COMMUNICATIONS

COMMUNICATIONS

Dear Readers,

It is a great honor for me to write an introduction to this issue of "Communications - Scientific Letters of the University of Zilina". This is a special edition compiled on the occasion of the 25th anniversary of the Faculty of Management Science and Informatics. Are 25 years of existence of the faculty many or few? I'll try to express the ratio.

Let me start by comparing the age of humanity to the age of our planet. If we compressed the age of the Earth, which is about 4.5 billion years, into one day, the first man on the Earth would appear one second before midnight.

Similarly, if we compare the age of the faculty and the length of the existence of higher education on the Earth, the launch of which can be dated somewhere around the year 800, when Charlemagne established the requirement for clergy to be educated, to learn and to teach, and compressed this time to one day, the age of the faculty could be 30 minutes.

When we realize what humans managed to accomplish in this 1 second of their existence, notional 30 minute existence of the faculty is the sufficient reason for a brief evaluation.

In these imaginary 30 minutes the faculty has produced nearly 3,400 engineers and more than 150 PhD graduates.

Of course, History of informatics and control at the University of Zilina is much older and dates back to 1972, when the Department of Technical Cybernetics was established.

Information and communication technology in the field of science and research is booming, although it is still a young science. This edition of the journal brings the latest knowledge in the field of informatics, information and communication systems, management, optimization, simulation, etc.

Informatics has been successful also because it helps to solve problems in other areas of research. Thanks to applied informatics, many difficult questions in economics, mechanical engineering, civil engineering, physics, etc. have been solved. Many questions are still open and waiting for our solutions mainly in an interdisciplinary approach and joint research projects.

This edition is an example of the connection of informatics and mechanical engineering. In the journal you will find articles from the field of machines and equipment, material research, production technologies, etc.

This edition of the journal will bring interesting information to all readers and will open new horizons in science and research.

Emil Krsak

Jakub Hrabovsky - Pavel Segec - Peter Paluch - Marek Moravcik - Jozef Papan *

USABILITY OF THE SIP PROTOCOL WITHIN SMART HOME SOLUTIONS

A smart home is being an important part of research due to its impact on the life of many people, who decide to try this partially known concept of building and controlling their houses. In this paper we present a little different approach to the implementation of a smart home structure and its methods in real environment. As a part of the work we specify main requirements on the architectural design, which should be considered in final solution. Our research is focused primarily on the use of open standards and tools. This approach became popular because it provides an alternative view in comparison to commercial ones, where mostly proprietary and publicly unknown protocols are used. According to this, three theoretical architectures are proposed and presented with their pros and cons. In the paper we concentrate mainly on the usage of SIP protocol and issues related to the communication between a remote user and a smart home. We are proposing two logical communication models. Each model describes the type of used SIP messages and their message flows. In addition, we consider financial aspects of simple and illustrative smart home solution, which as well pose issues to end users. Our solution offers implementation with minimal operational costs and high facility control of the smart home. The paper also includes a survey of potentially usable protocols for smart homes and the list of software (SW) and hardware (HW) components, which have been used to build a testing environment.

Keywords: *Home automation, smart home, SIP, session initiation protocol, central management.*

1. Introduction

The development of new technologies influences almost all fields of human life including home environment. By using new inventions and communication standards, we are able to automate operation of whole buildings, control their parts and achieve their optimal usage. The optimization causes a huge reduction of operational costs and time needed for flawless run of automated buildings.

An example of intelligent building is the smart home, where all core functions are controlled by special entity that is manageable by end users. Common functions are heating, lighting control, security and functions related to danger detection. The term “smart home” comes from the idea of automating actions of common house devices and interconnecting them to be capable of co-operation without user interaction.

The smart home has to be seen as a functional cluster of subsystems, which correspond with mentioned functions. Each subsystem is required to make its specified functionality and co-operate with other subsystems. They all together make one complex system.

In the heart of the system is the main controller, known as the central control unit. End users, using for example their mobile devices, are able to manage controller activities. Users send commands to the controller and as an answer they receive requested information about actual situation in their home. With this approach they are able to remotely access their houses independent of their actual place and time of day. The controller is also responsible for the management and interaction of each subsystem inside the smart home. This central control unit provides unified control of all electronic devices, which are members of the smart home. To ensure this, usually a special purpose software program is running inside the central control unit. It receives data from all connected devices, processes them and performs appropriate actions depending on actual situation.

There are other devices beside controller, which play a role within an architecture. They build a group where they are all interconnected via a special network. This way they are able to exchange information about different parts of home environment and so create its overall view available for end users. This service together with the remote access highly increases comfort of end users.

