in 1972, and it was the most reliable heuristic for solving travelling salesman problems for nearly two decades. More advanced variable-opt methods were developed at Bell Labs in the late 1980s by David Johnson and his research team. These methods (sometimes called Lin-Kernighan-Johnson) build on the Lin-Kernighan method, adding ideas from tabu search and evolutionary computing. The basic Lin-Kernighan technique gives results that are guaranteed to be at least 3-opt. The Lin-Kernighan-Johnson methods compute a Lin-Kernighan tour, and then perturb the tour by what has been described as a mutation that removes at least four edges and reconnecting the tour in a different way, then v-opting the new tour. The mutation is often enough to move the tour from the local minimum identified by Lin-Kernighan. V-opt methods are widely considered the most powerful heuristics for the problem, and are able to address special cases, such as the Hamilton Cycle Problem and other non-metric TSPs that other heuristics fail on. For many years Lin-Kernighan-Johnson had identified optimal solutions for all TSPs where an optimal solution was known and had identified the best known solutions for all other TSPs on which the method had been tried.

2.2.3. Randomised improvement

- Optimised Markov chain algorithms which use local searching heuristic sub-algorithms can find a route extremely close to the optimal route for 700 to 800 cities.
- Random path change algorithms are currently the state-of-the-art search algorithms and work up to 100,000 cities. The concept is quite simple: Choose a random path,

choose four nearby points, swap their ways to create a new random path, while in parallel decreasing the upper bound of the path length. If repeated until a certain number of trials of random path changes fail due to the upper bound, one has found a local minimum with high probability, and further it is a global minimum with high probability (where high means that the rest probability decreases exponentially in the size of the problem - thus for 10,000 or more nodes, the chances of failure is negligible).

TSP is a touchstone for many general heuristics devised for combinatorial optimization such as *genetic algorithms, simulated annealing, Tabu search, ant colony optimization,* and *the cross entropy method*.

3. Conclusion

There are many different applications and different TSP variations in the real world, therefore it is necessary to know, what kind of problem we are going to solve. First we have to study it, think about the results we demand, if we need the optimal solution or we need solution close to the optimal one, how close and how much time do we want to spend on computation. All these factors can lead us to different optimization techniques. Some of these techniques are universal and can be used for several tasks or several modifications; some are very special, designed for only one task. Genetic algorithms are used for a modification of TSP in a doctoral thesis *"Implementation of genetic algorithms into optimization of technological processes"*. In this thesis, GA based on TSP task are used as an optimization tool for improving the rapid move efficiency and reducing the length of toolpaths. This work deals with common problems of GA's like comparison of several techniques and searching for method, good enough for our purpose.

References

[1] ZBIGNIEW, M., *Genetic algorithms + data structures = evolution programs.* Berlin: SPRINGER, 1999,

[2] http://www.talkorigins.org/faqs/genalg/genalg.html

[3] http://en.wikipedia.org/wiki/Travelling_salesman_problem

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Video Camera Choice Criteria for the Following Image Processing

*Tibor Moravčík
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 1, 010 26 Žilina, Slovak Republic, tibor.moravcik@fel.uniza.sk,

Abstract. This paper focuses on digital video cameras used to capture the video signal. In the first part, image sensor as the main component of common camera, is described. Next, some features and parameters which are important for the appropriate camera choice from the point of view of image quality and data representation are introduced. At the end, the programmable camera platform for image processing is presented.

Keywords: Video camera, CCD, pixel, image processing

1. Introduction

In a film camera, the light is focused through the camera lens in varying levels of intensity. Then, the light hits a chemically-treated piece of film which reacts to the different intensities of light. The same basic principle is true for digital video cameras; however, instead of light hitting a strip of chemically-treated film, the light hits the special silicon chip that measures light photons on a panel of light-sensitive diodes and then converts that light information into a digital signal that represent image. To process and transfer the colour information, camera measures three separate light wavelengths of red, green and blue. Nowadays, the camera is used not only for a home video but also in industry applications, e.g. to capture and record the video for security purpose or for the following signal processing like motion detection or object recognition and tracing.

2. Image Sensor

Image sensors are usually based on one of two main technologies, which are CCD (charge coupled device) and CMOS (complementary metal oxide semiconductor). A CCD chip is an array of small electronic capacitors. These capacitors are charged by the electrons generated by the light. In fact, each light element (photon), reaching the CCD array's atoms, displaces some electrons which are providing a current source [1]. These current sources are localised in small delimited areas, called pixels. Digital colour cameras generally use a Bayer mask over the CCD. Each square of four pixels has one filtered red, one blue, and two green (the human eye is more sensitive to green than either red or blue). Three-chip cameras however posses one such chip for each prime colour red, green and blue. This results in an approximately 20 percent better image quality.

CCD and CMOS image sensors are two different technologies for capturing images digitally. Each has unique strengths and weaknesses giving advantages in different applications. [3].

CCD	CMOS
- Expensive.	- Cheap (1/3 of the CCD).
- Requires more power.	- Requires more light.
- Separate chips for processing.	- Integrated with circuits doing signal processing.
- Higher quality than CMOS.	- Slower due to more noise.

3. Properties of Cameras

3.1. Monochrome versus Colour

Monochrome is not used in many application of camera, even the cheapest TVs today feature colour. Monochrome is still common in the CCTV universe, but a shift is under way there toward both colour and higher resolutions. Then, there is machine vision which stills largely a monochrome. Imaging in industrial production is typically used for such tasks as quality control. The general goal is to identify a feature or features on an object, where are requirements like edge detection and blob (lighter vs. darker) analysis often do not require colour [5]. These processes typically work to maximize the data relevant to the specific task at hand, while minimizing unnecessary background data. Colour generally generates a higher data volume, requires more data processing, and has reduced sensitivity. There are a variety of additional colour issues that come in to play, including white balance and lighting, that affect the performance of a colour camera but have less impact on the monochrome camera [3].

3.2. Pixel Resolution and Size of a CCD Sensor

Each CCD sensor has a vertical and horizontal array of pixels that determine the overall resolution achievable. The CCD sensor size is reported as the diagonal measurement of the sensor, typically referred to as 1/2”, 1/3”, 2/3”, 1”, or higher chip size formats. Fundamentally, increasing pixel number for a given sensor size results in higher image resolution. Aspect ratio is relation between horizontal and vertical resolution, for example resolution 640x480 has 640 horizontal pixels and 480 vertical pixels with aspect ratio 4:3. Typically used resolutions are for example 352x288, 512x492, 512x582, 640x480 and 768x582. Pixels are usually square but can sometimes be rectangular what define pixel aspect ratio.

3.3. Dynamic Range and Sensitivity

The size of the pixel itself can affect overall dynamic range. Reducing pixel size can significantly reduce the well capacity of a pixel what is the limit on the number of electrons a pixel can hold. Typically, larger pixels have larger well capacities. The dynamic range (usually expressed in EV values) of the CCD determines the range over which a camera can record simultaneously very low light signals and very bright signals and is determined by the ratio of signal to noise [4]. Sensitivity (ISO) is the number indicating digital camera sensors sensitivity to light. The higher the sensitivity, the less light is needed to make an exposure, but with higher amount of noise.

3.4. Lens and Iris

The lens of a video camera gathers and concentrates light that falls from objects onto an image sensor in the camera. Between the lens and the image sensor is the iris, which regulates how much light falls onto the chip. An almost closed iris allows more objects at different distances to be in focus while with an almost fully opened iris only few objects close together are in focus.

3.5. A/D Converter and Bit Depth

Every CCD camera uses an A/D converter to transform the variable charges in the CCD into the binary data that represents pixels. The higher the bit depth of the A/D converter (10 bit, 12 bit, 16 bit, etc), the more accurate the analogue to digital conversion will be because the analogue signal will be sampled more frequently. This will make it easier to distinguish differences in intensity values, even if they are very close. The bit depth of the A/D converter can limit the dynamic range of the CCD chip itself if it has a smaller dynamic range than the CCD chip.

3.6. Video Compression

For image transfer or its record in digital form is appropriate reducing the quantify of data represent digital video. Some forms of data compression are lossless (for example Run-Length Encoding, RLE). This means that when the data is decompressed, the result is a bit-for-bit match with the original. Most video compression is loss and it operates on the premise that much of the data present before compression is not necessary for achieving good perceptual quality [1]. Usually used compression algorithms are Motion-JPEG and MPEG. By MJPEG is every frame of video signal compressed with the same algorithm JPEG what provide eases video editing, higher quality, but is less effective. MPEG-4 use interframe prediction for achieve higher data compression ratio and is preferred for live capture and application with low throughput. Video compression is a trade off between disk space, video quality, and the cost of hardware required to decompress the video in a reasonable time.

3.7. Frame Rate and Image Refresh Rate

One feature of the CCD camera often overlooked in discussion of camera specifications is the image refresh rate. Several factors related to signal processing, camera electronics, and determine the frame rate and image refresh rate, commonly referred to as Frames per Second or FPS. Frame rate is the frequency at which an imaging device produces consecutive images in video sequence. Low refresh rate values mean that there will be a lag between the image the camera captures and what is displayed on the computer screen. A slow refresh rate will increase the amount of time it takes to focus on an image. A minimum frame rate of 3 FPS is required for ease of use.

4. IR Illuminating for "Night Vision"

The infrared (IR) spectrum with wavelength from 700 nm to 1200 nm is for human invisible and is nearly 1,000 times as large as the visible spectrum (350 – 700 nm). Most CCD and CMOS cameras have excellent near-IR sensitivity [1]. Warm objects emit NIR light, but only with temperature about 500°C, and colder objects (like human) produce IR wavelengths higher then 1200nm, from this reason we need to create this NIR light. Infrared Cameras (Fig. 1a) have infrared LED’s positioned around the outer edges of the camera lens, which illuminate area with NIR light and gives the camera its "Night Vision". Infrared Cameras have the capability to capture quality video in low light and acceptable video in no light (0 Lux) areas. Example of low light scene is on Fig. 2b and the same scene with IR Fig. 2c.

a) b) c)

Fig. 1 a) IR illuminated camera, b) Scene without IR, c) Scene with IR

5. Application Camera CMUCAM3 Embedded Colour Vision Platform

This platform from Active Robots Limited (UK) is self-contained unit that perform image processing (Fig. 2b) on board and is able to produce control instructions or a measurement result without the need for additional hardware (Fig. 2a). The CMUcam3 provides a flexible and easy to use open source development environment that complements a low cost hardware platform [7]. The CMUcam3 is based on fully programmable embedded computer vision sensor. The main processor is connected to an Omnivision CMOS camera sensor module. For CMUcam3 can be developed custom C code and executables flashed can be flashed onto the board. The concept is more flexible and cost-effective than a PC-based system solution. Some application is object recognition and tracking, robotics or surveillance.

a) b)

Fig. 2 a) camera platform board, b) Example of application (Face detection)

6. Conclusion

In principle, the choice of an appropriate video camera for intended industrial application depends on few basic criteria, which are: the distance from sensing objects, objects character and movement (lens, shooter, frame rate), required image quality (monochrome – colour, resolution and compression), captured environment (light conditions, outer or inter use) and the price, which plays an important role in the decision. Let's assume an application of camera in car driver capturing for his vigilance monitoring. The distance between the driver and the dashboard is short and relatively does not vary; therefore the use of fixed focus is preferred. Frame rate must be sufficient to capture eyes blink (15fps or higher), compression (e. g. M-JPEG, MPEG-4) is chosen depending on the method of following image processing. Monochrome camera is more suitable for this kind of application. The image resolution should be at least 512x492 or higher, light condition can vary so that the auto iris and IR illumination for low light capture are required. Image processing is realized on PC or on another machine vision platform connected to it through the analog or digital interface.

References

[1] LOVEČEK, T., NAGY, P. *Kamerové bezpečnostné systémy*. EDIS Žilina, 2008, ISBN 978-80-8070-893-1

[2] CCD vs. CMOS http://www.dalsa.com/corp/markets/CCD_vs_CMOS.aspx (5.3.2009)

[3] Practical Color Machine Vision http://articles.directorym.net/Practical_Color_Machine_Vision-a906592.html (5.3.2009)

[4] HUDEC, R. *Snímanie obrazu.* EDIS Žilina, 2008, ISBN 978-80-8070-880-1

[5] SONKA, M., HLAVAC, V., Boyle, R. *Image Processing, Analysis and Machine Vision*. Thomson, 2008, ISBN 978-0-495-24428-7

[6] Face Detection in the Near-IR Spectrum www.cse.unr.edu/~bebis/DowdallSPIE03.pdf (3.3.2009)

[7] CMUcam Cameras http://www.active-robots.com/products/sensors/cmucam.shtml (5.3.2009)

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Security Technologies in Wireless Networks

*Tomáš Ondrašina
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 1, 010 26 Žilina, Slovak Republic, {tomas.ondrasina}@fel.uniza.sk

Abstract. Wireless networks use a variety of cryptographic techniques such as encryption and authentication to provide barriers for infiltrations. This paper deals with the basic security principles in WLANs (Wireless Local Area Networks). The main focus of this work is on WEP (Wired Equivalent Privacy), WPA-WPA2 (Wi-Fi Protected Access) and AES (Advanced Encryption Standard). VPN (Virtual Private Network) is described in last part of work as another way to security in wireless networks.

Keywords: Wireless, security, WLAN, advanced encryption standard, authentication.

1. Introduction

The growing demand and implementation of all types of wireless networks in homes, companies and industry (safety-related applications) is an accomplished fact. This type of networks offers a wide range of advantages over traditional cabled networks. Ease of use, wide coverage, mobility and simple extensions are just a few. It is to these characteristics that wireless networks owe the boom they are experiencing at the moment.

However, these advantages have a flipside in the form of security problems which, although they are rising to the surface, most administrators do not bear in mind. At this stage, it is clear and proven that wireless networks are intrinsically insecure and that a greater effort is needed to secure them than cabled networks. Security solutions must be seamlessly integrated, more transparent, flexible and manageable.

Security usually refers to ensuring that users can perform only the tasks that they are authorized to do and can obtain only the information that they are authorized to have. Security must ensure that users cannot cause damage to the data, applications or operating environment of a system. The word security involves protection against malicious attacks. Security also involves controlling the effects of errors and equipment failures. Anything that can protect against a wireless attack will probably prevent other types of trouble as well.

The balance between allowing authorized access and preventing unauthorized access is illustrated in Tab. 1.

Transparent Access	Security
• Connectivity	• Authentication
• Performance	• Authorization
• Ease of Use	• Accounting
• Manageability	• Assurance
• Availability	• Confidentiality
	• Data Integrity

Tab. 1. Balance between transparent access and security access.

2. Basic WLAN Security Technologies

Security Requirements for WLANs can be summarized by the following Tab. 2.

First generation security
▪ SSID *(Service Set Identifier)* ▪ MAC *(Media Access Control)* ▪ Static 40 or 128-bit WEP *(Wired Equivalent Privacy)*
Second generation security
▪ Centralized used-based authentication ▪ Dynamic 128-bit WEP ▪ VPN *(Virtual Private Network)* ▪ ACL *(Access Control Lists)*
Leading edge security
▪ TKIP *(Temporal Key Integrity Protocol)* ▪ MIC *(Message Integrity Check)* ▪ AES *(Advanced Encryption Standard)* ▪ Rogue AP detection

Tab. 2. Security requirements for WLANs.

The main characteristics of the protocols used in wireless networks are below in Tab. 3.

	WEP	**WPA**	**WPA2**
Encryption	RC4	RC4	AES
Key length	40 bits	128bits enc. / 64bits auth.	128 bits
Key duration	24-bit IV	48-bit IV	48-bit IV
Data integrity	CRC-32	Michael	CCM
Header integrity	None	Michael	CCM
Key control	None	EAP	EAP

Tab. 3. Basic security technologies.

You can see the gradual tightening of the encryption protocols up to WPA2, which finally changes RC4 as the protocol used to implement AES. It is also clear that an effort has been made to reinforce the integrity of the datagram, at both data level and header level.

2.1. First generation wireless security

Security was not a big concern for early WLANs. The equipment was proprietary, expensive and hard to find. Many WLANs used the Service Set Identifier (SSID) as a basic form of security. Some WLANs controlled access by entering the media access control (MAC) address of each client into the wireless access points. Neither option was secure, since wireless sniffing could reveal both valid MAC addresses and the SSID [1].

The SSID is a 1 to 32-character American Standard Code for Information Interchange (ASCII) string that can be entered on the clients and access points. Most access points have options like "SSID broadcast" and "Allow any SSID". SSIDs should not be considered a security feature.

MAC based authentication is not specified in the 802.11 specifications. However, many vendors have implemented MAC based authentication. Most vendors simply require each access point to have a list of valid MAC addresses. MAC addresses are not a real security mechanism, since all MAC addresses are unencrypted when transmitted. In certain cases, MAC address authentication can supplement security features, but this should never be the primary method of providing wireless security.

2.2. Wired Equivalent Privacy (WEP)

The IEEE 802.11 standard includes WEP to protect authorized users of a WLAN from casual eavesdropping. The IEEE 802.11 WEP standard specified a 40-bit key, so that WEP could be exported and used worldwide. Most vendors have extended WEP to 128 bits or more. When using WEP, both the wireless client and the access point must have a matching WEP key. WEP is based upon an existing and familiar encryption type, Rivest Cipher 4 (RC4).

The IEEE 802.11 standard provides two schemes for defining the WEP keys to be used on a WLAN [2].

- In the first scheme, a set of up to four default keys are shared by all stations, including clients and access points, in a wireless subsystem.
- In the second scheme, each client establishes a key mapping relationship with another station. This is a more secure form of operation, because fewer stations have the keys.

2.3. Wi-Fi Protected Access (WPA & WPA2)

WPA includes mechanisms from the emerging 802.11i standard for improving wireless data encryption. WPA is also called Simple Secure Networking (SSN) [3]. WPA has TKIP, which uses the same algorithm as WEP, but it constructs keys in a different way.

TKIP is also called WEP Key hashing and was initially referred to as WEP2. TKIP is a temporary solution that fixes the key reuse problem of WEP as illustrated in Fig. 1. WEP periodically uses the same key to encrypt data. The TKIP process begins with a 128-bit temporal key that is shared among clients and access points. TKIP combines the temporal key with the client MAC address. It then adds a relatively large, 16-octet initialization vector to produce the key that will encrypt the data. This is illustrated in Fig. 2. This procedure ensures that each station uses different key streams to encrypt the data. WEP Key hashing protects weak Initialization Vectors (IVs) from being exposed by hashing the IV on a per-packet basis.

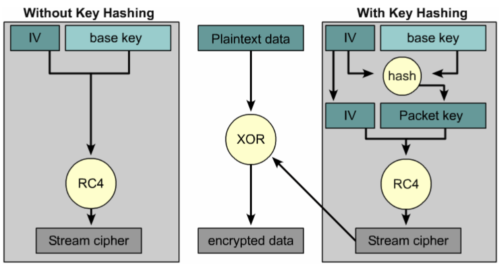


Fig. 1. Temporal key integrity protocol (TKIP).

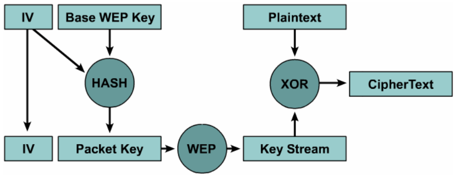


Fig. 2. Per-packet keying operation.

3. Enterprise Wireless Security

3.1. Advanced Encryption Standard (AES)

In addition to the TKIP solution, the 802.11i standard will most likely include the Advanced Encryption Standard (AES) protocol. AES offers much stronger encryption. In fact, the U.S. Commerce Department National Institute of Standards and Technology (NIST) organization chose AES to replace the aging DES. AES specifies three key sizes, which are 128, 192 and 256 bits. It uses the Rijndael Algorithm [4]. If someone where to build a machine that could recover a DES key in a second, then it would take that machine approximately 149 thousand-billion (149 trillion) years to crack a 128-bit AES key. To put that into perspective, the universe is believed to be less than 20 billion years old.

3.2. Virtual Private Networks (VPNs)

IP Security (IPSec) is a framework of open standards for ensuring secure private communication over IP networks. IPSec Virtual Private Networks (VPNs) use the services defined within IPSec to ensure confidentiality, integrity, and authenticity of data communications across networks such as the Internet [5]. VPN deployment is illustrated in Fig. 3. IPSec also has a practical application to secure WLANs. It does this by overlaying IPSec on top of 802.11 wireless traffic.

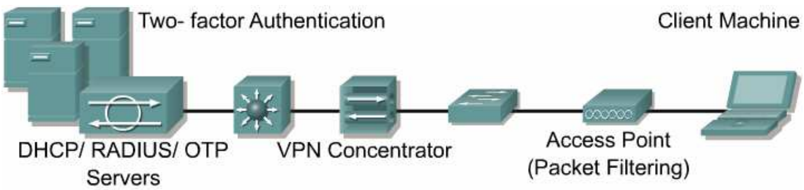


Fig. 3. Virtual private networks deployment.

IPSec provides for the confidentiality of IP traffic. It also has authentication and anti-replay capabilities using Message Digest 5 (MD5) or Secure Hash Algorithm (SHA). Confidentiality is achieved through encryption, which uses Data Encryption Standard (DES), Triple DES (3DES) or AES.

4. Conclusion

The main focus of this work was WLAN security. To secure a WLAN, both encryption and authentication are necessary. Simply fixing the problems discovered in WEP is not enough. Wireless networks must authenticate both users and devices. 802.1x and TKIP provide the necessary security mechanisms to protect WLANs. VPNs were discussed as another way to successfully secure WLANs.

Acknowledgement

This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0040/08 “Mathematic-graphical modelling of safety attributes of safety-critical control systems”.

References

[1] SANKAR, K., SUNDARALINGAM, S., BALINSKY, A., MILLER, D. *Cisco Wireless LAN Security*. Cisco Press, Indianapolis, IN, USA, 2005. ISBN 1-58705-154-0

[2] GEIER, J. *Wireless Networks first-step*. Cisco Press, Indianapolis, IN, USA, 2004. ISBN 1-58720-111-9

[3] MALM, T., KIVIPURO, M., HÉRARD, J., BOEGH, J. *Validation of Safety-related Wireless Machine Control Systems*. NT Technical report TR 605, Oslo, Norway, 2007. ISSN 0283-7234

[4] FRANEKOVÁ, M., KÁLLAY, F., PENIAK, P., VESTENICKÝ, P. *Komunikačná bezpečnosť priemyselných sietí*. Edis, Žilina, Slovakia, 2007. ISBN 978-80-8070-715-6

[5] CISCO SYSTEMS, Inc. *Next Generation Wireless – 802.11n Wireless Technology*, electronic file available at http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps6973/ps8382/prod_brochure0900aecd806b8a92_ns767_Networking_Solutions_Brochure.html, 15.03.2009

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Route Prediction on Tracking Data to Location Based Services

[*]Attila István Petróczi
[*]Budapest University of Technology and Economics, Faculty of Electrical Engineering and Informatics,
Department of Telecommunications and Media Informatics,
Magyar tudósok krt. 2, 1117 Budapest, Hungary, {petroczi}@tmit.bme.hu

Abstract. Wireless networks have become so widespread, it is beneficial to determine the ability of cellular networks for localization. This property enables the development of location-based services, providing useful information. These services can be improved by route prediction under the condition of using simple algorithms, because the limited capabilities of mobile stations. This study gives alternative solutions for this problem of route prediction based on a specific graph model. Our models provide the opportunity to reach our destinations with less effort.

Keywords: mobility modeling, route prediction, Markov model, pattern matching model, location-based services

1. Introduction

In our rushing world a modern man needs communication and information channels, that is why they enjoy the benefits of using mobile phones and location systems. The opportunity to use location-based services (LBS), which are information services accessible with mobile devices through the PLMN (Public Land Mobile Network) and utilizing the ability to make use of the location of a mobile device, is given by these equipments. It does not depend on the device is used. These services are based on localization, they use GPS coordinates or information of GSM cells. It is very useful to know where the nearest preferred restaurant or petrol station is, but it can be even more useful to know which one we can reach with the least effort, (e.g. go to forward instead of turn back) . It means that the system offers the place which will be the closest to our route in the following few minutes. It is significant in these applications that the algorithm must be simple, because the mobile devices have small capacity of memory, storage and computing. Our aim is to give such solutions that are very simple and easy to implement, to predict mobility.

2. Mobility Modeling

Examining past cases of prediction procedures [1][2][3], it is observable that they work efficiently only in a given environment. It means that known models are produced by taking account of the characteristics of the networks, so enormous changes are needed to use them in another environment. Our study would like to give such a general solution to predict mobility of mobile equipments and their owners that is simple and easy to implement. In our realization the basis of each single method is a graph-model, which can be generated by converting the tracking data into graphs. This means that the single directed graph representative of the tracking data is examined, that can be generated from the data according to the required accuracy. The simplest graph creation method is to divide the area into provinces. Let us map the provinces onto the vertices of the graph. We may draw the a_{ij} directed edge of the graph by examining the fixed routes. If we find a route from province i to province j, then we draw a directed edge from the point p_i to the point p_j. You can see such a directed graph representative in Fig. 1.

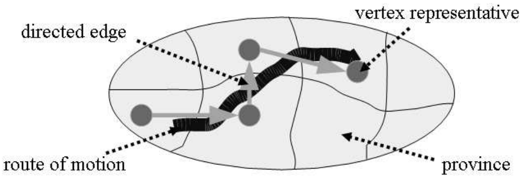


Fig. 1. A given route and its graph representative.

The methods described below give alternative solutions to predicting routes in such a graph model.

3. Mobility Prediction Models

As we have already mentioned, our aim is to define and to analyze general prediction methods, which can be applied in a wide range, and they can cooperate with the infrastructure of any tracking systems. Probably they are not the most efficient ones in a special environment, but with small-scale transformations they work with any kind of systems. Our elaborated models are classified into 3 groups: simple statistical model, Markov models and pattern matching models.

Conventionally nominate with $p_i{\cdot}p_j$ that case, when in a given route at time t our position is the vertex p_i, and at time $t+1$ the vertex p_j is our position. In the following cases we nominate with $p_i{\rightarrow}p_j$ those incidences where at a given t_1 time the position is vertex p_i, then at a given $t_2>t_1$ time the position is vertex p_j. The models predict the next vertex based on the available information, nominate it with X.

The prediction is based on a well-known indicator of confidentiality level. We count it with the following formula:

$$C(X) = \frac{\text{the number of routes crossing series of vertices } \mathrm{P_X}}{\text{the number of routes crossing series of vertices } \mathrm{P_i}} \,. \tag{1}$$

The models can give a probability on a confidentiality level, what the next station of the movement will be.

3.1. Simple Statistical Model

One of the tasks attributed to the data mining is the mining of the frequent element sets. The Simple Statistical Model tries to use this to predict routes. $|p_i{\cdot}p_k|$ nominate the number of the set of routes, contains p_i and p_k too. At a given vertex p_i the model predicts $X=p_k$ vertex for the next station of the route, which satisfy in all cases the following inequality:

$$\left|p_i \cdot p_k\right| > \left|p_i \cdot p_j\right| \qquad \forall j \neq k; \qquad j = 1;2;3;\ldots \tag{2}$$

3.2. Markov models

Markov models are originated with Markov chains. We assume that the system and its graph representative are given. In this case the directed graph - if there are enough given vertices and edges - behaves similar to the Markov chains. Let us map the vertices of the graph to states of the Markov chain, let us build a matrix from its edges, which can be considered the matrix of transition probabilities.

These models are originated with the Markov chain. We may map them to the terminology of the graph model, so Markov models predict the vertex $X=p_{i+1}$ which has the highest confidentiality level in series $(p_{i-z}\cdot...\cdot p_{i-2}\cdot p_{i-1}\cdot p_i)\cdot p_{i+1}$, where the expression in parentheses is known and $z=0;1;2;...$. In mathematical phrases: the model predicts that vertex $X=p_{i+1}$, with the highest corresponding $C(p_{i+1})$ value.

Markov models are not capable of prediction in any situation, when there is vertex, which is not a head of any directed edges or when there are no $z-1$ former vertices of the walk.

In practice we implemented and tested the First-, the Second- and the Third-order Markov Models. The depth of the Markov models can be increased, but the demand of memory and storage monotonously rises.

By analyzing the former models the following question was formulated in our minds: what kind of results can we get with easing the rigorousness of direct succession or can it help to improve the accuracy of the former models. The pattern matching models were formulated as the result of this intuition. We hope that this will be compensates in increased accuracy and efficiency compared with our former models.

3.3. Pattern matching models

Pattern matching models are special cases of frequent sequence mining belonging to the category of frequent sample mining. Frequent sequence mining means that we would like to define part series, which are often appearing in given series. The often expression indicates that in the case of the original task we only deal with the series if the number of the existing series among the routes is higher than a certain threshold. Unfortunately this algorithm is exponential, the frequent sequence mining demands much time and resources in large datasets.

In the case of models drawn up by us, task is not entirely this. On one hand we do not demand a minimal incidence threshold from the series, on the other hand we examine only a certain long incidence of series. We expect longer running time and bigger memory claim from these algorithms, than we have experienced in Markov models.

We implemented and tested four Pattern Matching Models. With the Pattern Matching Model No. 1 we ease the constraints of the First-order Markov model, instead of $(p_i)\cdot p_{i+1}$, we examine $(p_i)\rightarrow p_k$. The Pattern Matching Model No. 2 eases the constraints of the Second-order Markov model, instead of $(p_{i-1}\cdot p_i)\cdot p_{i+1}$, it examines $(p_k\rightarrow p_i)\cdot p_{i+1}$. We ease the constraints of the Third-order Markov model with Pattern Matching Model No. 1, instead of $(p_{i-2}\cdot p_{i-1}\cdot p_i)\cdot p_{i+1}$, we examine $(p_k\rightarrow p_{i-1}\cdot p_i)\cdot p_{i+1}$. Pattern Matching Model No. 4 is interpreted as the extension of the Pattern Matching Model No. 3 with another former vertex, it means we examine $(p_{k2}\rightarrow p_{k1}\rightarrow p_{i-1}\cdot p_i)\cdot p_{i+1}$. The base of the prediction is the highest corresponding $C(p_{i+1})$ value.

4. Simulation

Public tracking dataset of location-based services can not accessible, therefore simulation dataset was used to demonstrate the effectiveness of our solutions. Our data source simulates customers of a hypermarket, where the topology of shelves, products and their locations, and product sets bought by the costumers are based on real life. The sufficient quality of the simulated dataset is guaranteed by the complex artificial behavior of customer agent, which contains the following aspects:

- Customers enter the hypermarket with an explicit aim represented by a set of products, but they can buy items impulsively as well.

– When an agent enters to the hypermarket, it has rough map of the topology of the hypermarket and the positions of products, and during its movements this map is improved by new data.
– The agent has a searching strategy to find products with unknown locations, but it has a probability model to give up the searching as well.

Our dataset represents a 90×50 customer place, which is divided to 121 provinces. There were 14 000 training routes and 2 200 routes for validation. The result of the training the generated graph has got 121 vertices and 888 directed edge.

In our case the number of the vertices of the routes in the training set is 1 510 349, and the validation set has 284 915 vertices. The average length of routes is 126.4. We can consider this dataset enough big to treat the measured accuracy as general.

Table 1 shows the next vertex's prediction time and the accuracy in percent of the single models. Intuitive requirement is taken, that a real-time application has to give result in 1 second. Our models have to satisfy this statement. The accuracy of the methods is varied, but over the 50% it is acceptable, because of the complicated task.

Model	Prediction Time	Accuracy
Simple Statistical Model	0.18 ms	32.4%
First-order Markov Model	0.18 ms	49.7%
Second-order Markov Model	0.18 ms	64.7%
Third-order Markov Model	0.17 ms	69.2%
Pattern Matching Model No. 1	0.16 ms	6.1%
Pattern Matching Model No. 2	2.53 ms	51.6%
Pattern Matching Model No. 3	2.74 ms	65.2%
Pattern Matching Model No. 4	2.51 ms	51.9%

Tab. 1. Comparison of the prediction models

5. Conclusion

The success of the location-based services can be improved with mobility prediction. We gave an alternative solution for this. We defined 3 types of models, which are based on the representative graph that can be built up by the area separated into provinces. The Markov models are originated with the Markov chains. The models use only the current and some direct former vertices for prediction. If we use these models we can reach about 70% accuracy. Many models were published in which the accuracy is about 60%, still they are used. The pattern matching models are to improve the efficiency of Markov models. This approach can help a bit, but the offline running time increase more.

The actual published models can be improved, because the accuracy can be increased by training at a specific environment. The effects of the environment can be trained by the models, and it makes them more useable in the marketing world, or finds acquaintances in the users' neighborhood.

References

[1] LIU, T., BAHL, P., CHLAMTAC, I. *Mobility Modeling, Location Tracking, and Trajectory Prediction in Wireless ATM Networks*. IEEE Journal on Selected Areas in Communications, vol. 16, no. 6, 1998

[2] CHAN, J., ZHOU, S., SENEVIRATNE, A. *A QoS Adaptive Mobility Prediction Scheme for Wireless Networks.* Global Telecommunications Conference, vol. 3, 1998

[3] JUNIUS, M., KENNENMAN, O. *New Methods for Processing GSM Radio Measurement Data: Applications for locating, Handover, and Network Management.* Vehicular Technology Conference, vol. 44, no. 1, 1994

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Using Standards – a Way to E-content Interoperability and Reusability

*Katarina Pilatova
*University of Zilina, Faculty of Management Science and Informatics, Department of Information Networks, Univerzitna 8215/1, 01026 Zilina, Slovakia, katarina.pilatova@fri.uniza.sk

Abstract. A lot of technologies, applications, tools and platforms are used in Technology Enhanced Education. Their interoperability is necessary for effectiveness of their use in the education. Acceptation of standards and specifications is a way to enable such interoperability. This paper dealt with three of widely used specifications for e-content creating - IMS Content Packaging Specification, ADL SCORM and IMS Common Cartridge.

Keywords: standard, specification, study materials, SCORM, Content Packaging, Common Cartridge.

1. Introduction

Information and Communication Technology (ICT) is used in every area of our life, the education is not an exception. Using of ICT in learning known as E-Learning or Technology Enhanced Learning is routine for universities, educational institutions, secondary schools and commercial sector as well.

The base for effective E-Learning is good quality study material. E-study materials can be understood as digital processed contents for educational purposes. In commercial sector, there are enough finances to support their educational needs. But situation is more complicated for schools. They depend on their own staff – on teachers, who are responsible for creating of study materials. Teachers created those e-study materials according to their possibilities and such materials used to be in various quality, size and format and used by creators only.

Standards can change this. They bring some simple rules for content structure and storage format. Standards will allow finding, editing and using materials created by another person. In ideal case, then well-prepared content can be found and available through a central repository – library of learning contents.

Some of standards for content creation are described in this paper. They come from organizations IMS GLC (Instructional Management Systems Global Learning Consortium) and ADL (Advanced Distributed Learning).

2. IMS Content Packaging Specification

The IMS Content Packaging Specification provides the functionality for describing and packaging learning materials, such as an individual course or a collection of courses, into interoperable, distributable packages. Content Packaging addresses the description, structure, and location of online learning materials and the definition of some particular content types. Learning materials described and packaged using the IMS Content Packaging (IMS CP) XML format should be interoperable with any tool that supports the IMS CP. Content creators can develop and distribute material knowing that it can be delivered on any compliant system, thereby protecting their investment in rich content development. IMS uses XML as its current binding, and XML-Schema as its primary XML control document language. Some IMS bindings use parts of other IMS XML bindings; for example, the Content Packaging specification also uses the IMS Meta-Data.

A Content Package defined in the Specification consists of two main parts – Manifest and Content. The Manifest is an XML file describing structure of content and resources of package and contains Metadata as well. In the Content, there are real files with learning content. The Content Package is for distribution packaged to PKZip v2.04g format (.zip) and is called Package Interchange File (PIF). Fig. 1 shows this PIF.

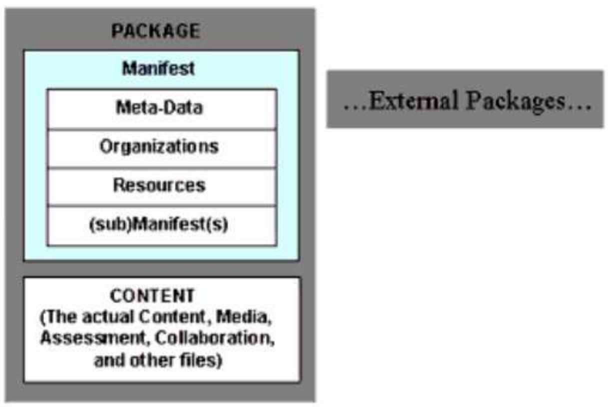


Fig. 1. The Package Interchange File

The IMS Content Package Specification is a part of some other specifications for content creating. ADL SCORM and IMS Common Cartridge are two of them. They are worked on the base of IMS Content Package, but they are more complex.

3. ADL SCORM

SCORM is a collection of standards and specifications adapted from multiple sources to provide a comprehensive suite of e-learning capabilities that enable interoperability, accessibility and reusability of Web-based learning content.

It has the base in the work of AICC (Aviation Industry Computer-Based Training Committee), IMS GLT, IEEE (Institute of Electrical and Electronics Engineers), ARIADNE (Alliance for Remote Instructional Authoring and Distribution Networks for Europe) and others. Its aim is to create one unified "reference model" of interrelated technical specifications and guidelines that meet Department of Defense high-level requirements for Web-based learning content.

SCORM consists of four distinct books that contain the critical elements of SCORM as follows:

1. The SCORM Overview book contains high-level conceptual information, the history, current status and future direction of ADL and SCORM and an introduction to key SCORM concepts.
2. The SCORM Content Aggregation Model (CAM) book describes the components used in a learning experience, how to package those components for exchange from system to system, how to describe those components to enable search and discovery and how to define sequencing rules for the components.

3. The SCORM Run-Time Environment (RTE) book describes the Learning Management System requirements in managing the run-time environment.
4. The SCORM Sequencing and Navigation (SN) book describes how SCORM conformant content may be sequenced to the learner through a set of learner-initiated or system-initiated navigation events.

The CAM book is the most important for learning content creating. In the book, there are discussed Content Model (defines the content components of a learning), Content Packaging (defines how to represent the intended behaviour of learning and how to aggregate activities of learning resources for movement between different environments), Metadata (defines how to describe the components of the content model) and Sequencing and Navigation (it is a rule-based model for defining a set of rules that describes the intended sequence and ordering of activities).

4. IMS Common Cartridge

IMS Common Cartrige (IMS CC) defines open format for rich online content distribution designed to ensure the correct installation and operation of content among various platforms and tools conformant with IMS CC. It implements several standards and specifications such as Dublin Core Metadata Element Set mapped to IEEE Learning Object Metadata, IMS Content Packaging Specification, IMS Question and Test Interoperability and IMS Authorization Web Service.

Two problems are solved in IMS CC. The first is to provide a standard way to represent digital course materials for use in online learning systems so that such content can be developed in one format and used across a wide variety of learning systems (often referred to as course management systems, learning management systems, virtual learning environments, or instructional management systems). The second is to enable new publishing models for online course materials and digital books that are modular, web-distributed, interactive, and customizable. The focus of Common Cartridge is interactive collaborative learning situations, typically with a teacher, professor, or instructor involved in guiding a cohort. The learning materials can be online, offline, or both - a situation often referred to as hybrid or blended learning.