* **Jakub Hrabovsky, Pavel Segec, Peter Paluch, Marek Moravcik, Jozef Papan**
Faculty of Management Science and Informatics, University of Zilina, Slovakia
E-mail: jakub.hrabovsky@kis.fri.uniza.sk

In our proposal we are relying only on open communication standards, mainly the Session Initiation Protocol (SIP) [1]. This approach is different in comparison to methods used in commercial sphere, where proprietary and publicly unknown protocols are mostly used. Our approach brings another usage proposal of known protocols, SIP in this case. The choice of common and well-defined protocols brings us considerable advantages. The primary advantage allows us to design clear and relatively simple architectural model. As the second advantage, it provides us with good accessibility of all elements needed for its realization, such as HW and SW entities.

When we created an architectural model, we were considering the system autonomy, comfort of end users, simplicity and minimal costs as the main model parameters. In the paper we propose three slightly different design architectures and a solution based on one of them to achieve considered properties.

The rest of the paper presents related works, analysis of components for building a smart home, as well as proposed architectures and communication models. Brief overview of reference research papers is introduced in section 2. Section 3 gives the description and tasks of each logical entity included in the smart home. The section also presents four potential communication protocol candidates, which can be used in environment of the smart home. Section 4 describes our three architecture proposals, which utilize the SIP protocol. The section also includes description of two logical communication models, used messages and theirs flows. Section 5 presents a particular proof concept implementation with a list and short description of used hardware devices and software tools. Results and experiences coming from the realization and subsequent conclusion with proposals for the future work are presented in section 6.

2. Related work

There are several researches done in the field of the smart home development and the usage of common open protocols [2] presents the use of SIP and the propagation of extended location information gathered from various sensors (wired and wireless) via presence service as a context for dynamic behavior of home devices. The paper emphasizes important ability of end users to have full control over the home and its customization in the form of various profiles. A uniform model of heterogeneous entities of the smart home, which are using a common SIP interface with extensions as an abstraction layer, is described in [3]. The author specifies three basic communication forms (commands, events and sessions) with their particular realization via SIP interface. At the same time, author provides an abstract interface between a high-level programming (applications) and a low-level communication (SIP) defined in an architecture description language (DiaSpec). DiaSpec hides details of used technology and simplifies creation of new applications. To take advantage of open protocols like SIP, some approaches are using SIP infrastructures and applications already in existence to implement a smart home model. Such approach is presented in [4], where existing IP Multimedia Subsystem (IMS) provides SIP services (registration, session management, instant messaging and presence [5]) for an instance of the smart home [6] and [7]. Such approach allows a researcher to focus more on the design of a signaling gateway between SIP and hardware communication protocols like KNX. In [8] is provided an example of signaling gateway implementation deployed as a SIP adapter. As the gateway platform an ATmega board with an Ethernet shield is used. The platform is programmed in C that uses functions of the oSIP library. Another example of such gateway and its implementation to Raspberry Pi is described in [9]. Authors propose the usage of summarized SIP Uniform Resource Identifier (URI) of the smart home in comparison with the usage of several separated SIP URIs, each for individual home device. The main focus of the paper is placed on the security and the access control for home private data, where the concept of security layers is proposed. In our approach, we follow results and principles identified in mentioned papers, mainly the last one. We provide our own solution driven by the analysis described in next sections.

3. Structural view of the smart home

The smart home consists of elements that perform various functions - central control unit, sensors, control elements and remote mobile clients. We also consider a communication bus as a separate part of home structure. Here we provide a brief description of each entity and its role. It serves as the reference of architectural design.

A. Central Control Unit

Central control unit (CCU) is the main part of smart home. It controls other subsystem parts and also serves as a gateway between a remote client and an internal system. It achieves this by receiving and collecting data from home sensors, processing them to gain some valuable information and depending on the results by executing appropriate actions. Actions can be classified into two groups. Classification of an action depends on its initiator, which provides stimulus. If the initiator is the controller itself, we talk about scheduled actions. Otherwise (initiator is another device, i.e. one of sensors or remote users) the action is identified as executed. Two standard actions are sending information about a specific subsystem to the remote client and the modification of parameters of specific subsystem. In this paper the term home gateway (HG) is used to identify CCU.

B. Sensors and control elements

Sensors are sensible devices, which monitor a specific subsystem of the home environment. They extract appropriate data from the environment and send them to HG. Control elements, known as actors, are responsible for making desired changes in the behavior of home appliances. They are based on commands received directly from HG. Sensors and actors create the lowest layer of the system.

C. Remote mobile client

As a remote client is usually used a mobile device (Smartphone or tablet) with a software client. User uses the client to send requests (commands) to HG and eventually receive responses (depend on the type of request). The client serves as a user interface for remote access to HG. Software client has to be appropriate protocol application. This depends on a protocol, used for a message exchange between the client and HG. More information about our application is provided in section 5.

D. Communication bus

The communication bus is very important because it connects all subsystems and their elements together into one system. The format of communication via the bus is defined by used communication protocols. In our solution, we need to handle two independent communications applied at two separate places of the system because of their distinct requirements on network parameters. Main parameters that have to primarily be taken into account when choosing a right protocol are speed, distance, security and power-consumption.

1) The communication between HG, sensors and control elements

The communication relates to the lowest layer. Communication protocol focuses on the hardware connection with simple data exchange, where data format has the form of signals with two possible levels (on/off). Protocol is responsible for the interconnection of intern environment of the smart home – sensors and control elements. Therefore, it puts emphasis on power-consumption aspects. For completeness, examples of open protocols suitable for this type of communication are X10, ZigBee Home Automation, Z-Wave, Insteon, Bluetooth Low Energy (BLE), Wi-Fi and Near Field Communication (NFC). Wired or wireless technology is used as a transfer medium and depends on the specific protocol and its support.

2) The communication between HG and a remote client

In comparison to the previous type of communication, this type relates mostly to the application layer. It focuses on an application data exchange where data format has the form of messages. Protocol is responsible for the communication between HG and a remote client. Therefore, this communication puts emphasis on speed, reliability and security.

Our solution considers the usage of SIP right for this type of communication.

E. Communication between HG and clients

To support the communication, we think about three different types of abstract messages, SET, GET and NOTIFY that relate to distinct operations, which are performed by HG. The support of messages should be considered as the requirement on selecting correct protocol. The message exchange proceeds through interface of specific protocol in the process of implementation.

The GET message is used to send requests from the side of a remote client. HG reacts by sending requested data back in the next GET message. This way the client can get information about actual state of the home. To be able to send command and set a state, remote client uses the SET message. After receiving it, HG reacts by executing desired action. The initiator of communication in both previous cases is the remote client. To be able to initiate communication from the side of HG, the NOTIFY message is needed. Using this message HG informs remote clients about events, which just happened, e.g. detection of fire or breach of security. The communication protocol must allow simple and efficient implementation of all of these abstract messages and ensure their correct exchange. There are some other requirements. For example, video cameras are often used parts of the smart home. Therefore, there is also a requirement to support real-time transmission of video data from HG to the remote client. Other requirements come out from the security and innovations. There must be ensured that only authorized clients can communicate with HG and control its action. Similarly, we should be able to add new functionality without bigger problems, therefore the solution has to be technologically independent.

According to these requirements together with the openness, there are four communication protocol candidates – HyperText Transfer Protocol (HTTP) [10], eXtensible Messaging and Presence Protocol (XMPP) [11], MQ Telemetry Transport (MQTT) [12] and SIP.

The HTTP protocol is used for transfer of web content and it natively supports the GET method. It allows the implementation of GET and SET abstract messages. However, the protocol does not provide native and straightforward support of the NOTIFY message. One of possible ways how to implement this feature lies in uninterrupted sending of requests by client to HG. Once an event happens, HG will inform a client using following GET message. HTTP enables us to transfer video via streaming servers and Web Real-Time Communication (WebRTC) technology.

XMPP is a protocol based on XML objects. XMPP supports the implementation of all three abstract messages using appropriate sections of XML message – presence, message and iq (info-query). XMPP supports video transfer with the use of different extensions, as for example Jingle.

MQTT is an open communication protocol based on the client-server architecture. Since 2014 MQTT is one of OASIS standards. MQTT allows simple message transfer between various devices. Its implementation is very suitable for devices, which have limited performance, memory and power. These properties are typical for devices used within Internet of Things (IoT) or machine-to-machine (M2M) devices, for example those as various sensors, smart appliances and other embedded systems located in the home. Thanks to used communication model a MQTT server, which is called broker, is able to handle a great number of concurrently connected clients. The MQTT communication is based on the exchange of publish/subscribe messages and it uses a data-centric approach. A MQTT client does not send data directly to another client, but its data is forwarded only to the corresponding MQTT server. The server is afterwards responsible for "sharing" data to subsequent subscribed clients. The MQTT protocol supports enhanced security features like authentication, authorization and encryption through the implementation of the Transport Layer Security (TLS) protocol. As an advanced additional feature, the support of quality of service (QoS) can be included. MQTT usually runs on top of the TCP/IP architecture. However, taking into account actual progress in the field of IoT, there was created a new extension called MQTT for Sensor Networks (MQTT-SN) [13], which is designated for connectionless networks, as for example wireless sensor networks. MQTT-SN can be used on top of any kind of network, which provides lossless bidirectional connection with preservation of message order. Thanks to the network independence various transport protocols, like Zigbee known in the area of embedded devices, are available for MQTT-SN. Main goal and advantage of MQTT-SN is that it allows the deployment of MQTT in networks with low bandwidth connections and high link failures - wireless sensor networks. MQTT in its nature is only a simplified version of other here mentioned protocols, such as SIP and XMPP, however MQTT mainly concentrates on one specific function, the transfer of simple messages by minimizing their size and simplifying their transfer and processing needs. MQTT is appropriate and nowadays often used protocol alternative for a sensor data distribution within smart homes. On the other hand, its simplicity doesn't offer any method to meet some requirements mentioned above, such as video data transmission. To allow such services the collaboration with another protocol is required.