Common Cartridge specifies six things [4]:

1. A format for exchange of content between systems so that there is a common way to interpret what the digital learning content is and how it is organized. The content is described in a manifest and the components that make up the manifest may be in the exchanged package or external to the package (referenced by URL).
2. An authorization standard (access rules) for each component of the package. This allows "protected" content or applications (those requiring a license) to be contained in a cartridge in a flexible way along with unprotected content.
3. A standard for the metadata describing the content in the cartridge - based on Dublin Core. Common Cartridge is extensible to allow other metadata schemas.
4. A standard for test items, tests, and assessments. This standard allows learning systems to understand imported assessments as natively - so they can be manipulated (such as deciding what items are to be used and where in the flow of a course) as needed in the learning system.
5. A standard for launching and exchanging data with external applications so that they can be part of a single learning experience orchestrated through the learning system. These can be literally any type of application in any location, such as social

networking, wiki, external assessment systems, adaptive tutors, varieties of web-based content libraries, or other learning systems.

6. A standard for populating online discussion forums for collaboration among students. This allows such forums to be pre-populated with potential exercises, discussion threads, and so forth.

5. SCORM versus Common Cartridge

SCORM was developed to support portability of self-paced computer-based training concept while IMS CC to support the use of online digital course materials and digital books in the instructional context. Current educational scenarios require advances in assessment, interactive content, sequencing of content, collaboration, facilitation and authorization that SCORM was not designed to address, but Common Cartridge was. IMS CC provides more functionality than SCORM, but does so in a different way. However IMS CC is more complex than SCORM, it is easier to implement and to test for conformance.

Common Cartridge was designed to obtain much higher levels of interoperability what is done by removing the run-time component associated with SCORM and by achieving agreement on specific subsets of widely used specifications. But SCORM content don't have to be wasted. SCORM content can be converted to IMS CC or it can be contained within a cartridge, but it is up to the learning platform to be able handle it. In the second case, the learning platform has to be SCORM and IMS CC both compliant.

6. Conclusion

In field of Technology Enhanced Learning as well as in other fields, standards and specifications are a base platform for interoperability and reusability. Their use within content creating enables to save money, time and energy because high quality learning materials can be found through the e-content library, adapted and used many and many times in various situations. Standards support market efficiency and open up the market for greater choice in both content and platforms.

Acknowledgement

The support by project KEGA number 3/5205/07 is gratefully acknowledged.

References

[1] *ADL Sharable Content Object Reference Model* (SCORM®), 3rd Edition Content Aggregation Model, Version 1.0, 2006

[2] *IMS Content Packaging Information Model*, Version 1.1.4, 2004, http://www.imsglobal.org/content/packaging/cpv1p1p4/imscp_infov1p1p4.html

[3] *IMS Learning Design Information Model*, Version 1.0, 2003, http://www.imsglobal.org/learningdesign/ldv1p0/imsld_infov1p0.html

[4] *Common Cartridge Frequently Asked Questions*, http://www.imsglobal.org/cc/ccfaqs.html

[5] *IMS Common Cartridge Profile*, Version 1.0, 2008, http://www.imsglobal.org/cc/ccv1p0/imscc_profilev1p0.html

[6] LEGINOVA, K., Štandardizácia študijných zdrojov pre e-learning, in: *ITlib* 3/2007, pp. 36-42

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

New Assignment and Reasignment Techniques for OVSF Codes in WCDMA Systems

[*]Peter Poloha, [*]Norbert Majer, [*]Ján Haring, [*]Vladimír Wieser
[*] University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia,
poloha.peter@gmail.com,{majer, haring, wieser}@fel.uniza.sk

Abstract. In this paper we present most important adaptive schemes for OVSF codes in WCDMA system. In an effort to reduce the decision time of OVSF codes assignment and capacity enhancement we are trying to explain the best performance of each scheme. Some schemes are easy to use but with high code blocking probability like Leftmost code assignment (LCA) scheme is [1]. Another is very compact to use like in the event of Dynamic Cost Assignment (DCA) scheme [2]. On behalf of capacity enhancement and code assignment time depression is very important to find out the best performance scheme.

Keywords: Assignment Scheme, OVSF Codes, WCDMA, Spreading Factor (SF), code blocking

1. Introduction

Mobile Communications have come into our daily life recently. The second generation of mobile communication systems was limited because of Orthogonal Constant Spreading Factor (OCSF). Services provided in existing 2G systems are typically bounded to voice and low-bit-rate data.

In third generation (3G) are expected higher data rates, QoS by file transfer, and new services. To satisfy different requirements, the system has to provide variable data rates [1]. WCDMA has been selected as the technology for use in the UMTS terestrial radio access (UTRA). WCDMA can flexibly support mixed and variable-rate services. Such flexibility can be achieved by the use of Orthogonal Variable Spreading Factor (OVSF) codes as the channelization codes. Because of this fact we present in this paper a study of OVSF code assignment schemes and their performances. We show that these chosen schemes manage to mitigage the problem of wireless system capacity. By selection of suitable assignment scheme we can reduce the call blocking probability [4].

2. Problem statement

In WCDMA two operations are applied to user data. The first one is channelization, which transforms every data bit into a code sequence. The length of the code sequence per data bit is called the spreading factor (SF). The second operation is called scrambling, that is used to separate the signals from different sources [1], [4].

OVSF codes are presented in the code tree by, where SF is the spreading factor and k is the branch number (Fig.1.).

We are trying to find out the best performance assignment scheme to reduce call blocking with capacity increasing.

In comparation we are introducing following schemes.

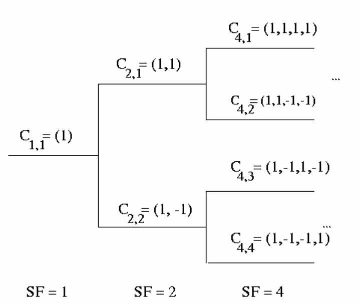


Fig. 1. OVSF code tree.

2.1. Assignment Schemes

One of the most used assignment scheme for OVSF codes is well known as a Crowded first scheme (CF). Assignment procedure is follows. If there is one or more than one code in the code tree with a rate kR, pick the one whose ancestor code has the least free capacity. When there are ties between two or more codes with the same SF factor, we will go one level up by comparing code's ancestors. This is repeated until the subtree with the least free capacity is found. In case where two codes represent the same code tree, the leftmost rule i sused [2].

Leftmost code assignment (LCA) scheme is using at first the left side of code tree (Fig.2.). These codes are checked for an availabelity, and when they are free, they are assigned to a call. Otherwise the call is rejected. LCA scheme is very simple and do not require a reassignment. The complication is by call blocking, that can be avoided by dynamic code assignment (explained lower). The scheme will be usefull when number of calls is limited.

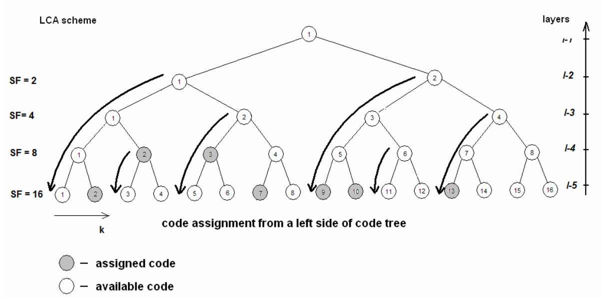


Fig. 2. LCA assignment scheme.

Basic point in Dynamic Code Assignment Scheme (DCA) is a cost function of each code Cl,k is defined as a capacity od occupied descendants od the code Cl,k. If the code is available, it is assigned to a new call. Otherwise, when we have no available code and the capacity is within the maximum, cost is checked for every blocked code with rate equal to rate of incoming call. The call with minimum cost is picked and descendants of minimum cost are reassigned to other branches (Fig. 3.). Now code is available and can be assigned to a new call. This assignment scheme has the best performance to remove blocking probability. The signaling overhead increases proportional to the number of reassignments.

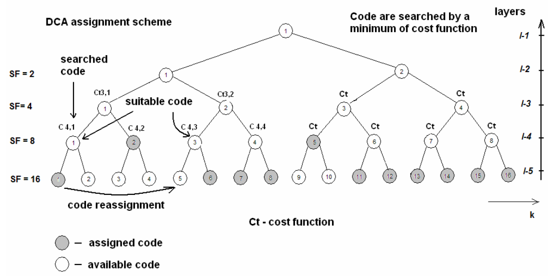


Fig. 3. DCA assignment scheme.

Ancestor Cost Assignment Scheme (ACA) is based on ancestor cost function. It is also called as a compact scheme because the code trees are compact in terms of locations of busy codes. No reassignments are necessary in ACA scheme, because the new call is always given code location so that the available capacity for future calls is highest. This scheme is the best assignment scheme in terms of having minimum blocking probability. For a vacant code Cl,k ancestor cost ACtl,k is calculated for a number of descendants of code Cl,k (Fig. 4.). All the vacant codes are checked for a finding of ancestor code with maximum number of descendant codes. If tie occurs by vacant codes in one layer l-1, searching will continue in higher layer, till suitable parent code is chosen.

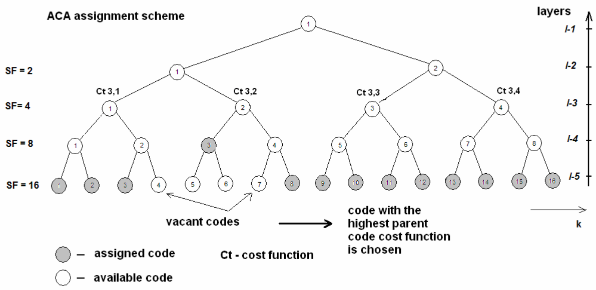


Fig. 4. ACA assignment scheme.

Class partition assignment scheme (CPA) divides the users to different classes (rates). So to this users are assigned different set of codes. Let us define the T as a total number of code trees and L as a different classes of codes. R is the basic rate. If the arrival rate of all types of classes is equal, than the trees T1, T2…TL are equal. The major benefit of CPA scheme is less number of code searches and also a less decision time for finding appropriate code (Fig. 5.).

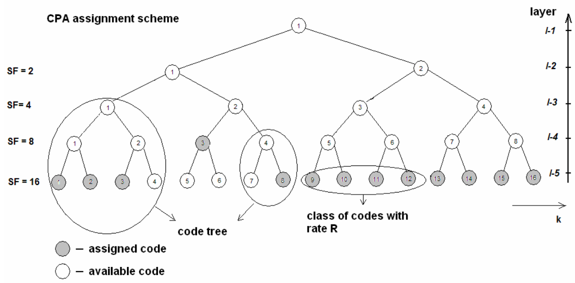


Fig. 5. CPA assignment scheme.

Class Partition Assignment and Reassignment scheme (CPAR) presents an improvement of CPA scheme. It reduces the call blocking probability towards The CPA assignment scheme. Total number of trees T is divided into assign tree set (TA) and reassign tree set (TR). If Tl is the number of assignable trees for class l, TA group can be defined by T1 + T2 + ...+Tl = TA. By CPAR scheme the assign algorithm is as follows:

- • To enter an arrival rate, number of calls (users) and number of trees for assign set and reassign set.
- • generate a new call
- • assign available code from assign tree set, if in this set no code is available, then search for code in reassign code set
- • discard the call, when no code is available.

2.2. Equations

For simulation was used MATLAB 7.0.1 [2]. Different probabilities for different rate users as(p1,p2,p3,p4), where p1,p2,p3,p4 are the probabilities of arrival rate R, 2R, 4R, 8R, users. Fig. 6. shows the blocking probability for proposed assignment and reassignment schemes. Results shows [2] that the blocking probability of call blocking is least for ACA scheme, followed by DCA scheme, then CPAR and with CPA scheme at last.

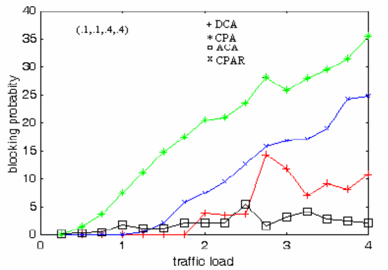


Fig. 5. Blocking probability for distribution (.1,.1,.4,.4) [2].

3. Conclusion

In this paper we present most used assignment and reassignment schemes, they are used for WCDMA system. Scheme selection is based on parameter which is helpful for system we need to have. These results will be used in hybrid scheme to increase the capacity of WCDMA system. It must be carefully chosen which scheme will be most effective for WCDMA system.

References

[1] TSENG, Y - C., CHAO,C. – M. *Code Placement and Replacement Strategies for Wideband CDMA OVSF Code Tree Management*, IEEE transactions on mobile computing, VOL. 1, NO. 4, october-december 2002.

[2] DAVINDER S SAINI, SUNIL V BHOOSHAN, *Assignment and Reassignment Schemes for OVSF Codes in WCDMA*, IWCMC’06, July 3-6, 2006, Vancouver, British Columbia, Canada.

[3] SHUMING GUO, JIANHUA PENG, JIANGXING WU, *A reassignment Code Allocation Scheme for Multirate Services on OVSF CDMA Cellular Networks*, 1-4244-1312-5/07 IEEE 2007.

[4] ANGELOS N. ROUKUNAS, DIMITRIOS N. SKOUTAS, *Management of Channelization Codes at the Forward Link of WCDMA*, 1089-7798/05, IEEE 2005

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Postprocessing of Sonographic and MR Images of Temporomandibular Joint

*Zdeněk Smékal, *Jiří Přinosil, **Ondřej Liberda, **Andrea Šprlaková
* Faculty of Electrical Engineering and Communication, Brno University of Technology
Purkynova 464/118, 612 00 Brno, Czech Republic
** The Faculty Hospital Brno, Faculty of Medicine, Masaryk University, Komenskeho nam. 2
Brno 662 43, Czech Republic

Abstract. The temporomandibular joint is one of the most complex joints in the human body; in its examination, non-invasive diagnostic methods can be used that are based on magnetic resonance or ultrasound examination. The output is always an image signal, which, however, is often degraded due to noise and interference from both the scanning instruments themselves and their environment. In the paper, the application is described of the method of thresholding the wavelet transform coefficients in order to enhance image signals from an MR tomograph or a sonograph and to attenuate noise. The method was used to enhance the image of a relatively small temporomandibular joint, which facilitated the determination of a correct diagnosis of the afflicted joint.

Keywords: Temporomandibular joint, MR tomograph, sonograph, wavelet transform, image enhancement.

1. Introduction

Disorders of temporomandibular joint are the second most frequent cause of pain in the facial region. The laterodeviation of temporomandibular joint is one of the most frequent temporomandibular joint disorders, usually associated with characteristic clinical findings such as pain, joint crepitus, and lateral function of mandibles [1]. Usually, patients are first treated by conservative, non-invasive methods but if these methods prove inefficient, surgery is resorted to. Its indication, however, requires that a correct diagnosis be made. The surgeon must therefore have at their disposal all the necessary data in order to be able to choose an optimum treatment method and thus avoid any damage to the temporomandibular joint or potential post-operation complications. Because of the relatively small size of the joint, a correct diagnosis of the joint disorder often proves very difficult [2].

In the image examination of temporomandibular joint today, the “golden standard” is seen in MR tomography, which yields correct diagnoses in 95% of cases [3]. The main disadvantage of MR tomography is the high financial cost (purchase and running costs), which is reflected in the low clinical availability of this method of examination. Moreover, this type of examination can be applied on only a limited scale to patients suffering from claustrophobia, patients with pacemakers and metallic prostheses, and to children. According to current scientific studies, ultrasound can be employed in the diagnosis of temporomandibular joint disorders, which enables a dynamic representation of temporomandibular joint. An advantage is here the low purchase and running cots, and the simple examination method. However, well trained attendants are indispensable here in order to achieve correct diagnosis, which usually happens in 80% of cases [4].

Image information from the MR and from the sonograph is often distorted due to interference and noise coming from the instruments themselves and from the environment.

The present paper is concerned with a method for enhancing the image information for both types of temporomandibular joint examination.

2. Methods for Image Enhancement

An image obtained from the tomograph or sonograph is often degraded by wide-band additive noise of the input amplifier and other circuits. To enhance image information thus means to remove this additive noise component while preserving the original data to the greatest possible degree. To remove noise from the image, standard methods of digital image filtering are used, the Wiener and the median filters in particular [5]. Considering, however, the wide-band nature of additive noise, the median filter does not appear to be very suitable for the enhancement itself since this filter is basically a lowpass filter that attenuates only the higher spectral components of the image (noise). The Wiener filter, on the other hand, works on the principle of creating a filter that is inverse to the noise model; this, however requires an apriori knowledge of noise features and the probability distribution of noise in the image, which in most cases is not realistic and therefore white noise with normal distribution is often considered instead. But noise occurring in images from MR tomography and from sonography does not exhibit white noise features and the application of this filter is thus again not very suitable.

Another method that can be used to enhance image information is the wavelet analysis, which is often used in image data compression (JPEG 2000) or in enhancing speech signals hidden in noise [6]. Thanks to the wavelet transform, noise-embedded image can be decomposed into individual parts and wide-band noise separated from useful information contained in the image.

2.1. Image Enhancement via Wavelet Analysis

The wavelet transform is an integral transformation that assigns to a real signal the image of two variables, which represent the position (translation) and the scale (dilatation) of the wavelet in comparison with the mother wavelet [6]. What is concerned here is a decomposition of input signal into a series of wavelet coefficients representing the expression of this signal by the mother wavelet using its shifting and the change in scale. This can be achieved by filtering the signal by a pair of orthogonal filters of the type of highpass and lowpass filter, where the highpass filter output consists in detailed (scale) coefficients, and the lowpass filter output in the approximation (wavelet) coefficients. The approximation coefficients can likewise be further decomposed to the required decomposition level. In the decomposition of a two-dimensional signal (image) the resultant decomposition is given by the way the coefficients are arranged into individual rows and columns.

Noise can be removed from the image under analysis by thresholding detailed coefficients of individual decomposition levels; the threshold magnitude should be chosen with a view to sufficient noise attenuation while as much as possible relevant information of the original image is preserved. Here, the so-called hard thresholding can be applied, where detailed coefficients that are lower than the given threshold are zeroed and the others are left without change. Alternatively, the so-called soft thresholding is used, when from coefficients that are higher than the given threshold the value of this threshold is subtracted according to eq. 1. When soft thresholding is used, the time-frequency discontinuity of the output image is reduced and therefore this thresholding is for the enhancement itself more suitable:

$$\hat{c}[k] = \begin{cases} c[k] - T & \text{pro } c[k] > T \\ 0 & \text{pro } c[k] \leq T \end{cases}, \tag{1}$$

where $c[k]$ are the wavelet coefficients prior to thresholding, $\hat{c}[k]$ are the wavelet coefficients after thresholding, and T is the threshold value for the given decomposition level. Figs 1 and 2 give examples of the results of using the wavelet analysis to enhance the temporomandibular joint images from a sonograph and an MR tomograph in comparison with results obtained via the Wiener and the median filtering.

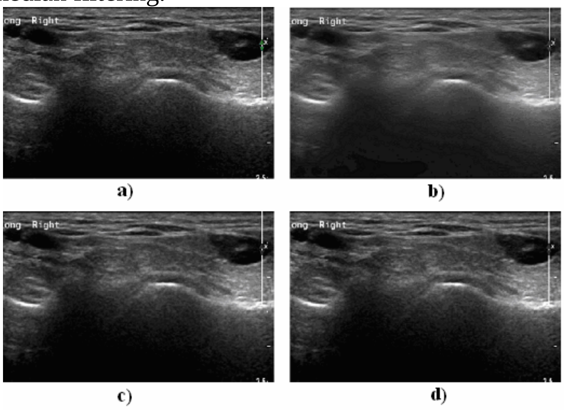


Fig. 1. a) original noise-embedded image from sonographic examination of temporomandibular joint; the other images are enhanced via b) wavelet analysis, c) Wiener filtering, and d) median filtering.

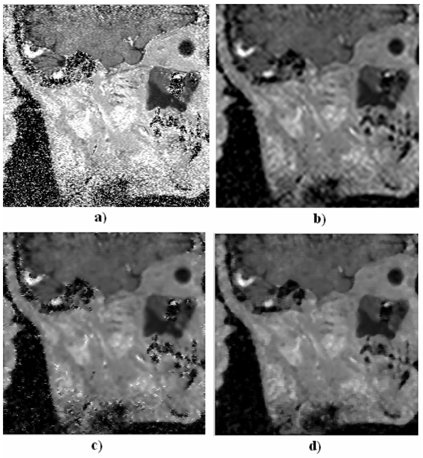


Fig. 2. a) original noise-embedded image from MR tomography examination of temporomandibular joint, the other images are enhanced via b) wavelet analysis, c) Wiener filtering, and d) median filtering.

2.2. Image Enhancement via Sharpening

As can be seen from the above pictures, the attenuation of noise led simultaneously to a certain decrease in the sharpness of the resultant image. It would therefore be appropriate to subsequently sharpen the image. The local, so-called gradient operators can be used to this purpose. Basically, they are filters of the highpass filter type that make use of the convolution mask [5]. Applying these filters to the image will enhance the edges of the object. But the disadvantage is their high sensitivity to the noise in the image, which must therefore be first suppressed. Fig. 3 gives the results obtained by using the gradient operator to enhance temporomandibular joint images from sonograph and MR tomograph, with the noise removed via wavelet analysis.

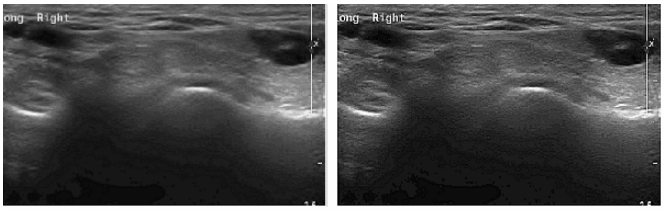


Fig. 3. Enhancement of sonographic image of temporomandibular joint, using gradient operator.

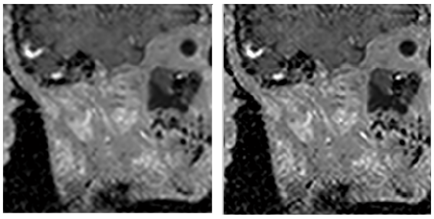


Fig. 4. Enhancement of MR image of temporomandibular joint, using gradient operator.

3. Conclusion

A new method was introduced in the paper for enhancing temporomandibular joint images from MR tomograph and from sonograph with the aid of wavelet analysis for noise attenuation and gradient operators for image sharpening. In the process of noise attenuation the method introduced was compared with the regularly used methods (the Wiener and the median filtering). It could be seen from the results obtained that the Wiener filtering was quite unsuitable for this kind of images. Using the median filtering, the noise was sufficiently attenuated but, in contrast to the wavelet analysis, there was a major attenuation of the high-frequency components, which resulted in reduced image sharpness.

Acknowledgement

This work was supported within the framework of the research plan MSM 0021630513 and project of the Grant Agency of the Czech Republic 102/07/1086

References

[1] Emshoff R, Brandlmaier I, Bosch R, Gerhard S, Rudisch A, Bertram S.: Validation of the clinical diagnostic criteria for temporomandibular disorders for the diagnostic subgroup - disc derangement with reduction. *J Oral Rehabil*. 2002 Dec., 29(12) pp. 1139-45.

[2] Manfredini D, Tognini F, Melchiorre D, Cantini E, Bosco M.: The role of ultrasonography in the diagnosis of temporomandibular joint disc displacement and intra-articular effusion. *Minerva Stomatol*. 2003 Mar., 52(3), pp.93-100, 100-4.

[3] Brandlmaier I, Rudisch A, Bodner G, Bertram S, Emshoff R.: Temporomandibular joint internal derangement: detection with 12.5 MHz ultrasonography. *J Oral Rehabil*. 2003 Aug., 30(8), pp. 796-801.

[4] Smékal, Z.; Liberda, O.; Bartušek, K.; Šprláková, A.; Bulik, O. Sonographic evaluation of disc displacement of temporomandibular joint in comparison to MRI. *Journal of Cranio-maxillofacial Surgery*, 2008 (36), No. 9, p. 278. ISSN: 1010-5182.

[5] Jan, J.: *Digital filtering, signal analysis and restoration*. Brno: VUTIUM Press Brno, 2002. ISBN: 80-214-1558-4. (In Czech)

[6] Smékal, Z.; Sysel, P. Single-channel noise suppression by wavelets in spectral domain. In *Proceedings of COST Action 2102 International Workshop*. Vietri sul Mare, Italy: Springer, 2007. pp. 150-164. ISBN: 978-3-540-76441-0.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Using of Knowledge System in Diagnosis

*Jana Šebeňová
*University of Žilina, Faculty of Electrical Engineering, Department of Control and Information Systems, Univerzitná 1, 01026 Žilina, Slovakia, {jana.sebenova}@fel.uniza.sk

Abstract. The article is dealing with embedding of knowledge system in existing diagnostic system. Possible system architectures are concerned.

Keywords: Knowledge system, knowledge base, reasoning, working memory, diagnosis.

1. Introduction

The diagnostic process consists generally of fault detection, fault identification and fault localization. If there is some diagnosis already applied, it is supposed that the process is mapped and incorrect function is detected. Fault detection and diagnosis is discussed for example in [1].

The task is then to identify, localize and dispatch the fault so that the system can be available as soon as possible. In order to improve the availability of system in such cases, knowledge-based system implemented in diagnostic level of system can be significantly helpful. Look at AI and knowledge based approach is given, among others, in [2].

2. Knowledge System Creation

The most significant part of knowledge system creation is creation of its knowledge base (Fig. 1). It is needed to collect all available data about system structure and behaviour. The appropriate shell should be chosen according to the character of data being at disposal. To be able to process this data and bring them into desired formalism, it can be necessary to transform the data, from their source form into some well structured form, in order to be able to process it automatically.

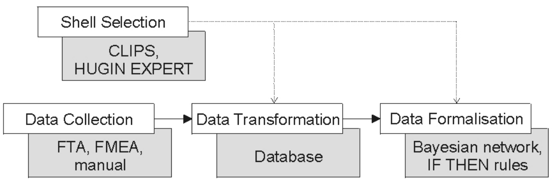


Fig. 1. Knowledge acquisition.

Another task is to gain the facts from process and provide the knowledge system with them. As the diagnostic system is assumed to be at disposal, it is needed to analyze the used diagnostic data, select the information relevant to the fault identification and to transfer them to the knowledge system using the corresponding form.

After determining of solution it is needed to offer it to user.

3. Architecture of Knowledge System

Most of available software tools for creation of knowledge system offer the possibility to make the full-fledged stand-alone system disposing of all main features of knowledge system including ability to infer, make decisions with reasoning, explain, work with uncertain or incomplete data, or hold a dialogue with user.

Besides it is appropriate for application in existing diagnostic unit, if the knowledge system can be embedded in existing software, is able to communicate with external programs, gains the facts from the process being diagnosed in particular from accessible data and requires just minimum of necessary information from user.

3.1. Basic Architecture

The system should consist of the core executing the reasoning itself attended with a service application forming an interface to the process being observed (Fig. 2).

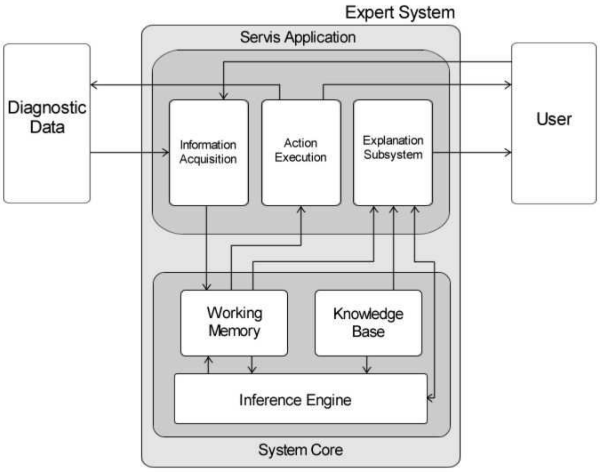


Fig. 2. Basic architecture of the knowledge system.

The core of the knowledge system includes working memory saving the facts about the actual state of the process, knowledge base to store the knowledge about the process rules and relations and inference engine, which is in position to bring possible solutions on the basis of information in working memory and dependences in knowledge base.

The service application should access the process, get data from it, transform them into facts using required formalism suitable for chosen shell and provide the user with the results of reasoning. Actually it should be able to handle the base of facts – to provide the supplying with actual facts, recognize the requests to fill in the missing facts, to look for the absent facts in available data from process as well as ask the user for them when necessary.

3.2. Mechanism for Validation of Knowledge

Using the system, new knowledge, dependences or relations can be recognized. When acquiring new knowledge, knowledge base should be enabled to be supplemented potentially adjusted or corrected.

The service application should thus also be able to create the interface to access the knowledge base (Fig. 3).

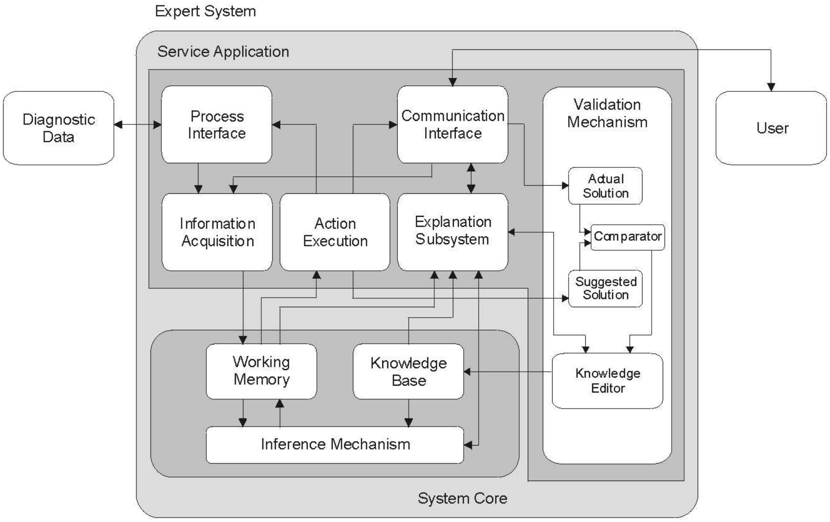


Fig. 3. Architecture of knowledge system extended by validation mechanism.

3.3. Using More Knowledge Approaches

Ordinarily there are more than just one sources of knowledge about the observed system, monitored processes and processes of fault arise and consequences. According to form of source information, various knowledge-based approaches are appropriate to build the knowledge base of the system. Hence it could be purposeful to build a system disposing of more knowledge-bases and corresponding inference mechanisms. In general two architectures of such a system can be considered.

In the first case all the decision mechanisms are working simultaneously in the table architecture, using the same space of table to store the facts (Fig. 4). All the inference engines can access it at the same time sharing all facts and partial solutions.

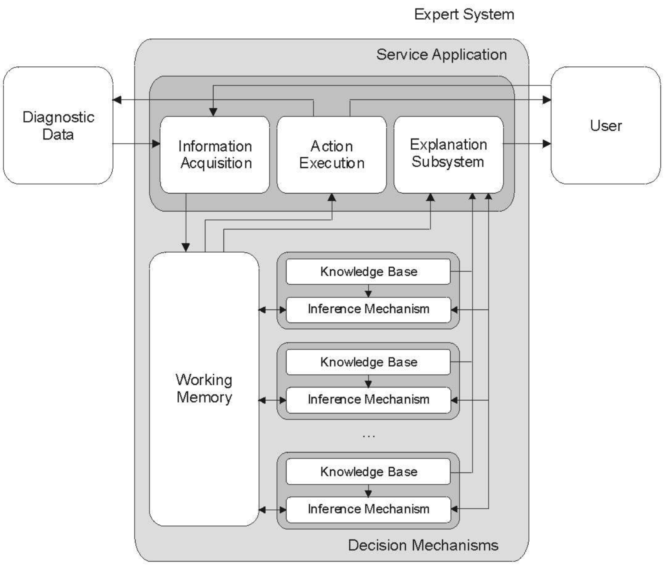


Fig. 4. Extended architecture using common working memory.

The other architecture is based on condition, that the diagnosed system itself or the observed process can be divided into parts, whose faults or dysfunctions can be separated and do not affect and also are not affected by faults of other parts. In this case the task is to decide on the level of the system as the whole, which part has caused the malfunction and the individual decision mechanisms are then considering the contribution of corresponding part (Fig. 5).

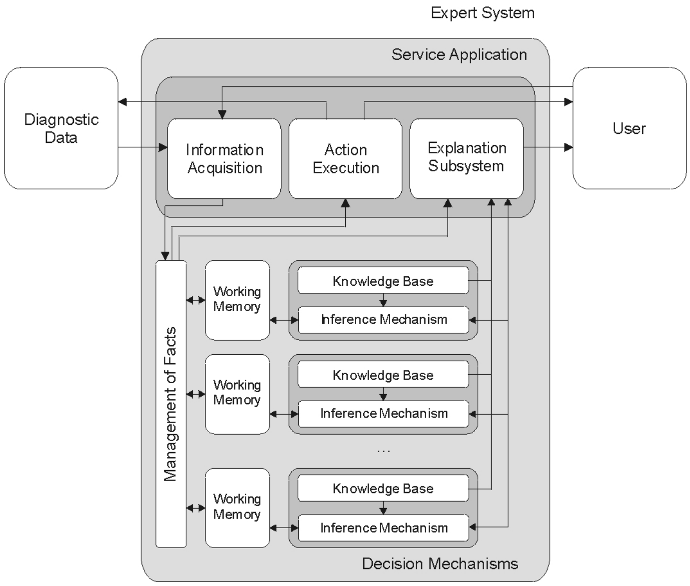


Fig. 5. Extended architecture using separated working memories.

4. Conclusion

The paper was written with motivation to present the methodology of creation of knowledge system for diagnostic purposes, which should be embedded into existing diagnostic system. Basic architectures of such system were considered.

Acknowledgement

This work has been supported by the scientific grant agency VEGA, grant No. VEGA-1/0040/08 “Mathematic-graphical modelling of safety attributes of safety-critical control systems”.

References

[1] CHIANG, L. H., RUSSEL, E. L., BRAATZ, R. D.: *Fault Detection and Diagnosis in Industrial Systems*. Springer, London, 2001

[2] SPALEK, J., JANOTA, A., BLAŽOVIČOVÁ, M., PŘIBYL, P.: *Rozhodovanie a riadenie s podporou umelej inteligencie*. EDIS, Žilina, 2005.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

General Vectorization of Line Objects in Maps

*Peter Tarábek
* University of Žilina, Faculty of Management Science and Informatics, Departement of Transportation Networks, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, {Peter.Tarabek}@fri.uniza.sk

Abstract. Maps consist of multiple object types. The most important are line objects which represent infrastructure. Attributes of these objects are essential for many tasks but in raster format they provide only low level information. Vectorization must be used to obtain vector data. In this paper general vectorization process consisting of five stages is proposed. For these stages short discussion and basic recommendations are given and some proper methods are presented.

Keywords: Vectorization, skeletonization, skeleton, thinning, raster-to-vector.

1. Introduction

Nowadays extraction of graphical information from raster maps is a complex problem. Vector data which represent the necessary information on networks and need less space for their representation are used in many tasks and are essential for GIS and CAD applications. Maps consist mainly from line objects, which the length is much larger then the thickness. These objects represent road infrastructure and in raster format they provide only low level information. Vectorization process needs to be used to transform this data to vector form. Conservation of connectivity and shape of the original objects are two most important conditions for vectorization process. Conserving connectivity means that if two objects are connected in raster image they should be connected also in their vector representation. In other words number of components and their structure in raster image should match number and structure of vector components.

2. Vectorization process

Process of vectorization proposed in this paper consists of 5 stages (Fig. 1) and it is focused on maps that mainly contain line objects. These maps are sometimes called drawing maps. Because of huge scope of possible representations of individual objects, there is no specific vectorization process which can deal with all raster maps. Multiple object types which can overlap each other, different color schemes and sometimes poor quality of input maps make this process very difficult. When dealing with line objects this situation is easier. These objects are represented in very similar manner so if the right methods are applied for each stage this process can be used for wide range of different maps.

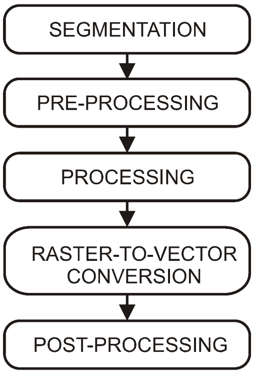


Fig. 1. Stages of proposed general vectorization process.

2.1. Segmentation

Besides of line objects, maps usually contain other objects like text, symbols and regions. In this stage line objects are separated from others. The result is binary image where all line objects should be represented by black pixels and all other objects including background as white pixels. This reduces amount of data and it also speed up and simplify future processing.

One of the most widely used techniques for segmentation is threshold. Threshold use parameter called threshold value to separate objects of interest from others. General condition for threshold can be defined as follows:

$$g(x,y)=\begin{cases}0, if & f(x,y)<T\\1, if & f(x,y)\geq T\end{cases}. \tag{1}$$

where $g(x,y)$ is value of pixel with coordinates x, y in resulting binary image, $f(x,y)$ is original value of pixel and T is threshold value.

More threshold values are usually used to correctly separate objects in image. There are many threshold techniques [1, 2] which can be used. These techniques can be local or global and each of them can be adaptive based on different criteria. Although correct automatic set up of threshold parameters for all kinds of maps is not possible some type of automation is achievable [3].

2.2. Pre-processing

In this step binary image is processed to remove imperfections and to amplify desired features of line objects. Pre-processing is closely associated with the processing step. Binary image should be improved accordingly to weaknesses of techniques used in processing step to prevent future errors. Imperfections in binary image are caused by bad condition of paper maps, process of scanning, usage of segmentation methods, complexity of map and variety of representation possibilities. For line objects conserving connectivity and shape are two most important features that should be considered when the methods and parameters are selected. Accurate pre-processing should fulfill these tasks:

- remove isolated small objects
- reduce boundary noise (contour intrusions and protrusions)
- filling small holes in objects
- connect disconnected objects

The importance of each task can vary depending on processing step as will be described later. In my experiments good results were obtained using the binary morphology operators opening and closing [4].

2.3. Processing

In this step binary image is prepared for raster-to-vector conversion. Features like thickness of objects, information about their contours and number of components are extracted. Also some errors created in segmentation step can be fixed e.g. removing incorrectly segmented text. For processing step more methods can be used and some level of feed back is possible by measuring quality of result and analyzing extracted features.

Skeletonization techniques which produce skeleton are used for line objects. Skeleton is ideal for this type of objects because it is represented by set of one pixel thick lines which are natural representation of line objects such as roads. Skeletonization should fulfill these requirements:

- Skeleton should be one pixel thick
- Connectivity should be preserved
- Shape and position of the junction points should be preserved
- Skeleton should lie in the middle of a shape (medial axis)
- Skeleton should be immune to noise (especially to boundary noise)
- Excessive erosion should be prevented (length of lines and curves should be preserved)

There are many papers about different skeletonization techniques [5, 6, 7]. Some of these techniques can directly produce vectors and don't need to use another algorithm for raster-to-vector conversion. Each technique has its pros and cons [5] but for line objects thinning method can provide good results [8]. Thinning algorithms conserve connectivity and shape of the objects which are most important features for this task. On the other hand thinning is sensitive to boundary noise and to other imperfections like holes in objects. These problems can be solved by binary morphology operators recommended in previous chapter.

2.4. Raster-to-vector conversion

Conversion to vector form for one pixel thick skeleton can be performed in two main steps: nodes recognition and segments extraction. Local approach where 3x3 neighborhood of each pixel is inspected can be used. In these case candidates for nodes can be recognized by having connectivity number (CN) > 2 (number of black to white transitions in 3x3 neighborhood). Nodes are selected from these candidates based on additional rules. Special case of nodes are end points which can be recognized by having CN = 1. In next step segments are extracted. Segment pixels have CN = 2 so they can be easily traced.

Some of the conversion techniques can directly recognize straight lines and arcs and also some level of approximation can be done in this step.

2.5. Post-processing

After raster-to-vector conversion is made vector data usually contain large number of vertices that can be reduced by some kind of polygonal approximation. Also straight lines and arcs can be recognized in this phase. Vector data can be used to perform additional processing including pruning, removing incorrectly separated objects, improving quality of junction points and recognizing attributes. In this step highest level of control is accomplished.