SIP has been proposed for a session management. Its role lies in creating, controlling and removing sessions between two and more end nodes. It describes each session with distinct parameters exchanged in specific messages. SIP is independent of used transport layer, provides unique abstract address for each device in the form of SIP URI and gives us means for simple implementation of all previously discussed abstract messages. SIP distributes individual functions among different logical entities as clients, registrar and proxy servers. The list of functions and their responsible entities includes registration, authentication and authorization of remote clients on the side of HG (SIP proxy), basic security, message exchange on the side of remote client (SIP Agent) and an interface between SIP and intern control plane of the smart home (SIP Gateway). Section **5** will provide a detailed description of proposed communication models, which incorporates SIP entities and their co-operation. An example of relation between SIP entities is shown in Fig. 1.

To ensure a secure connection, SIP supports several mechanisms. The simplest one uses a basic authentication based on a login name and a password. Others include the support of secure technologies like TLS, Message Digest 5 (MD5) and Internet Protocol Security (IPSec). All of these may be added to the communication model as means of enhanced security.

Considering all advantages of SIP over other three protocol candidates and our overall experience and knowledge gained by working with it, we decided to use the SIP protocol.

4. Architectural design

SIP allows several ways of implementation of communication between the remote client and HG. We propose three slightly different communication architectures. They describe relations between individual smart home components. They differ from each other in the placement of SIP server and in additional use of Virtual Private Network (VPN) entities.

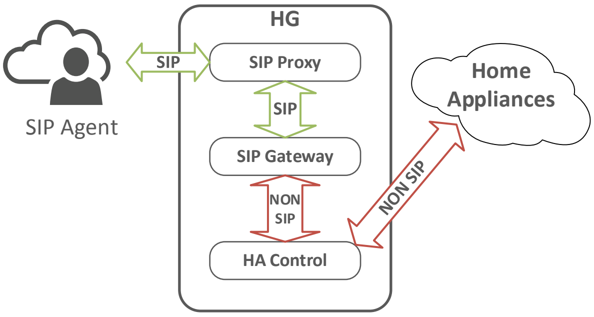


Fig. 1 Relations between SIP entities

F. SIP-only public server-based architecture

This architecture relies on the use of publicly placed SIP server with a public IP address assigned. Therefore, the server is globally accessible. The HG entity is a part of a private network located behind a NAT device inside the smart home. Both, the remote client (SIP Agent) and HG have to register to the public SIP server capabilities. After successful registration, communication between the client and HG goes via SIP server.

The advantage of this architecture is public SIP server accessibility where one individual SIP server may be used to control the communication of a group of smart homes. On the other hand, the charge for renting a public SIP server by some hosting company causes that the overall costs increase and requirement on the low price isn't more met. Other issue is the reliability of public SIP server, which represents a Single Point of Failure (SPOF) [14]. The risk of SPOF is usually carried out by implementing a cluster of public SIP servers where each is a backup for the others. This, of course, rises costs again. Figure 2 gives a better understanding of relations between given entities.

Considering advantages and disadvantages of this architecture, it could be effectively used in the case of multi-domain server deployment model (with more customers). However, for our view of the smart home it isn't suitable.

F. SIP-only private server-based architecture

As shown in the brief descriptive picture (Fig. 3), this architecture uses a private SIP server. The SIP server is implemented together with HG and both are placed in a private network of the smart home (i.e. behind customer's internet router with NAT). It means that each smart home has its own SIP server with assigned private IP address and with complicated remote reachability. The main issue lies in the translation of private address to the public one because the household gets from an ISP usually only one dynamically allocated public IP address. As a solution, which allows the access from a remote client to the SIP server placed in a private network, we can use Dynamic Domain Name Server (DDNS) together with Port-forwarding. DDNS informs the remote client about actual public IP of the home gateway. Port forwarding ensures that home gateway forwards each incoming SIP message (with destination port equal 5060) to the private address of SIP server.

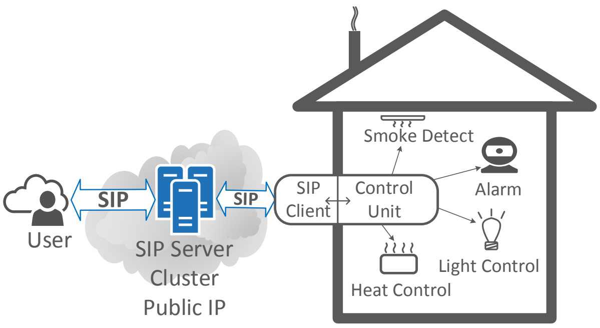


Fig. 2 SIP-only public server-based architecture

The existence of NAT presents especially for SIP a serious issue. The SIP protocol in its messages uses parameters of lower network layers like IP addresses and transport ports. There exist ways that address NAT in SIP like STUN and TURN. However, the overall functional implementation of aforementioned methods for NAT traversal is not straightforward.