3. Conclusion

In this paper general vectorization process for line objects in maps was proposed. For each of its step a short discussion and basic recommendations were given. To create more specific vectorization process some methods were developed. In this specific process thinning which create skeleton is the most important method and all other methods should respect its features. Thinning algorithms conserve connectivity and shape of line objects which are the most important requirements for this process. On the other hand thinning algorithms are sensitive to boundary noise and sometimes produce distortions in junction points. The result of the proposed specific vectorization process is shown in Fig. 2.

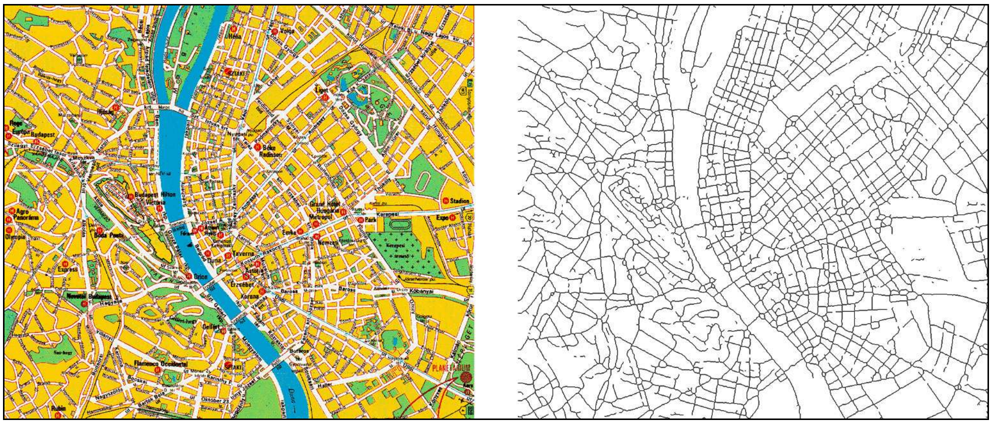


Fig. 2. Result of proposed vectorization process.

Acknowledgement

This work has been supported by grant VEGA No. 1/3775/06.

References

[1] MARTIN, A., TOSUNOGLU, S. *Image processing techniques for machine vision.* Conference on Recent Advances in Robotics, 2000.

[2] TRIER, I. D., TAXT, T. *Evaluation of binarization methods for document images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995.

[3] TARÁBEK, P. *Pre-processing road maps with threshold.* Transcom, 2007.

[4] TARÁBEK, P. *Morphology image pre-processing for thinning algorithms.* Journal of Information, Control and Management Systems, Vol. 5, No.1, 2007.

[5] TOMBRE K. *Vectorization in graphics recognition: To thin or not to thin.* Proceedings of 15th International Conference on Pattern Recognition. Volume 2., 2000.

[6] PENG, J. *An efficient algorithm of thinning scanned pencil drawings.* Journal of Image and Graphics.

[7] ZOU, J. J., HONG, Y. *Vectorization of cartoon drawings.* Selected papers from the Pan-Sydney workshop on Visualisation - Volume 2, 2000.

[8] NG, G. S., ZHOU, R. W., QUEK, C. *A Novel Single Pass Thinning Algorithm.* IEEE Transaction on System Man and Cybernetics, 1994.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Video Quality Evaluation

[*]Miroslav Uhrina, [*]Martin Vaculík
[*]University of Žilina, Faculty of Electrical Engineering, Department of Telecommunications and Multimedia, Univerzitná 1, 01026 Žilina, Slovakia, {miroslav.uhrina, martin.vaculik}@fel.uniza.sk

Abstract. The article focuses on the problem of video quality. In the first part of this article subjective assessment is described. The second part deals with objective assessment and metrics.

Keywords: subjective assessment, objective assessment, metrics, test sequence.

1. Introduction

The rapidly expanding field of multimedia technology exploits the image information of varying quality. The image and also the video quality are one of the most important parameters in the field of multimedia technology, related to the overall Quality-of-Service (QoS). The fixed part of multimedia is the audio and its relation to video, so that the overall quality can not be evaluated without complex evaluation of the content. The compression technology and the transmission link imperfection (packet loss, delay, jitter) are the most common factors that influence the overall quality.

2. Subjective assessment

The “subjective quality assessment“, consists of the use of human observers, who should score video quality during experiments called "quality assessment tests“. Since video quality is a subjective notion, subjective quality assessment can be considered as the best. But it is a time consuming method for measuring video quality and human resource is needed.

2.1. Quality factors

In order to be able to design reliable visual quality metrics, we have to understand what “quality” means to the viewer. A viewer’s enjoyment when watching a video depends on many factors [2]:

- Individual interests and expectations (everyone has his favourite programs).
- Display type and properties (CRT screens, LCD’s, plasma displays, front and back projection using various technologies — with different characteristics in terms of brightness, contrast, colour rendition, response time etc.).
- Viewing conditions (the viewing distance, the ambient light, the exterior light).
- The fidelity of the reproduction.
- The accompanying soundtrack (the sound should be synchronized with the video).

2.2. Subjective methods

<u>Double Stimulus Continuous Quality Scale (DSCQS)</u>

In this method viewers are shown multiple sequence pairs consisting of a “reference“ and a “test“ sequence, which are rather short (typically 10 seconds). The reference and test sequence are presented twice in alternating fashion, with the order of two chosen randomly for each trial. Subjects are not informed which the reference and which the test sequence is [1]. They rate each other separately DSCQS is the preferred method when the quality of test and reference sequence is similar, because it is quite sensitive to small differences in quality.

<u>Double Stimulus Impairment Scale (DSIS)</u>

As opposed to the DSCQS method, the reference is always shown before the test sequence, and neither is repeated. Subjects rate the amount of impairment in the test sequence [1]. The DSIS method is well suited for evaluating clearly visible impairments such as artefacts caused by transmission errors.

<u>Single Stimulus Continuous Quality Evaluation (SSCQE)</u>

Instead of seeing separate short sequence pairs, viewers watch a program of typically 20-30 minutes duration which has been processed by the system under test; the reference is not shown. Using a slider, the subjects continuously rate the instantaneously perceived quality. [1]

For all of these methods, the ratings from all observers (a minimum of 15 is recommended) are the averaged into a Mean Opinion Score (MOS), which represents the subjective quality of a given clip.

3. Objective assessment

The second group of video quality evaluation methods is called “objective quality assessment“. These methods consist of the use of a computational method called “metric“ (or “quality metric“) which produces values that score video quality.

3.1. Metric classification

Based on the amount of information required about the reference video quality metrics can be classified into the following categories:

- Full-reference (FR) metrics – sometimes referred to as fidelity metrics – perform a frame-by-frame comparison between a reference video and the video under test; they require the entire reference video to be available, usually in uncompressed form, which is quite an important restriction on the practical usability of such metrics.
- No-reference (NR) metrics look only at the video under test and have no need of reference information. This makes it possible to measure video quality anywhere in an existing compression and transmission system, for example at the receiver side of television (TV) broadcasts or a video streaming session. The difficulty here lies in telling apart distortions from actual content, a distinction humans are usually able to make from experience.
- Reduced-reference (RR) metrics lie between the above two extremes. They extract a number of features from the reference video (e.g. the amount of motion or spatial detail), and the comparison with the video under test is then based only on those features. This makes it possible to avoid some of the pitfalls of pure no-reference metrics while keeping the amount of reference information manageable [2].

3.2. Metrics

The mean squared error (MSE) and the peak signal-to-noise ratio (PSNR) are the most popular difference metrics in image and video processing.

The MSE is the mean of the squared differences between the grey-level values of pixels in two pictures or sequences I and $\tilde{I}$:

$$MSE = \frac{1}{TXY}\sum_t \sum_x \sum_y \left[I(t,x,y) - \tilde{I}(t,x,y)\right]^2 \tag{1}$$

for pictures of size X x Y and T frames in the sequence.

The PSNR in decibels is defined as:

$$PSNR = 10\log\frac{m^2}{MSE} \tag{2}$$

where m is the maximum value that a pixel can take (e.g. 255 for 8-bit images).

Also other metrics are known:

- MSAD – the value of this metric is the mean absolute difference of the colour components in the correspondent points of image. This metric is used for testing codecs and filters [5].
- Delta – the value of this metric is the mean difference of the colour components in the correspondent points of image. This metric is used for testing codecs and filters [5].
- Blurring Measure – this metric allows you to compare power of blurring of two images [5].
- Blocking Measure – this metric was created to measure subjective blocking effect in a video sequence. This metric also contains heuristic method for detecting objects edges, which are placed to the edge of the block. In this case, metric values is pulled down, it allows measuring blocking more precisely [5].
- SSIM Index – is based on measuring the structural information degradation, which includes three comparisons: luminance, contrast and structure. It's defined as [3]:

$$SSIM = l(x,y).c(x,y).s(x,y) \tag{3}$$

where $l(x,y)$ is Luma comparison, $c(x,y)$ is Contrast comparison and $s(x,y)$ is Structure comparison. They are defined as:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1},\ c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$
$$s(x,y) = \frac{(\sigma_{xy} + C_3)}{\sigma_x + \sigma_y + C_3} \tag{4}$$

where x and y are two nonnegative image signals to be compared, μ_x and μ_y are the mean intensity of image x and y respectively, σ_x and σ_y are the standard deviation of image x and y respectively, σ_{xy} is the covariance of image x and y.

In fact, without C_3, the equation with an element $s(x, y)$ is the correlation coefficient of image x and y, and C_1, C_2 and C_3 are small constants to avoid the denominator being zero. It is recommended by:

$$C_1 = (K_1 L)^2,\ C_2 = (K_2 L)^2,\ C_3 = C_2 / 2 \tag{5}$$

where $K_1, K_2 << 1$ and L is the dynamic range of the pixel values (255 for 8-bit greyscale images). In addition, the higher of $SSIM(x, y)$ is, the more similar the image x and y are.

- VQM – this metric is based on the Discrete Cosine Transform (DCT) [5]. It incorporates aspects of early visual processing, including light adaptation, luminance and chromatic channels, spatiotemporal filtering, spatial frequency channels, contrast masking and probability summation. VQM uses DCT in order to perform decomposition of the original data into spatial channels. This provides a powerful advantage towards the implementation of this metric, since efficient hardware and software are available for this transformation and because in many applications the transform may have already been done as part of the compression process [4].

4. Conclusion

This article has presented basic information about video quality assessment. We mentioned some subjective and objective methods used for evaluating the video quality. Also some metrics has been brought to mind. In the future, the best method and metric will be chosen and used for measuring and evaluating the video quality, i.e. for measuring the differences between video codecs and determining the best video codec for TV transmission and as well for measuring and simulating the influence of the features of the transmission channel on the TV transmission.

References

[1] WINKLER, S. *Digital Video Quality: Vision Models and Metrics.* John Wiley & Sons Ltd, 2005, ISBN 0-470-02404-6.

[2] WU, H.R., RAO, K.R. *Digital Video Image Quality and Perceptual Coding.* Taylor & Francis Group, 2006, ISBN 978-0-8247-2777-2.

[3] ZHI-YI MAI, CHUN-LING YANG, LAI-MAN PO, SHENG-LI XIE. *A New Rate-Distortion Optimization Using Structural Information in H.264 I-Frame Encoder.* http://www.ee.cityu.edu.hk/~lmpo/publications/2005_NLSC_SSIM_H264.pdf.

[4] H. KOUMARAS, G. GARDIKIS, A. KOURTIS, D. MARTAKOS. *Quantitative Perceptual Comparison of Variable Bit Rate over Constant Bit Rate Encoding Scheme for MPEG-4 Video.* http://aias.iit.demokritos.gr/~koumaras/Quantitative%20Perceptual%20Comparison%20of%20VBR%20over%20CBR%20Encoding%20Scheme%20for%20MPEG-4%20Video.pdf.

[5] *MSU Quality Measurement Tool: Metrics information.* http://compression.ru/video/quality_measure/info_en.html.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Performance of Parallel Queries on Distributed Databases and Data Warehouses

*Eva Valentíková, *Matúš Chochlík
[*] University of Žilina, Faculty of Management Science and Informatics, Department of Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovakia, {eva.valentikova, matus.chochlik}@fri.uniza.sk

Abstract. Unlike common transactional databases, data warehouses store especially read-only data for analytical processing. Due to performance, historical, organizational and other reasons some of these warehouses are distributed and require query distribution and parallelization. The aim of this article is to compare the efficiency of parallel query execution on distributed transactional database with execution of equivalent queries on a distributed data warehouse.

Keywords: Parallel SQL queries, parallel programming, distributed databases, distributed data warehouses, MPI, load balancing.

1. Introduction

Data warehouse is a repository for massive amounts of data from different sources or systems. They are usually used for analytical tasks such sales statistics, cost statistics and many other analytical requirements. Unlike common transactional databases (OLTP) they store more and especially historical data, which are not common for transactional databases. The idea of data warehouse [6], [7] is to extract data from operational databases and store them separately. One of the major differences between a data warehouse and a transactional database is in the data access and modification patterns.

Transactional databases are relational databases with highly normalized tables, largely optimized for common operations like inserts, deletions and updates. They deal with daily produced data and users can manipulate them. Data warehouses are designed for analytical purposes and the tables are de-normalized. They contain historical data for analysis received from transactional databases or other sources. Data are read-only and only select statements are allowed for users, so users should not be able to manipulate data there. To design data warehouses the Star Schema or the Snowflake Schema is used in most cases [6], [7].

Usually there is not enough space for storing a large data warehouse on one server. That's why distributed data warehouses are used in such cases. "In distributed data warehouses architecture, multiple nodes are connected to each other and may freely share data issue aggregation queries against other nodes participating in a distributed scenario." [8].

Furthermore in some cases the data warehouse is naturally distributed for historical, organizational or other reasons. In the aforementioned cases parallel and distributed queries are required.

1.1. Related work

There are many publications, for example [1], [8], [11], [13], concerning distributed data warehouses, but we didn't find any comparison of parallel execution of equivalent queries on warehouses and transactional databases.

1.2. Motivation

The goal of this article is to compare number of parallel query executed on distributed transactional database with execution of the same queries on distributed data warehouse schema (to compare queries returning the same results, but on another schemata). We also investigate the advantages of parallel query execution on several nodes, compared to serial query executed on a single node.

2. Testing environment

The measurements were performed on a network of workstations (NOW) [5] which were however not specifically dedicated for parallel programming experiments. The individual computers used in this experiment were regular computer laboratory workstations used for educational purposes, connected via a standard 100 Mbit Ethernet network, all with the following, identical, hardware configuration: single core Intel® Celeron Processor @2,8 GHz, 512 MB of system memory (8 MB shared with the graphics card) and a 80 GB HDD. The total number of 20 computers was used in the test. Each of the machines was running the Windows XP Professional SP3 operating system. We used MS SQL Server as the database system and the MPICH2 parallel programming middleware [8] as the communication facility employed in the parallel query execution and result gathering.

2.1. Test database and data warehouse

For our measurements we have used two database schemas. One transactional database schema (named CEHZ) and one „family of stars" schema for data warehouse (named CEHZ_DW). In [14] can find more about schemas designed as family of stars. Both of these databases store data of same character. Each machine's SQL server stored identical replicas of both the database and the data warehouse.

2.2. Test suite

Our test applications employed two types of processes. A single instance of a *Manager* process which coordinated a variable number of *Worker* processes. The manager was responsible for test query preparation and distribution among the workers and for query result gathering and presentation. The workers were executing the queries and providing the manager with the results. The tests were focused on measuring the execution times of a set of SQL queries, depending on the number of used worker processes and on various properties of the queries. The final goal was to determine the *Speedup* (1) and *Efficiency* (2) [15] of the parallel execution compared to sequential execution of the same test queries with N workers:

$$Speedup_N = \frac{ExecutionTime_1}{ExecutionTime_N}. \tag{1}$$

$$Efficiency_N = \frac{Speedup_N}{N}. \tag{2}$$

The measurements were performed in two basic scenarios:

- Load balancing (LB): Individual queries of a large batch (consisting of several hundreds of queries) were distributed by the manager among the workers. The execution time of the whole batch was measured.

- Query parallelization (QP): In this scenario, the manager semi-automatically preprocessed the test queries before assigning them to the workers. Each worker selected a disjoint subset of rows which were combined into the resulting dataset by the manager. Here, the execution time of the parallelized query was measured.

Furthermore, the tests were performed separately on the CEHZ database (DB) and the CEHZ_DW data warehouse (DW) and the results for equivalent test cases were compared.

2.3. Test cases

In both of the scenarios we have executed several test cases with different query sets, to determine the impact of the following two query characteristics on the parallel speedup:

- Query run-time
    - Short (TS): less than 10 sec.
    - Medium (TM): 10 to 30 sec.
    - Long-running (TL): more than 30 sec.
- Resulting dataset size
    - Small (SS): less than 1000 values
    - Medium (SM): 1000 to 100 000 values
    - Large (SL): more than 100 000, up to several million values

3. Results

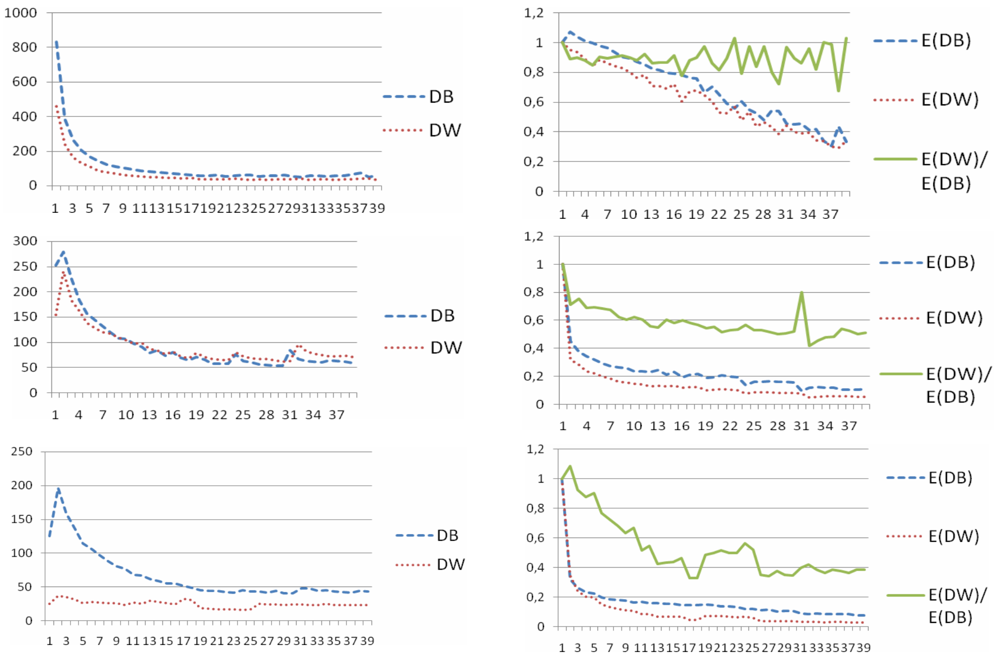


Fig. 1. Results of the measurements and the calculated parallelization efficiency.

The charts in the left column of the figure 1 show the execution times, charts on the right show the parallelization efficiency and the comparison of efficiencies for the database and the warehouse, depending on the number of workers.

The first row contains the results of the load balancing scenario with medium to long running queries, returning small datasets. The second row contains the results of query parallelization with medium run-times returning small datasets and the third pair of charts shows the results of query parallelization with short run-times returning small sized datasets.

4. Conclusion and future work

The results suggest that the load balancing and parallelization is efficient for queries with medium to long run-times, which are returning small datasets. These kinds of queries are very common in analytical data processing using aggregation, performed on data warehouses.

As the third pair of charts on figure 1 suggests, parallelization of short running queries is not effective for the data warehouse and quickly drops with the regular database too. Other measurements show that the efficiency also drops if the resulting datasets are large, because the manager becomes a bottleneck and is unable to handle the incoming data from the workers in appropriate time.