H. SIP-VPN private server-based architecture

Looking at disadvantages of previous architectures, we can make an enhancement by adding additional features. For the interconnection of a client and HG a separate virtual private network (VPN) connection is created. To make it possible, VPN server runs beside SIP server. Both are placed in a private network. On the other side of connection, a VPN client is used as a part of remote client. Using VPN, the issue related to NAT is cleared, because both communication nodes (client and HG) are in the same network. In addition, VPN improves the security by using encrypted connection. The problem of learning actual public IP address of home gateway is also real in this architecture. To overcome the issue, DDNS and Port-forwarding techniques can be used again. Connections between individual parts of the architecture are in Fig. 4.

SPOF issue mentioned in the first proposed architecture occurs also in remaining two. In the case of second and third design, we have to back up HG by using backup power supply and a duplicated hardware device of implemented HG. Loss of internet connection opens a security problem, because without a connection the smart home cannot inform remote client about occurred events. Therefore, it is important to implement other ways of informing remote clients, which run simultaneously. For example, SMS message can be sent to the client's phone to inform or alert him.

According to requirements on the smart home, advantages and disadvantages of individual architectures, especially difficulty of implementation and security issues, the third one seems to be the right choice. Therefore, our implementation presented in the next section comes out from the SIP-VPN private server-based architecture.

I. Models of SIP message flows

As was already mentioned in section 3, there are several ways or models how to incorporate and use SIP to satisfy our needs. We again propose two communication models, which describe types and flows of individual SIP messages. Video streams aren't considered in following models because Real-time Transfer Protocol (RTP) is responsible for their transfer immediately after successful session establishment.

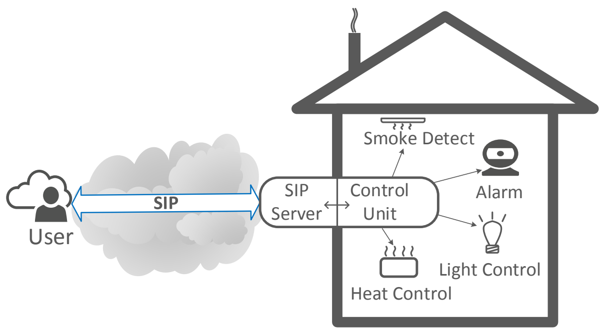


Fig. 3 SIP-only private server-based architecture

Communication models refer to various logical entities where each plays a particular role. SIP proxy manages SIP communication and includes client registration. SIP GW serves as an interface between a SIP proxy and a Home automation (HA) controller, which communicates directly with sensors. SIP GW consists of SIP server and intern SIP user agent (SIP

UA). The second SIP UA is used at remote client. From the SIP communication perspective, both act as end clients that communicate on a peer-to-peer basis. For clarity, HA controller is also put into visual figures of models and is responsible for management of sensors and control elements. Usually several logical entities are realized in one physical HW device. In our models SIP proxy, SIP GW and HA run as processes on one physical device.

Both models assume that both clients (SIP UA on remote client and SIP UA on SIP GW) are already registered at SIP proxy. Logical entities and their relations for each of following models are in Figs. 5 and 6.

1) Message-only model

The first model uses only one type of SIP request to implement all three abstract messages described in section 3 (GET, SET and NOTIFY). The SIP MESSAGE request is originally a part of Instant Messaging service (IM), which provides primitive chat functionality.

When a client wants to check actual value or a state of specific subsystem (GET operation), it sends request in the form of SIP MESSAGE (see Fig. 5). SIP MESSAGE includes requested parameters and instructions. The request is received at SIP GW.

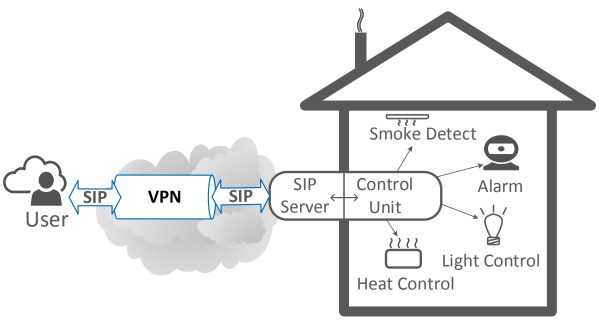


Fig. 4 SIP-VPN private server-based architecture

SIP GW extracts instructions from the message and sends them on the input of HA. The HA entity processes incoming instructions, reads data from desired sensor and as a response informs back SIP GW. In our solution we are using for this exchange means of inter-process communication. As a next step, SIP GW uses internal SIP UA to generate a SIP MESSAGE with desired value and sends it to the remote SIP UA.

In the case of SET abstract message almost the same flow is used. However, at HA there is distinct processing of the request where HA takes desired action (apply setting) instead of just reading values from sensors.

The issue emerges in the case of NOTIFY abstract message where we need to initiate SIP communication from HG. In this model we propose a simple solution based on regular sending of requests from remote client about actual state of monitored event.

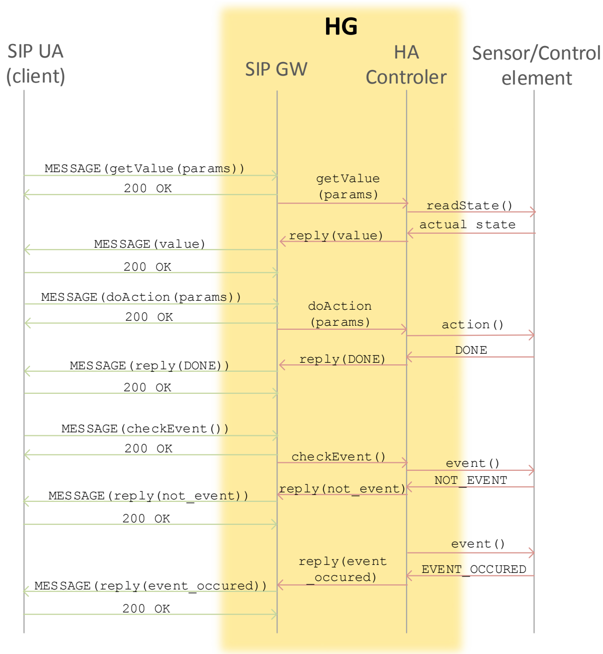


Fig. 5 Message-only model

The main advantage of this model lies in its simple implementation. On the other hand, proposed solution of NOTIFY message implementation is inefficient. Situation gets worse when more remote clients check occurrence of some event simultaneously. In that case, HG must handle each of them separately and that causes significant increase of HG load.

2) Message-presence model

The second model adds presence service with corresponding PUBLISH and NOTIFY SIP message types. This separates the implementation of NOTIFY and GET abstract messages. The SET message is implemented in the same way as it was in the previous model.

Used presence service requires deployment of additional SIP entities - watcher, presentity and presence server. Watcher is a subscriber that applies for the presence information about a specific subsystem or an event. Presentity is a monitored entity specified by its state. When the state is changed, information is sent to the presence server in the PUBLISH message. Presence server (PS) is responsible for keeping a list of subscribers. It maintains actual states of registered presentities. In the case of status changes, PS sends information about this change to all subscribers in the list. NOTIFY message serves this purpose. The body of both message types has unique format called the Presence Information Data Format (PIDF) and it has the form of XML document. The SIP presence extension is used for the implementation of NOTIFY and GET abstract messages.

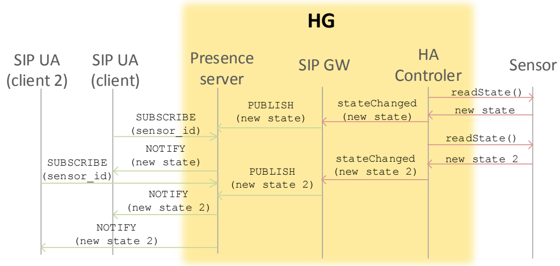


Fig. 6 Message-presence model

Monitoring of sensors is done by HA. Each change of a state is automatically reported to SIP GW (Fig. 6). SIP GW then creates a PUBLISH message, which includes this new state and sends it to the presence server. PS updates corresponding records. Subsequently PS sends the NOTIFY message to all registered subscribers. Therefore, the remote client has to subscribe to particular subsystem in order to get desired information about it.

This behavior of PS helps to solve the issue related to simultaneous checking of an event occurrence by more remote clients. Thanks to the deployment of SIP presence service, the whole responsibility is shifted from HA to the presence server. This solution is more efficient compared to the first solution and is considered as more preferable. Disadvantage of this model is its demanding implementation, which requires additional configuration of presence server and the support of presence service at both SIP UAs.

According to proposed models, their requirements and properties, we had to decide in our solution for the first model. The main reason comes from the lower difficulty of its implementation and the availability of suitable software components in comparison to the second model.

5. Our implementation

Summarizing our choices, as the proof concept we implemented the solution based on SIP-VPN private server-based architecture and message-only communication model. Description of the solution's components is provided in this chapter. The solution meets requirements specified in the introduction of the paper. We split the description into two parts - hardware and software. Each part contains specification of used tools and devices that correspond to individual smart home entities.