Generally, the efficiency of parallelization and load balancing is similar for both the regular database and the data warehouse, but is slightly worse for the latter, because warehouses are already optimized for the aggregating queries.

Future work in this area may be focused of the automation and optimization of the query parallelization technique. In the measurements made for this article we have used the naturally indexed attributes for the separation of the individual sub-queries. The rows are however not evenly distributed which has negative impact on the parallelization efficiency.

References

[1] BAUER, A., LEHNER, W., *On Solving the View Selection Problem in Distributed Data Warehouse Architectures,* International Conference on Scientific and Statistical Database Management, 2003

[2] GRONDŽÁK, K., MARTINCOVÁ, P., CHOCHLÍK, M., *Performance Analysis of Parallel Algorithm for Backtracking.* 4th International Workshop on Grid Computing for Complex Problems.

[3] GRONDŽÁK, K., MARTINCOVÁ, P., CHOCHLÍK, M., *Static Job Scheduling in the Grid.* 4th International Workshop on Grid Computing for Complex Problems.

[4] GUPTA, K. C., GARG, R., BAHL, I., BHARTIA, P. *Microstrip Lines and Slotlines*. 2nd ed. Norwood: Artech House, 1996.

[5] HANULIAK, I., *Paralalné architektúry*, Knižné centrum v Žiline, ISBN 80-88723-64-7, 1997.

[6] INMON, W., *Building the Data Warehouse*, Wiley & Sons, New York, 1996

[7] KIMBALL, R., *The Data Warehouse Toolkit*, John Wiley & Sons, New York, 1996

[8] MA, H., SCHEWE, K., ZHAO, J. *Cost Optimization for Distributed Data Warehouses*. Hawaii International Conference on System Sciences, 2005

[9] MPICH-A Portable Implementation of MPI, http://www.mcs.anl.gov/mpi/mpich/

[10] NICOLE, R. *Title of paper with only first word capitalized*. J. Name Stand. Abbrev., in press.

[11] STÖHR, T., HOLGER, M., RAHM, E., *Multi-dimensional Database Allocation for Parallel Data Warehouses.* Proceedings of the 26th International Conference on Very Large Databases, Cairo, Egypt, 2000.

[12] VALENTÍKOVÁ, E., ZÁBOVSKÁ, K., ZÁBOVSKÝ, M.: *Dátový sklad na vlastnej koži*, Systémová integrácia 2008, zborník z medzinárodnej konferencie, Slovenská spoločnosť pre systémovú integráciu, Žilina, 2008, str. 78-84, ISBN 978-80-970057-0-2.

[13] ZHAO, J., SCHEWE, K. *Using Abstract State Machines for Distributed Data Warehouse Design. A,* First Asia – Pacific Conference on Conceptual Modelling, 2004.

[14] *Dátové sklady - Data Warehousing*, http://kaia.mtf.stuba.sk/predmety/vis/prednaska_dw.pdf , 12.03.2009

[15] Wikipedia, Speedup in parallel computing, http://en.wikipedia.org/wiki/Speedup

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Search Engine Optimization

*Ing. Martin Vaško, *Ing. Michaela Gajňáková
*University of Žilina, Faculty of Operation and Economics of Transport and Communications,
Department of Communications, Univerzitna 2, 01026 Žilina, Slovakia,
{martin.vasko, michaela.gajnakova}@fpedas.uniza.sk

Abstract. Search Engine Optimization (SEO) is often considered the more technical part of Web marketing, which encompasses a wide variety of tasks that improve a website's presence on search engines. The article focuses on some basic tasks performed by search engines on the internet and also on basic factors, which affect on web page ranking in their search result pages.

Keywords: SEO, search engines, robots, crawler, optimization.

1. Introduction

The Internet is a worldwide network of computers that contains millions of pages of information. Everyday, millions of people search the web to find out what they look for.

A lot of people spent a good deal, worked hard for many days and nights and built a great website and want of their website to come up top in the results of search engine. There are millions upon millions of pages of web content out there and website is totally lost in the shuffle, like the proverbial needle in a haystack. When search engines ignore site, site becomes non-existent in the cyber world. The real problem with website is that it failed to harness the most cost-effective and powerful Internet marketing strategy: Search Engine Optimization! [1]

2. What is SEO

Search Engine Optimization is often considered the more technical part of Web marketing. This is true because SEO does help in the promotion of sites and at the same time it requires some technical knowledge – at least familiarity with basic HTML. SEO is sometimes also called SEO copyrighting because most of the techniques that are used to promote sites in search engines deal with text. Generally, SEO can be defined as the activity of optimizing Web pages or whole sites in order to make them more search engine-friendly, thus getting higher positions in search results.

Simply put, Search Engine Optimization is about making website visible on search engines. SEO is important not only because it brings lots of visitors to website, but also because it helps to increase the return on investment, if harnessed properly. [2] SEO encompasses a wide variety of tasks that improve a website's presence on search engines. [3]However, Search Engine Optimization is a process that requires patience, careful planning, and a long-term approach.

One of the basic truths in SEO is that even if someone does all the things that are necessary to do, this does not automatically guarantee him top ratings but if person neglects basic rules, this certainly will not go unnoticed.

Although SEO helps to increase the traffic to one's site, SEO is not advertising. Of course, people can be included in paid search results for given keywords but basically the idea behind the SEO techniques is to get top placement because their site are relevant to a particular search term, not because they pay.

SEO can be a 30-minute job or a permanent activity. Sometimes it is enough to do some generic SEO in order to get high in search engines – for instance, if someone is a leader for rare keywords, then he do not have a lot to do in order to get decent placement. But in most cases, if customer really want to be at the top, he need to pay special attention to SEO and devote significant amounts of time and effort to it. Even if person plan to do some basic SEO, it is essential that he understand how search engines work and which items are most important in SEO. [2]

3. How search engines work

Search engines are the key to finding specific information on the vast expanse of the World Wide Web. Without sophisticated search engines, it would be virtually impossible to locate anything on the Web without knowing a specific URL. We could say, that Internet search engines are special sites on the Web that are designed to help people find information stored on other sites.

There are differences in the ways various search engines work, but they all perform three basic tasks:

- They search the Internet (or select pieces of the Internet) based on important words.
- They keep an index of the words they find, and where they find them.
- They allow users to look for words or combinations of words found in that index.

Early search engines held an index of a few hundred thousand pages and documents, and received maybe one or two thousand inquiries each day. Today, a top search engine will index hundreds of millions of pages, and respond to tens of millions of queries per day. [4]

But what the search engine really is and how does it work? And what makes some search engines more effective than others?

Nowadays, when people use the term *search engine* in relation to the Web, they are usually referring to the actual search forms that search through databases of HTML documents, initially gathered by a robot. There are basically three types of search engines:

- **Crawler-based search engines** (powered by robots, called crawlers; ants or spiders) are those that use automated software agents (called crawlers) that visit a Web site, read the information on the actual site, read the site's meta tags and also follow the links that the site connects to performing indexing on all linked Web sites as well. The crawler returns all that information back to a central depository, where the data is indexed. The crawler will periodically return to the sites to check for any information that has changed. The frequency with which this happens is determined by the administrators of the search engine. Here is important to realize, that crawler-based search engines are not humans. Unlike them, these search engines are text-driven. Although technology advances rapidly, these search engines are far from intelligent creatures that can feel the beauty of a cool design or enjoy the sounds and movement in movies. Instead, crawler-based search engines crawl the Web, looking at particular site items (mainly text) to get an idea what a site is about.
- **Human-powered search engines** rely on humans to submit information that is subsequently indexed and catalogued. Only information that is submitted is put into the index. A human-powered directory, such as the Open Directory, depends on humans

for its listings. Web creator submits a short description to the directory for his entire site, or editors write one for sites they review. A search looks for matches only in the descriptions submitted. Changing web pages has no effect on visitors listing. Things that are useful for improving a listing with a search engine have nothing to do with improving a listing in a directory. The only exception is that a good site, with good content, might be more likely to get reviewed for free than a poor site.

- **Hybrid search engines** are those that are a hybrid of the previous two types. In the web's early days, it used to be that a search engine either presented crawler-based results or human-powered listings. Today is extremely common for both types of results to be presented.

3.1. Robots deliver

How we can see above, for purposes of SEO are the most interesting crawler-based search engines, which have three major elements. First is the spider, also called the crawler. The spider visits a web page, reads it, and then follows links to other pages within the site. This is what it means when someone refers to a site being "spidered" or "crawled." The spider returns to the site on a regular basis, such as every month or two, to look for changes.

Everything the spider finds goes into the second part of the search engine, the index. The index, sometimes called the catalog, is like a giant book containing a copy of every web page that the spider finds. If a web page changes, then this book is updated with new information.

Sometimes it can take a while for new pages or changes that the spider finds to be added to the index. Thus, a web page may have been "spidered" but not yet "indexed." Until it is indexed (added to the index) it is not available to those searching with the search engine.

Search engine software is the third part of a search engine. This is the program that sifts through the millions of pages recorded in the index to find matches to a search and rank them in order of what it believes is most relevant.

All crawler-based search engines have the basic parts described above, but there are differences in how these parts are tuned. That is why the same search on different search engines often produces different results. [4][5]

3.2. How to build a high traffic website

Search engines are mainly focused on next few factors that affect on web page ranking and position on Search Engine Result Page (SERP):

- **Keywords** are the most important SEO item for every search engine – actually they are what search strings are matched against. From this point of view, is important to put right keywords to right place: to <title> tag, URL, document text, anchor text, headings, <alt> tags, metatags and so on.
- Very important factor what is needed to check is keyword density in document text. 3-7 % for major keywords is best, 1-2 for minor. Keyword density of over 10% is suspicious and looks more like keyword stuffing, than a naturally written text.
- **Links** are extreme important for good page ranking. Having inbound and outbound links from/to similar sites is very useful. Also, is good to have internal links between particular pages of the site.
- **Metatags** are becoming less and less important but if there are metatags that still matter, these are the <description> and <keywords> ones. Is useful to use the <description> metatag to write the description of site. Besides the fact that metatags still rock on some search engines such as MSN and Yahoo!, the <description> metatag has one more advantage – it sometimes pops in the description of the site in search

results. The <keywords> metatag also matters, though as all metatags it gets almost no attention from Google and some attention from MSN and Yahoo!.

- **Content**. Having more relevant content, which is different from the content on other sites (both in wording and topics), is a real boost for site's rankings. When a keyword in the document text is in a larger font size in comparison to other on-page text, this makes it more noticeable, so therefore it is more important than the rest of the text. The same applies to headings (<h1>, <h2>, etc.), which generally are in larger font size than the rest of the text. Bold and italic are another way to emphasize important words and phrases.
- **Visual extras** as a Java Scripts, images, videos, frames and flash. If they are used wisely, it will not hurt. But if the main content is displayed through these extras, this makes it more difficult for spiders to follow and in some cases spiders can't follow it at all. This can definitely hurt site's ratings. [2]

4. Conclusion

Although Search Engine Optimization seems to be so complicated at the outset, it really is a simple and interesting process. Search Engine Optimization is the most cost-effective, easy to implement Internet marketing strategy that can get more traffic to optimized site and in effect more revenue.

References

[1] http://www.rankquest.com/seotips/2005/10/what-is-seo-and-why-optimize.html

[2] http://www.webconfs.com/seo-tutorial/introduction-to-seo.php

[3] Couzin, G., Grappone, J. Search Engine Optimization AN HOUR A DAY. 2nd ed. Wiley Publishing, Inc. 2008.

[4] http://searchenginewatch.com/2167891

[5] http://www.webopedia.com/DidYouKnow/Internet/2003/HowWebSearchEnginesWork.asp

[6] http://computer.howstuffworks.com/search-engine-optimization4.htm

[7] http://www.monash.com/spidap4.html

[8] http://www.webconfs.com/seo-tutorial/seo-and-keywords.php

[9] COUSIN, G., GRAPPONE, J. SEO – Search Engine Optimization. Zoner Press, 2007

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Implementing Gateway Mobile Location Center

*Petr Vlacil, *Robert Bestak
*Czech technical university in Prague, Faculty of Electrical Engineering, Department of Telecommunications Engineering, Techicka 2, 16000 Prague 6, Czech Republic
{vlacip1, bestar1}@fel.cvut.cz

Abstract. Location based services provide new opportunities for service providers to attract customers. Main part of location base services in GSM network is Gateway Mobile Location Center. The essential task of this signaling node is to provide a position of mobile terminals to the location clients that request it. The node also represents a gateway from the IP network side to the mobile network. This article describes implementation of GMLC entity into mobile networks.

Keywords: Location Based Services, Mobile Location Protocol, XML-RPC-C, Gateway Mobile Location Center, Abyss webserver, libxml2.

1. Introduction

Standardized way to implement location based services in mobile networks is to upgrade the radio access network with localization procedures such as Time Difference of Arrival (TDOA) or Enhanced Observed Time Difference (E-OTD). The next step is to implement GMLC (Gateway Mobile Location Center). The GMLC queries position of terminals in a network. The GMLC also converts position information to the standardized form and provides it to location clients.

This article describes implementation of GMLC by using C/C++ programming language. The presented GMLC is developed for Solaris operating system and is deployed at a server that is connected to a SS7 signaling network. The server is as well as connected to an IP LAN (Local Area Network). Testing location client is installed on the server that is connected to the same LAN. Interconnection of location client and GMLC is performed using MLP (Mobile Location Protocol) [1] carried in HTTP (HyperText Transfer Protocol).

2. Implementation of GMLC

2.1. Used telecommunication standards

Location services are defined in the same way for 2G and 3G networks [2]. To provide location based services, the 3GPP updated the architecture of mobile networks by adding new signaling nodes (See figure 1). One of the most important parts of the update is entity GMLC that queries terminal positions from a network by employing different methods [3, 4]. There can be more than one GMLC in a network. Other important part of the update is a location client. The location client requests position of terminals from GMLC. There can be many location clients. A location client can be for example a service provider or subscriber himself.

Because of introduction of new nodes to the network new interfaces have to be specified. These interfaces are defined in [2]. The interface “Lh” is used for communication between GMLC and HLR (Home Location Register). It is used to obtain address of signaling node

SGSN (Serving GPRS Support Node) or MSC (Mobile Switching Center) that serve a terminal that is being localized. This interface is based on MAP (Mobile Application Part) that is the 7th layer of SS7 signaling architecture.

The “Lg” interface is used between GMLC and MSC or SGSN. This interface is used to query subscriber's location from network. Lg interface is also based on SS7 signaling. The “Le” interface is used between GMLC and location clients. It can be implemented by MLP protocol over IP [1]. The “Lr” interface interconnects GMLC in home PLMN (Public Land Mobile Network) network and other PLMN. It is used to query a location of subscriber that is connected to a different PLMN.

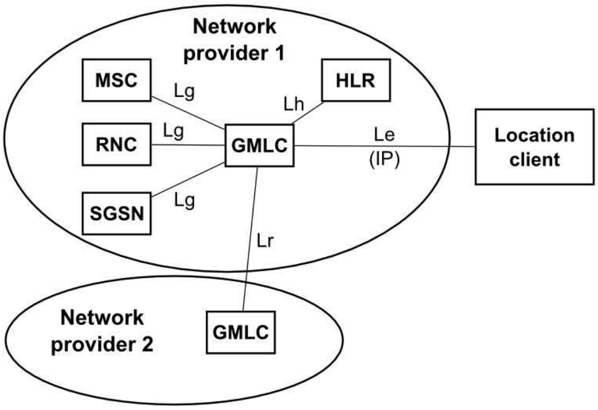


Fig. 1. Architecture of location system.

2.2. Technologies used for implementation of GMLC

Two independent software modules are cooperating together as one GMLC. Server that is running GMLC runs on SUN Solaris operating system. The server represents standalone SP (Signaling Point) in SS7 testing network and is equipped with SS7 signaling card that is connected to the testing GSM network.

The manufacturer of signaling card supplies software that implements all signaling layers of SS7 signaling system [5, 6, 7]. The software modules communicate through shared memory IPC (Inter-Process Communication). There are numbered queues created in this shared memory. Each software module has its own queue with a unique ID [8]. The software modules use asynchronous communication, where modules exchange messages via queues.

As mentioned above the GMLC is split into two modules. The first module processes queries from the Le interface and is called the mlp_module. The second module queries terminal's position through the Lg interface and is called the lbs_module.

The mlp_module is based on XML-RPC-C [9] project. This project is offering software implementation of RPC (Remote Procedure Call) based on XML (eXtension Markup Language) and is programmed in C/C++. The mlp_module uses Abyss HTTP server that is contained in XML-RPC-C software package. The mlp_module communicates with location clients via MLP. The MLP is protocol based on XML. Libxml2 C/C++ libraries [10] are used for parsing MLP.

The lbs_module acts as a MAP application. It communicates with the MAP signaling layer of SS7 signaling card. The lbs_module uses SS7 signaling procedures to query position of terminals from the SGSN.

The position of terminal is returned back to the network in the form of CGI (Cell Global Identity [11]). This parameter contains Cell ID of cell where the terminal is situated. The

GMLC uses a database of cells to find out the geographical position of network cell. MySQL database is used.

2.3. GMLC architecture

The GMLC is composed of two modules, lbs_module and mlp_module. Both modules are built on same programming template and classes described in [12] so they are manageable in the same standard way. These classes provide modules with ability to be configured by XML configuration files during their initialization or by configuration messages when already running.

The GMLC uses timer module for measuring timeouts. Timer module is another software module that is maintaining timer requests from other software modules. Every software module can register its timer to the common timer module. The timer module sends timeout messages to the given modules when their timer expires. The timer module is described in more details in [12].

2.4. Function of GMLC

This sections describes how the software modules cooperate together to act as a GMLC. The lbs_module waits for request from the mlp_module. Each request contains IMSI (International Mobile Subscriber Identity) of subscriber. When obtaining a new request the lbs_module creates a new MAP dialog, registers its timer and tries to find out the terminal's position from the SS7 network. The terminal’s position is requested form SGSN by MAP message called MAP-PROVIDE-SUBSCRIBER-LOCATION [13]. When answer from network arrives, appropriate dialog is found in dialog manager. SGSN answers with CGI. The Cell ID is picked up and sent to the mlp_module. Scheme of module is shown in figure 2.

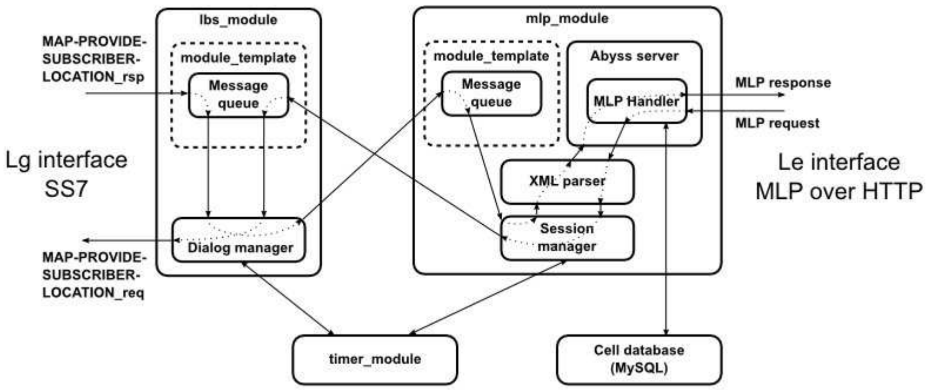


Fig 2. Scheme of lbs_module and mlp_module.

The mlp_module uses XML-RPC-C Abyss server. The Abyss HTTP server accepts requests from Location clients. Location clients request terminal's position from GMLC. Location clients send MLP requests over HTTP protocol. When new request arrives to mlp_module Abyss runs MLP handler which reads and processes MLP message using XML parser. A new session for incoming MLP message is created and stored in the session manager. Identification of subscriber is taken from the MLP message and is transmitted to the lbs_module with the session identification. The lbs_module tries to find out the terminal's position in the network. The obtained result is returned back to the mlp_module with the session identification. The mlp_module finds out the given session in the session manager. The returned Cell ID is transformed to the geographical coordinates by using MySQL Cell ID database. The MLP message is filled with the geographical coordinates and sent by Abyss server back to the location client.