A. Hardware components

The list of HW components includes HG, sensors and control elements. HG is the main physical entity of the smart home, as mentioned earlier. HG is a HW device where all logical entities (client and servers described in section 4) will run as software processes. Therefore, HG must have sufficient performance to be able to perform all requested processes at the same time. In addition to that, HG must support sufficient amount of physical connectors and their types. Through them, we interconnect other parts like sensors and control elements. To be as close as possible to a real environment, we expect and support the use of following connectors. Ethernet port (RJ-45) for the internet connection, a set of General Purpose Input/Output (GPIO) pins used to manage sensors and control elements, 1-Wire bus interface for thermal sensors, an analog output for smoke sensor and a Universal Serial Bus (USB) port for video camera. One of key requirements requested on the HG is its power-consumption, which has to be low. Considering all requirements, a single board computer (SBC) was chosen as a hardware platform for the solution implementation. There are many SBCs on the market today. We compared three different models – Raspberry Pi 2 [15], Cubieboard 2 [16] and BeagleBone Black [17]. Based on the results of comparison shown in Table 1 we picked BeagleBone Black. In comparison with other two options, it meets all requirements and brings us additional advantages in the form of operating system stability and simple handling of peripherals.

For demonstration purposes as sensors, we implemented one sensor type for each of subsystems. As a thermal sensor, we used DS18B20 by Dallas Semiconductor. Sensor supports 1-Wire bus and assigned unique address. Thanks to the shared access of 1-Wire bus we are able to attach more sensors without a need for additional pins or connectors. DSN-FIR800 PIR sensor was chosen as an example of a motion sensor. As a smoke detector, we chose the MQ-4 sensor. It allows us to detect a potential fire. In addition, the smoke sensor allows to detect occurrence of a gas in the environment. For a water leakage detection, we used simple water sensor placed on the floor.

As control elements a set of relays controlled by GPIO pins is applied. Individual home appliances can be controlled via relay by changing a value of specific GPIO pin. We also demonstrate an electrical control of temperature. This is performed by electro thermic radiator head which is connected to a relay.

B. Software components

Logical entities of the smart home, including SIP and communication entities, are implemented as software components running in HG and in a remote client (smartphone). We expect entities: SIP server, SIP client, VPN server and a VPN client. As HG's operating system, we decided for the GNU/Linux. In the case of remote client Android OS is used. As a SIP server running on HG we selected the Kamailio SIP server. Kamailio provides all required functions including IM and Presence and its demands on the system resources are low. As a VPN server, which runs on HG, we used the OpenVPN server. OpenVPN server is free and stable IPSec solution. On a remote client we are running CSipSimple (a SIP client) and OpenVPN (VPN client) applications. Relatively complicated issue is an effective generating of SIP messages from HG. We solved this by using separate SIP clients running on HG. This option brings minimal additional delay and processor load without the need for the development of own clients. We picked the Linphone application as one of freely accessible and usually deployed SIP clients. Linphone almost supports all required functions including command line interface (CLI) management. CLI management allows simple control and is necessary for our solution. All used software tools are open-source as specified in requirements.

Comparison of three distinct SBCs Table 1

	BeagleBone Black	Cubieboard 2	Raspberry Pi 2
Processor	ARM-Cortex A8	ARM-Cortex A7	ARM-Cortex A7
Core	Single-core	Dual-core	Quad-core
CPU Frequency	1 GHz	1 GHz	900 MHz
Operational memory	512MB DDR3	1GB DDR3	1GB LPDDR2
Ethernet	✓	✓	✓
USB	1x	2x	4x
GPIO	65x	86x	40x
1-Wire	✓	✓	✓
Storage	2GB +SD card	4GB + SD card	SD card
Power Consumption	3W	4W	3W

Implementation of the HA controller requires special functionalities. Therefore, we designed and built our own software program written in C programming language. Based on ad-hoc solution, our controller is able to communicate and co-operate with SIP server as is required in the proposed **architecture**.

XI. HA controller design

Structure of the control program consists of one master and more slaves. Master is the main program responsible for initial start of slaves and their central management. Each slave is a separate unit implemented as a thread that provides particular functionality. We chose POSIX threads, which are available directly from Pthreads library. Threads require fewer resources than processes and offer simpler way of managing and sharing data between individual slaves. Communication from SIP server to HA is realized through simple file access, where after receiving specific SIP message SIP server extracts its body and stores it as text in the file. Next, these data are read by master of HA. For opposite direction HA uses a separate SIP client placed in HG to create and send a SIP message with desired information. Initial requests for creation and transmission of messages are in competence of individual slave programs. The flow of messages and mutual relations between parts described above is drawn in Fig. 7.