3. Conclusion

This article describes an implementation of GMLC server architecture into a mobile network. The system is designed in such way that it includes independent software modules. Modules implement same basic functions and classes described in [12] so they are manageable in the same manner. The communication between the modules is done through numbered queues in a shared memory. This facilitate adding of new modules to the system.

The whole system has been developed and deployed in Research and Development Centre for Mobile Applications at CTU (Czech Technical University in Prague). The developed lbs_module was successfully tested with SGSN situated in the testing lab. The SGSN provides location of terminals in the form of Cell ID. The mlp_module is able to process the Cell ID by using MySQL database of cells and their positions. Position is then sent to location client in MLP message. Proper XML parser that reads MLP messages from clients is still under a development during writing of this article.

Acknowledgment

This research work was supported by grant of Czech Ministry of Education, Youth and Sports No. MSM6840770014.

References

[1] LIF TS 101 Specification: *Mobile Location Protocol Specification*, version 3.1, 2004

[2] 3GPP TS 23.271: *Functional stage 2 description of Location Services (LCS)*, Release 8, 2008-12

[3] 3GPP TS 25.305: *Stage 2 functional specification of User Equipment (UE) positioning in UTRAN*, Release 8, 2007-12

[4] 3GPP TS 43.059: *Functional stage 2 description of Location Services (LCS) in GERAN*, Release 8, 2008-08

[5] Dialogic®: *Intel® NetStructureTM SS7 Protocols SCCP Programmer's Manual*, 6th edition, 2003.

[6] Dialogic®: *Intel® NetStructureTM SS7 Protocols TCAP Programmer's Manual*, 6th edition, 2003.

[7] Dialogic®: *Intel® NetStructureTM SS7 Protocols MAP Programmer's Manual*, 6th edition, 2003.

[8] Dialogic®: *Intel® NetStructureTM SS7 Protocols Software Environment Programmer's* Manual, 6th edition, 2003.

[9] XML-RPC-C: <*http://xmlrpc-c.sourceforge.net/*>, online

[10] LIBXML2: <*http://xmlsoft.org/html/index.html*>, online

[11] 3GPP TS23.003: *Numbering, addressing and identification*, Realease 8, 2008-09

[12] DANIHELKA I.: *Aplikačni moduly pro SS7*, 2005

[13] 3GPP TS 29.002: *Mobile Application Part (MAP) specification*, Release 8, 2008-06

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Femtocells Integration into Mobile Networks

*Jaroslav Vrana, *Marek Neruda, *Robert Bestak
*Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Telecommunication Engineering, Technicka 2, 166 27 Prague 6, Czech Republic,
{vranaj1, nerudm1, bestar1}@fel.cvut.cz

Abstract. This paper investigates influence of femtocells to the traditional mobile network hierarchy and their features. This upcoming technology is supposed to be widely deployed in the upcoming wireless and mobile networks. These small base stations are designed to improve indoors capacity, reducing macro cell traffic or to provide services in locations without signal or with weak coverage. The paper discusses usage of femtocell and related challenges, such as interferences, frequency spectrum planning, handovers or integration of femtocells into mobile networks.

Keywords: Mobile network, femtocell, interference, handover

1. Introduction

Continuous advancement of 3G mobile networks applications and convergence to 4G networks calls up some new challenges. One of these challenges represents inside building coverage. Studies show that more over 50% of all voice calls and more than 70% of data traffic originates indoors [1]. Mobile networks 3G (e.g. UMTS) are distributed at higher frequency spectrum range, than 2G mobile networks do. This circumstance leads to worse signal propagation in buildings and through urban agglomeration in general. One of possible solutions how to provide coverage inside buildings is to place femtocell utilization.

Femtocells are based on Access Point Base Station called Home Node B (HNB) or Home eNode B (HeNB). HNB represents a very low power and small base station, designed for use in small rooms such as residential or small business environments. Femtocell can be placed only where DSL (Digital Subscriber Line) or cable connection to the core network is allowed. It consequently enables mobile service provider to extend signal coverage via broadband connection. Therefore, provider is able to improve his mobile services in locations where signal coverage was limited or unavailable.

The paper is structured as follows. Section 2 describes femtocell integration to the cellular core network and main connection and installation issues. Section 3 is focused on interferences and spectrum accuracy topics. Conclusion is described in Section 4.

2. Femtocell integration

Femtocells are small base stations as mentioned before. They increase signal strength and due to improved signal coverage they also increase hypothetical speed of data transmission. However, the femtocell installation needs an appropriate broadband connection. As femtocells are using licensed frequency spectrum, operators need to control Home Node B settings. Especially they need to check and set up power level and frequency of HNB [2].

2.1. Connection and installation

The femtocell is connected to operator’s core network via existing residential broadband service. This fact is considered like a Fixed Mobile Convergence (FMC). FMC leads to reuse of existing DSL or cable services. The customer who owns a Home Node B and broadband services should be able to manage an access list in his femtocell. There are two possible solutions of accessing to the femtocell:

- Public Subscriber Group (PSG)
- Closed Subscriber Group (CSG)

PSG means that everyone who is within the signal range of HNB (and his permission is allowed by operator) is able to connect through this HNB. On the other hand CSG provides connection only to those users who are in the access list of appropriate HNB. Users off the access list are able to connect to the specific HNB only in the case of emergency calls. There has to be software utility enabling emergency calls from all UEs (User Equipment).

2.2. Integration to core network

There are three possible femtocell integration principles (Fig. 1, Fig. 2 and Fig. 3). This section of the article describes these situations.

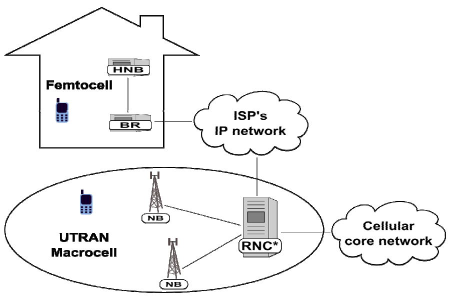


Fig. 1. HNB connection to RNC

Figure 1 shows an example of HNB that is connected throw ISP’s network (Internet Service Provider) directly to RNC entity (Radio Network Controller). This RNC differs from the basic RNC, that is why we label this RNC as RNC* in figures. Prime RNC is not designated to manage hundreds or thousands of Node Bs (NB) contrary to RNC*. The RNC* is enhanced about the HNB with HNB Gateway (HNB_GW). The 3G HNB and 3G HNB_GW in combination support all of the UTRAN (UMTS Terrestrial Radio Access Network) functions. These fundamental UTRAN functions are supported by RANAP (Radio Access Network Application Part). Requirements of new features related to HNB are supported by the new protocol HNBAP (Home Node B Application Protocol). This protocol is used between the 3G HNB and the 3G HNB_GW. Interface between HNB and HNB_GW is identified by abbreviation I_{uh}. RNC* is connected to the cellular core network throw the standard I_u interface.

Figure 2 describes an example of HNB integration based on a concentrator (CCTR). The use of CCTR makes possibility to provide solution without any intervention in the RNC. One of the main parts of concentrator is HNB_GW. In addition this gateway is connected to HNB via I_{uh} interface. The CCTR should be able to manage hundreds of HNBs.

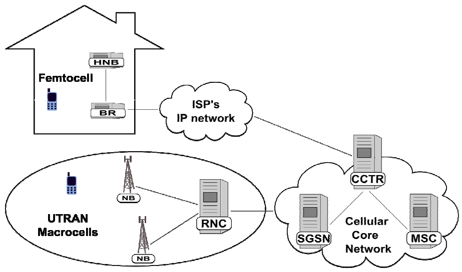


Fig. 2. Managing HNB with concentrator (CCTR)

Interface between CCTR and cellular core network is based on the type of data traffic. Packet switched data are sent from the CCTR via $I_{u\text{-}PS}$ interface to SGSN (Serving GPRS Support Node). The SGSN is an entrance point of the operator's packet domain part of network. Circuit switched data, such as call request, are transferred from the CCTR to MSC (Mobile Switching Center). The MSC is part of the operator's circuit switched network.

Figure 3 shows how the IMS/SIP (IP Multimedia Subsystem/ Session Initiation Protocol) clients are connected to the operator's network. Connection is established via SIP Application Server (SAS), which is directly connected to the operator's IMS domain and consequently to cellular core network.

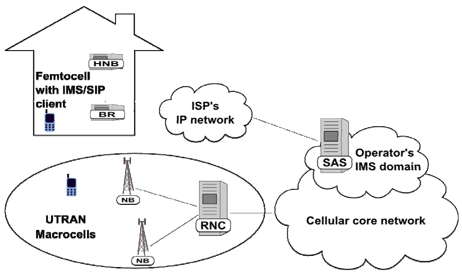


Fig. 3. Managing HNB with concetrator (CCTR) – IMS/SIP client

3. Frequency spectrum planning

Femtocells have critical effects on the performance of mobile networks. Without unique spectrum for the femtocell or very careful spectrum planning in the network, femtocells suffer from severe interference problems. This task is solved by three development alternatives. The first one is to assign different frequencies to femtocells and macro cells. The second solution is called partial co-channel, where macro cells use all frequencies and femtocells use only the part of it. The third alternative uses a full co-channel operation. In this case macro cell and femtocell share the same spectrum.

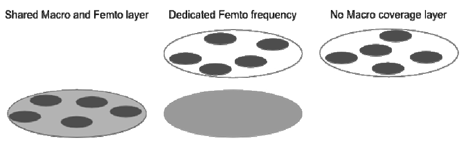


Fig. 4. Frequency spectrum planning

Interference can occur by using the same frequency band of femtocell and macro cell in a single frequency CDMA system caused by near far problem. If a mobile terminal increases transmit power to the femtocell being also in macro cell it can result in interference. Other example of this problem is the situation where neighbors use femtocells in the same area. One of the partial solutions is for example the mode-2 fixed power option which prevents from increasing power and causing interference.

3.1. Spectrum accuracy

A used spectrum of femtocell has to meet spectrum mask requirements. The femtocell must generate precise radio frequency signal, typically around 50 parts-per-billion (ppb) or better. Ppb denotes relative proportions in measured quantities. They need a very good, meant expensive, crystal oscillator for keeping this accuracy for a long time. These oscillators however require calibration every 12-to-24 months. This problem can be solved by replacing the unit or by using external, accurate signal to calibrate the oscillator. Conventional base stations often use GPS timing for synchronization. In residential environment there is crucial cost of hardware and also good GPS coverage. Another solution describes time synchronization standard providing 100 ns accuracy depending on the location of the master clock [3]. Also, NTP (Network Time Protocol) is possible solution.

4. Conclusion

3G networks use radio signal in higher frequency spectrum in comparison to 2G networks, hence this signal is worse propagated indoor. One of the possible solutions to improve indoor coverage is to place small cells inside the buildings. Femtocells have a potential to increase mobile networks coverage indoor and decrease numbers of coverage hole. The better indoor coverage the better QoS can be provided to users.

The research area of the femtocells is quite new and it offers many ways of improving. There is a big potential in network integration which can reduce operator’s costs. A lot of homes have 2-wire connection which can be used for DSL connection. It can be effectively used also for femtocells.

Acknowledgement

This research work was supported by grant of Czech Ministry of Education, Youth and Sports No. MSM6840770014.

References

[1] CHANDRASEKHAR, J., ANDREWS, J., GATHERER, A. *Femtocell networks: A survey*. Communications Magazine, IEEE, 2008.

[2] CHAMBERS, D. *Femtocell Primer: Quickly get up to speed on this new mobile phone technology*. North Carolina: Lulu, 2008. 168 s. ISBN 978-1409226475.

[3] EIDSON, J., TENGDIN, J. IEEE-1588 *Standard for a precision clock synchronization protocol for networked measurement and control systems and applications to the power industry*. In Distributech 2003. Las Vegas, 2003. s. 13.

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

The Secure Design Issue of Automated Electronic Appliances Used in Transport Infrastructure

*Ing. Monika Welterová, **Ing. Miroslav Bariak, Ing. Peter Piala
*University of Žilina, Faculty of Operation and Economics of Transport and Communications,
Department of Railway Transport, Univerzitná 1, 01026 Žilina, Slovakia,
Monika.Welterova@fpedas.uniza.sk
**University of Žilina, Faculty of Operation and Economics of Transport and Communications,
Department of Water Transport, Univerzitná 1, 01026 Žilina, Slovakia,
{Miroslav.Bariak, Peter.Piala}@fpedas.uniza.sk

Abstract. This contribution aims to review the main approaches towards physical security of electronically controlled appliances in transport processes and transport infrastructure in general. It is specially focused to transport processes in water and railway transport. It describes most common attack scenarios of physical attack to a device controller appliance and shows possible tamper response techniques.

Keywords: tamper attack, physical security, attack scenarios, secure design, tamper response

1. Introduction

During past few years, there has been an effort to bring the advantages of information technologies, advanced control of transport appliances and transport processes automation into water and railway transport technology. In case of communication environment, several broadband technologies based on wired and wireless transmissions were designed to address this issue in all common modes of transport. Generally there are huge differences in technological needs for transport appliances operating on the ground, in the air or on the water-line. The technological level of controllers, automation devices and its background based on proper and robust communication environment can indirectly affect balancing of the modal shares of transport systems in the future.

It's important to mention the issue of safety and information security in communication environment, controllers of automated transport appliances and other transport infrastructure often used even for mission critical transport applications. The secure design is nowadays usually well implemented in (mostly adopted) communication environment itself, but what concerns the automation controlling appliances in transport, the security needs are being somehow forgotten.

2. The operational environment of appliances

Once the invader gains physical access to the device itself, the security of the device depends on its construction and security implementations in its hardware and software (firmware). Physical security is very important in case a device controlling module can be directly accessed - especially in an unprotected environment. In applications of water and railway transport it has to be assumed, that each party involved in operation of appliances is in principle able to get physical access to each device controlling module.

The operational environment of an automation appliance can range from the secure server located in protected area on one side to security tokens which are handed over to the end user on the other side. It can be assumed that the secure servers are protected by some means of environmental security (e.g. guards, alarm systems, etc.), but this cannot be enforced in the case of the non-trusted device operator. Therefore the security tokens have to be constructed to be able to protect themselves well. The operational environment can be classified as protected, periodically controlled and non-protected. An example of periodically controlled environment would be the situation with random (or periodical) checking of the appliances whether they have been tampered or not.

3. Typical attack scenarios and the secure design approach

There are several attack scenarios of physical attack to a device controller appliance. These attacks can be roughly divided into groups of monitoring, manipulation, penetration, modification and substitution. Monitoring attack is a passive, non-invasive attack. Special kind of this attack is done trough the measuring/recording of the electromagnetic emission of the device (known as „listening“). Manipulation attack is also a passive and non-invasive attack, which aims to obtain a service in an unintended manner. Manipulation attacks are sometimes performed by applying abnormal environmental conditions (i.e. extreme temperature, vibrations, EM noise, radiation...). Penetration is an active, invasive attack against the appliance module. The aim is to intercept data at the internal lines or to read the memory of the device. Modification is also an active and invasive attack, but unlike the penetration attack, its aim is to modify internal connections of the device controller or the contents of its internal memory. Substitution attack is the removal of original controller module, which is then substituted by an emulating device with a modified functions or behavior. The removed module can be later used for an analysis of its internal construction.

There are two ways to be regarded in physical security design of a device controlling module. The first one aims to definitively prevent any disclosure or modification of the module. For this purpose the tamper-resistant or tamper-responsive design techniques are implemented. Tamper-responsive design leads to self-destruction of the module once an attack is detected. Tamper resistance design technique makes the module able to avert all attacks. The second way of design approach focuses on tamper-evident characteristics in case of attack. In this case the tamper evidence cannot prevent breaking into the appliance controlling module. The use of a tamper-evident protection scheme requires a control authority that periodically inspects the protected module. Typical tamper-evident designs include indicative security seals (holograms). The removal of these items should be difficult and/or leave remaining traces that can be later recognized by inspection personnel. The items should include special characteristics which are not commercially available and faking of them is difficult, expensive and/or can be easily recognized by trained personnel.

Monitoring attacks at the external interfaces of the controller module cannot be usually easily detected by the appliance itself. Unfortunately such attacks at the internal lines of the controller (done after some internal modifications during device is powered-down and off course only if there is no physical tamper protection applied) cannot be detected either. This kind of sophisticated attacks therefore cannot be avoided neither detected by common tamper-resistance nor tamper-responsive techniques described above. It is therefore necessary that these kinds of attacks based on electromagnetic or noise emissions of the controller are considered well at the design level and are made as difficult as possible for an interceptor to tamper. For the electromagnetic emission issue a special shielding might be certain solution, but even the most sophisticated tamper-protected modules cannot actively detect tamper attempts in all cases.

4. Tamper response

Once an appliance is being tampered, there comes a question of its tamper response. Besides classical invasive attacks described above, there might be some critical operational conditions applied to module which can lead to unexpected events. In this case, very common is an attack based on tampering with the power lines and environmental conditions (temperature/humidity/vibrations/EM). For their detection, special integrated sensors can be utilized by the protected module itself, which permanently supervises its operating conditions. One general requirement for such tamper response is that an internal power supply is capable of detecting and reacting to tamper attempts. It's important to notice that security-sensitive information (if any) has to be erased as fast as possible, so that the self-destruction process cannot be canceled by the invader. Therefore the destruction circuitry has to be fired before the critical state of the controller module internal power supply is reached.

At the design level of the appliance, it must be aimed to counteract the reverse-engineering as much as possible. The typical approach in the internal construction includes the encryption of internal bus lines and memory which contain critical data. During power-on self test (POST) of such tamper-resistant appliance control module it is recommended to include self tests which verify the integrity of critical internal data and the internal hardware-based random number generator (if any).

5. Conclusion

In this contribution the basic requirements for physical security of the controllers used in transport infrastructure appliances are summarized. The security of the appliance depends on its construction and security implementations in its hardware and software. Most of the attacks described above cannot be easily avoided, because detailed knowledge of the appliance combined with proper equipment and sufficient attacking time might still overcome even most sophisticated tamper-preventing solutions. Therefore the strength of physical security measures is limited in its principle. The secure design concept and the needs for adequate tamper response should be considered well at the point of transport appliance design.

References

[1] ROSS J. ANDERSON, *Security Engineering: A Guide to Building Dependable Distributed Systems,* John Wiley & Sons, Inc., 2001.

[2] STEVE H. WEINGART, *Physical Security Devices for Computer Subsystems: A Survey of Attacks and Defenses, in Cryptographic Hardware and Embedded Systems* – CHES 2000, LNCS 1965, Springer, 2000.

[3] ROSS ANDERSON, MARKUS KUHN, Tamper resistance – a cautionary note, *The Second USENIX Workshop on Electronic Commerce Proceedings,* USENIX Association, Oakland, California, November 18–21, 1996, pp 1–11, ISBN 1-880446-83-9

[4] SOSEDOVÁ, J. and col.: *Project VEGA MŠ SR 1/3831/06 – Prepravno–právne podmienky vnútrozemskej plavby a ich ekonomické dôsledky pre dopravcov po vstupe SR do EÚ,* Žilinská univerzita 2006

[5] DIRECTIVE 2005/44/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 7 September 2005 *on harmonized river information services (RIS) on inland waterways in the Community,* Official Journal L 255 , 30/09/2005 P. 0152 – 0159

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

Multiarea-eKanban Board (MKB)

Prof. Dr.-Ing. habil. W. Bärwald, Doz. Dr.-Ing. S. Baumann, Dr.-Ing. R. Keil,
Dipl.-Ing. K. Richter, Dipl.-Ing. R. Knofe
TU-Dresden, Fakultät Verkehrswissenschaften "Friedrich List", Institut für Verkehrstelematik, Professur für Verkehrsnachrichtensysteme

Abstract: Within the project "Sm@rt Logistics" and other projects as well the usability of RFID in a rough environment was demonstrated. By using RFID as an innovative enabler for an eKanban-System. But still there are some difficulties to deal with, like inductive fields of machines or a highly metallic surrounding, as well as economical problems. A solution can be provided by a modular and shielded type of eKanban Board.

Keywords: RFID, Kanban, Multiarea-eKanban Board(MKB), coupling and decoupling of antennas

1. Introduction

The Kanban process was developed by Toyota in the year 1947 and became one of the most used process controlling methods worldwide. The whole method is based on cards, the so called Kanban and can be used for the different functions production, transportation, purchase and storing.

But there are also some disadvantages. One of them is the number of cards. When there are more cards than necessary to much material is stored within the process, if there are fewer cards the process can became instable. But also if the number of cards is ideal, one main disadvantage last, the impossibility of recording data. A second one is the time between setting a card into a mailbox and the resulting action. The time is necessary to collect the cards and to find out the best reaction. Both disadvantages can be rectifying by using RFID and innovative data transmission technologies. Base is the combination of Kanban cards and transponder sat the one hand and the implementation of the RFID components inside a Kanban board at the other. To connect the RFID components with a central server, IT components are needed. Within the project "‘Sm@rt Logistics"’ one of the first so called eKanban-Systems was implemented in a factory of an automobile delivery firm. The whole system based on the RFID working frequency of 13,65 MHz. Each board is equipped with one antenna area, a RFID-reader and an interface module. The structure is made out of glass fibre reinforced plastic. The cards are combined with TTP-Label transponders of XIdent.

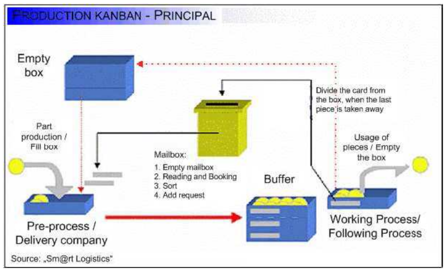


Fig. 1. Transportation Kanban Process[1] and "Sm@rt Logistics" Singlearea-eKanban Board

After the implementation and first practical tests some problems became evident. One of them was a malfunction next to a high inductive machine. Another problem is the special environment and needs in each enterprise. And so for every enterprise a special development has to be done. To resolve both problems a new concept for an eKanban-Board was developed.

2. Conception

The main aims by developing the MKB were the modular structure, an area-precisely reading of cards and a better behavior in a metallic or inductive environment. In the first step the components of the board has to be defined, in the second one the antenna to be developed and proofed and within the third step the whole board be built up.

The working frequency was set on 13,56 MHz. To get the best performance the following three concepts were compared.

- 3 Antennas with one RFID-Reader and one interface per Antenna
- 3 Antennas with one RFID-Reader, one RFID-Powersplitter and one interface for all antennas
- 3 Antennas with one RFID-Reader, one RFID-Multiplexer and one interface for all antennas

The second concept couldn’t be used in cause of the areaprecisely reading. The first concept couldn’t be realized as well. Reasons were a shutting down of all RFID-Readers in the nearer environment and the highest costs. As a result concept 3 was chosen and shown in fig. 2. The decision about the hardware components depends on the interfaces, the RFID standards (ISO 15693) and the necessary transponder energy on the field. Another important fact is the economic efficiency.

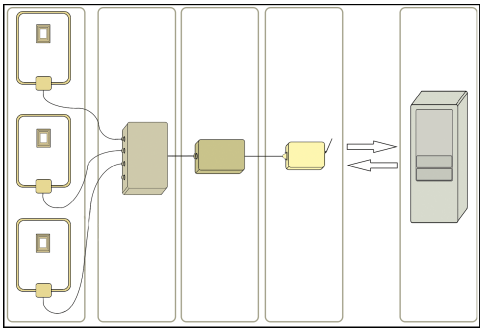


Fig. 2. Basic Multiarea-eKanban Board Concept

The following components for the tests and MKB prototype were chosen: RFID Reader (Feig ID.ISC MR200-A /-W /-EP), Multiplexer (Feig ID ISC.ANT.MUX), RFID Antennas, 3 areas pocket-table (with 24 pockets) and Kanban cards (with Transponder X-ident: TTP-Label 23 x 66-PH44).

The main task was to find a useable antenna which fulfils all the demands, like a full coverage of the area surface. But a done market analysis showed that no available antenna could stand the demands and so a special adapted antenna had to be developed and constructed. The construction based on the geometric demands between the antenna and the transponders or rather the positioning of the Kanban cards in the pocket-table. The loop antenna is positioned under the dividing line of the pocket-table, shown in fig. 2, so every transponder at the Kanban-board is getting enough energy for a safe and secure reading.

For reading out the transponders in very safe way special software, which reads more than one time before breaking up and waiting for the next term, was developed. By the software the used multiplexer gets the orders over the antenna cable from the RFID reader and is able to switch between the connected antenna areas. In this way it's also possible to read out separately all transponders from each area.

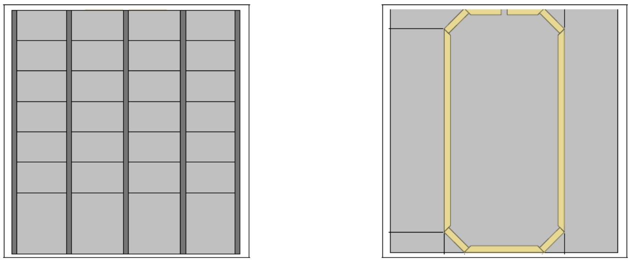


Fig. 3. Surface of an area with antenna

3. Coupling of Antenna Area

The right transponder identification of the antenna segments is necessary for the transponder separation. Therefore experiments of the identification range between a transponder and the loop antenna were done. Experiments of the identification range of a transponder with side by side positioned antenna segments are shown in fig.4.

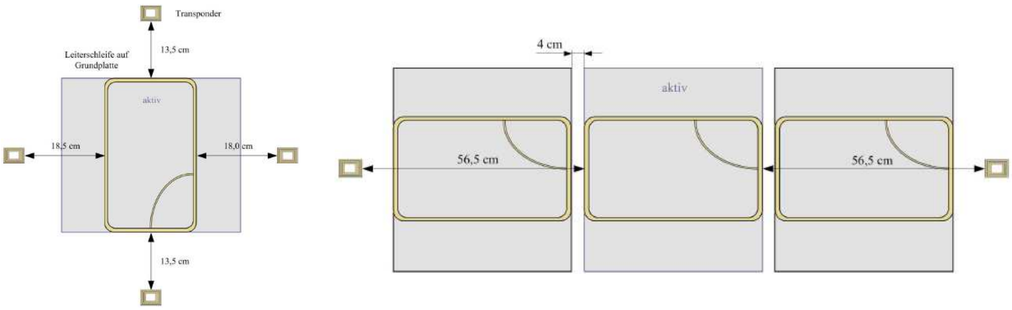


Fig. 4. Maximum transponder identification range of a single and three antenna areas

Background for the higher identification range is the coupling of the electromagnetic fields of the middle antenna segment with the other on resonance frequency calibrated antennas. This is comparable with the principle of a transformer. But it is necessary to prevent the coupling, so that the transponder separation is guarantee.

4. Decoupling of Antenna Areas

The used principle for the decoupling is the theory of eddy currents in electrically conducting materials. The electromagnetic alternating field of the antenna generates an induction current in the material. The consequence of this current is the education of an induction electromagnetic field which jams with the original field. Because of that develops a suppression of the original field. For the optimal separation of the antenna segments some experiments with different materials were done. The analyzed materials of the shield are steel, aluminium and copper. In a first experiment the influence of the antenna parameters, like resonance frequency, impedance, quality and identification rate, were analyzed.

The tests have shown that copper has the lowest influence on the antenna parameters. As well for the optimal shield material are also other factors like a high electrically conducting and magnetic permeability decisive. As a result cooper was chosen as shield material. In next experiments the positioning and the copper parameters, like thickness and copper material were tested.

The constructed 5 side copper shield allows the decoupling of the magnetic fields and the separation of the transponders in every antenna segment. Because of the shield the influence of metal objects, which normally jam the transponder identification, was minimized. Another advantage is the reduction of the influence of system strange electromagnetic fields. Therefore many experiments near induction machines, welding machines and high voltage machines were done. The complete MKB is shown in fig. 5.

Fig. 5. Front view of the MKB

5. Conclusion and Outlook

The MKB represents a new and innovative step in the direction of a flexible and adaptable eKanban-System. The basic is the modular usable antenna segment, which can also be used as singlearea eKanban board. By shielding the antennas the developed eKanban board is applicable also in highly electromagnetic polluted areas like within factory buildings. The other components like the RFID-Reader or the RFID-Multiplexer are changable and can be adapted to the special situation and needs of the company. The presented form of the eKanban board can be used for systems from 24 to 192 eKanban cards. The smallest board consists of just one antenna area and one RFID-Reader as well as one interface. The biggest board instead needs 8 antenna areas, one RFID-Multiplexer, one RFID-Reader and one interface. Additionally the robustness with regard to metallic and electromagnetic environment is much higher compared to unshielded eKanban boards. But there is also some work to do like calculating the optimal number of cards per segment and to integrate the antenna directly into the surface of a provider. After these steps an economical mass production can be established and a hugh market of Kanban using firms could be equipped. Also other segments could be interesting like presence controlling systems.

References

[1] SCHUH, G.; KAMPKER. A.; RIPP. S.: Gesamtvorhabensbeschreibung “Sm@rt Logistics”, 18.03.2005

[2] KNOFE, R.: “Prototyp eines elektronischen Mehrbereichs-Kanban-Board (MKB), Diplomarbeit, TU-Dresden, 2009

Transcom 2009, 22-24 June 2009
University of Žilina, Žilina, Slovak Republic

About Reflectometric (OTDR) Measurement of Light Tube Tracks

MSc Eng. Marcin Grabiński*, Dr. Eng. Roman Wiatr**
* University of Technology and Life Sciences, Faculty of Mechanics, Specialisation Construction and Operation of Machinery, Bydgoszcz, Poland, e-mail: mgrabinski@o2.pl
** University of Technology and Life Sciences, Postal Engineering Department, Al. Prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland, phone (+48) 523408115, e-mail: rowiat@utp.edu.pl

Abstract. This paper presents the basic details of construction and reflectometric physical phenomen used in the measurement by reflectrometer. On the chosen example the way of interpretation ot the results is discused.

Keywords: REFLECTOMETER, FIBER, RAYLEIGH SCATTERING, REFLECTION FRESNEL

1. Introduction

Nowadays information sending requires use of bigger and faster transportation channels, the solution is the use of light tubes as a means of data transportation. This research work contains information regarding methods of light tube examination to establish parameters of transportation track by reflectometer. The main measurement tools used during executive as well as operating work are OTDRs (Optical Time Domain Reflectometers). They provide clear form of obtained results and precise characteristics of optical tracks. With their use we can prescribe and verify:

- unit fibre attenuation [in dB/km],
- optical attenuation of splitable and non-splitable links [in dB],
- rear reflection attenuation of links, ie. reflectancy [in dB].

2. Principle of work of reflectometer

In principle reflectometers work like radar. OTDR sends short light impulses to the light tube and measures return pulse as a function of time. As a result of rear distribution, and reflections eg., from light tube links, part of the optical power returns and may be measured, giving length image of occurrences in the fibre. During propagation of the impulse sent from OTDR along the fibre, two phenomena occur: Rayleigh's distribution and Fresnel's reflection.

a). Rayleigh scattering

The light in the fibre is distributed in all directions; part of the power goes towards returning direction, which is used in the measurements. This phenomenon is caused by non-uniformity – fluctuations in density of the material, from which fibre is made (quartz glass), and also pollution, micro- and macro-bends, and changes in geometry.

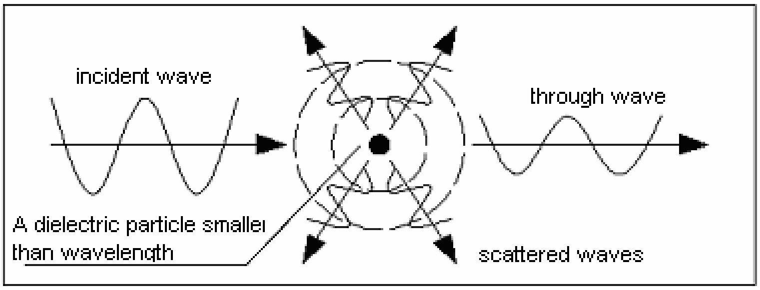


Fig. 1. Phenomenon of Rayleigh's distribution.

b). Fresnel reflection

This phenomenon occurs when the propagation light along the light tube meets sudden change in material density. These changes occur on links, welds, cracks and in places where air splits may be created. Great part of light is then reflected. Intensity of the reflection depends on differences in parameters of the bends in the neighbouring areas.

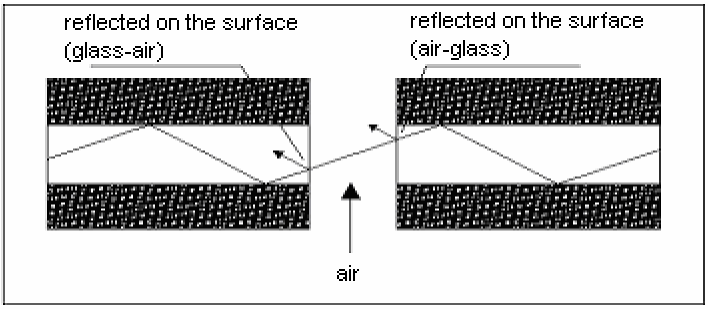


Fig. 2. Phenomenon of Fresnel's reflection.

c). Construction of reflectometer (OTDR)

Schematic construction of reflectometer is depicted in Fig. 3., incorporates the basic elements of functional devices.

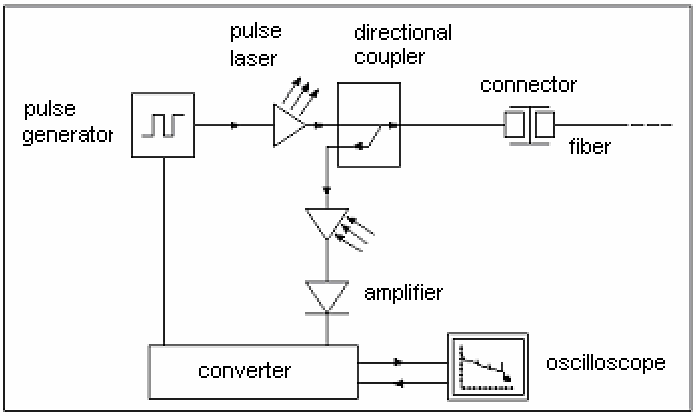


Fig. 3. Diagram of reflectometer.

Laser diode is moduled by the impulse generator. Optical impulses are injected to the fibre by the direction clutch. Amplitude of these impulses is chosen so that they don't override one another and ranges from 1 kHz to 20kHz, depending on the length of the tested fibre.

As a result of the Rayleigh's distribution or Fresnel's reflection, part of the energy returns to the reflectometer. Returned signal is directed through the clutch to fotodetector, where it is changed into electric signal, strengthened and recalculated into its digital version. Knowing the speed with which the optical pulse propagates in fiber, we can replace the analysis of the rear distribution as a function of time with an analysis as a function of length from the place of the measurement. Finally on the display of the reflectometer we derive: on axis x – length, on axis y – power of the returned signal (in decibells).

3. Reflectometer measurement

Below in Fig. 4. an example of the track section of fiber, which has been tested by reflectometer.

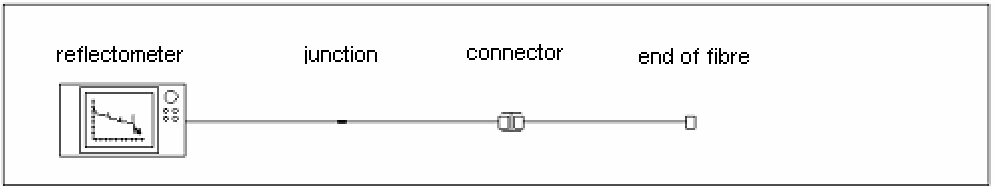


Fig. 4. Examined length of light tube track.

Test results are shown in reflectograph Fig. 5., it is a graphic representation of power as a function of distance measurement track.

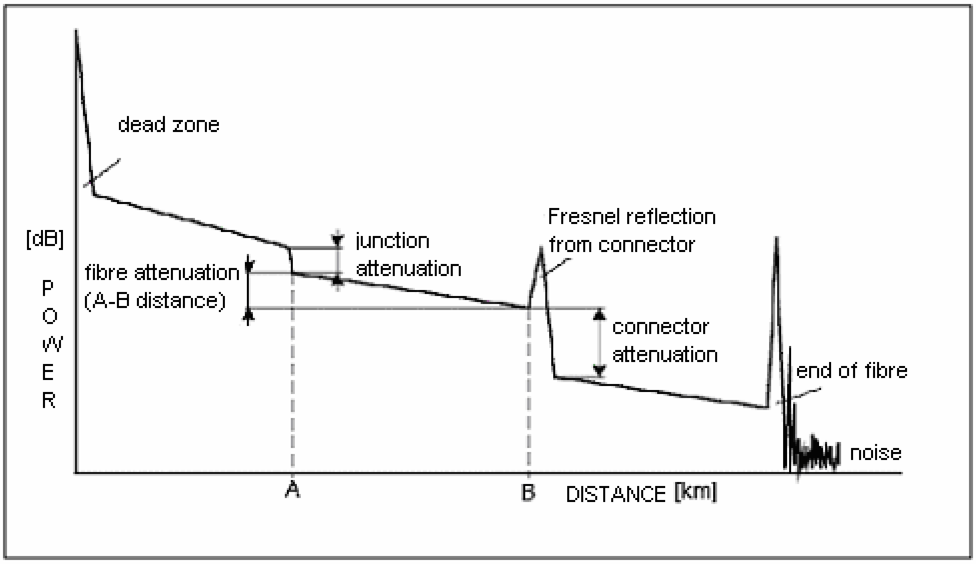


Fig. 5. Reflectogram of the examined length.

In places of non-uniformity of the light tube track an optical reflection of measured impulse occurs. These places are mechanical links of fibres, end of fibre, cracks.

a). Basic parameters of reflectograph

The correct interpretation of reflectograph allows you to specify the correct results of the reflectrometric measurement, below presents the basic parameters. Dead zone – level of reflected power is often several decibels higher than the level of rear distribution power. Reflected impulse may cause re-steering of the reflectometer receiver, which in turn causes incorrect optical-electrical recalculation. This length of the light tube is placed in such called Dead Zone on the reflectogram.

Dynamics - reflectometer is a distance between the level of rear distributed signal at its start and the level of noise in the fibre. The dynamics of the measurement increases with the increase of the measured impulse lengths and averaging time – the longer measurement the more accurate.

Reflectancy - occurs when there is a high step change of the parameters of light bends, eg. At the border glass-air. On the reflectogram we can see then sharp short impulse, which is a result of the Fresnel’s reflection.

This phenomenon occurs:

- at the end of the light tube, so ending or break of the cable,
- on mechanical links, so re-linkable,
- on incorrectly welded links.

In interpreting the results, we have to account measurment errors such as the phenomenon of ghost (multiple reflected signal) and negative attenuation, which could be wrongly interpreted as a reinforcement signal. This happens when the fiber which is located after the junction has bigger coefficient of reverse dispersion (for example, because of larger diameter) - this causes a biggest attenuation of that segment.

4. Conclusion

Undoubted merit of the reflectometer (OTDR) is the ability of examination of the transport track from one end of the light tube, which is great logistic facilitation to examine long length tracks. Accuracy of fault location depending on the quality of the instrument is up to 10-15cm, which allows to precisely plan the place of regeneration of the light tube. An important factor while working with reflectometer is adequate choice of parameters that will help in accurate interpretation of measurement results.

References

[1] LAMPERSKI, J. *Basis of light tube Telecommunication*, Lecture, University of Technology, Poznan, Poland 2003

[2] RATUSZEK,M. *Thermal links of one-mode telecommunication light tubes*, WUTP, Bydgoszcz, Poland 2008