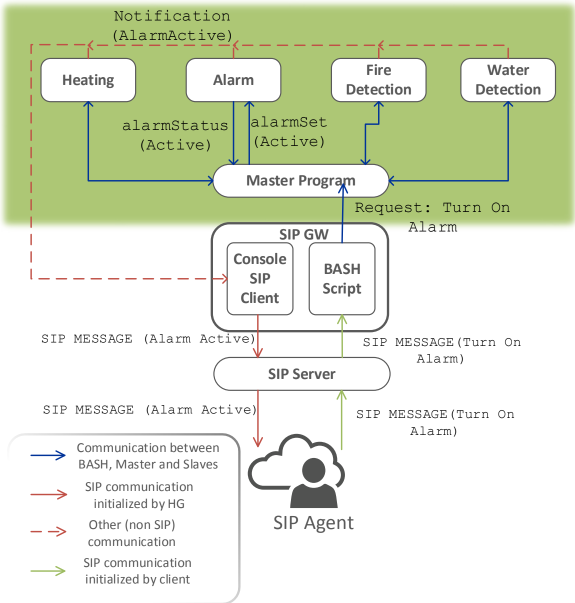


Fig. 7 Communication model of HA controller parts

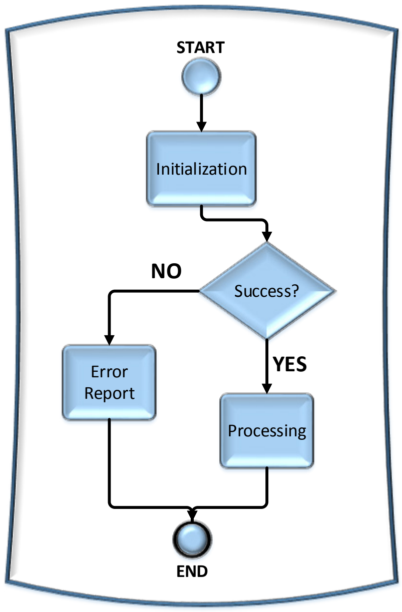


Fig. 8 Flow chart of common slave

Thanks to insertion of additional processing logic block into SIP server, every time it receives a SIP message destined for HA, it is able to identify action inside the message and store it in the text file. All slaves (heating, security, water and fire detection controllers) have very similar structure as shown in Fig. 8. They differ only in the way of processing particular function based on checking a state and reacting to its potential change. Slave periodically accesses data gathered by individual sensor to check its actual state. The activity has the form of access to the file. This abstraction is provided directly by operating system and its filesystem placed inside memory of our board PC. Subsequent reaction consists of creating appropriate message with measured value and sending it via SIP client interface to the desired destination. That is local SIP server or remote SIP agent.

The real power of this design comes from the distribution of functions into a set of simple programs running parallel in separate threads. It allows different subsystems to be controlled simultaneously and react to special events immediately. Although, our presented solution isn't optimal and requires additional improvements, it is good enough to show, that usage of the SIP protocol is so broad it can be used also in this manner.

6. Conclusion and future work

The paper describes the role of open protocols in the area of smart homes and is particularly dedicated to SIP. We addressed emerging issues related to this protocol. We analyzed recent papers focused on the problematic. Considering their results and principles, the paper introduces our basic smart home structure proposal with respect to present state of knowledge and research. The paper also provides an overview of solution's key components. We proposed three architectural models. Each of them uses slightly different approach to the realization of communication between the smart home and remote clients. Design, implementation and deployment of our solution were presented with emphasis on the realization of individual software and hardware components. We deployed a particular solution, which satisfies specified requirements, such as open-source approach, system autonomy, simplicity and minimal costs. Our testing solution costs about 120 €and provides functional system for basic management of the simple smart home.

There are some improvements of our solution, which should be realized as future works. For example, the realization and deployment of the second model mentioned in the paper. The intern SIP client, which is a part of SIP GW in actual solution, could be replaced with the use of functions directly from available SIP library inside of the HA control program.

References

[1] ROSENBERG, J. et al.: SIP: *Session Initiation Protocol*, IETF, RFC 3261, June 2002.

[2] SCHULZRINNE, H., WU, X., SIDIROGLOU, S., BERGER, S.: Ubiquitous Computing in Home Networks, *IEEE Communications Magazine*, pp.128-135, Nov. 2003.

[3] BERTRAN, B., CONSEL, C., KADIONIK, P., LAMER, B.: *A SIP-based Home Automation Platform: an Experimental Study*, Proc. of 13th Intern. Conference on Intelligence in Next Generation Networks, IEEE, Oct. 2009.

[4] ACKER, R., BRANDT, S., BUCHMANN, N., FUGMANN, T., MASSOTH, M.: *Ubiquitous Home Control based on SIP and Presence Service*, iiWAS '10 Proc. of the 12th Intern. Conference on Information Integration and Web-based Applications & Services, ACM, Oct. 2010.

[5] ROSENBERG, J.: *A Presence Event Package for the Session Initiation Protocol (SIP)*, IETF, RFC 3856, August 2004.

[6] KOVACIKOVA, T., SEGEC, P., BRUNCKO, M.: Standardization Paths for NGN IMS-based Architecture, *Communications - Scientific Letters of the University of Zilina*, vol. 10, No. 4, pp. 15-18, ISSN 1335-4205, 2008.

[7] SEGEC, P., KOVACIKOVA, T.: Implementation of IMS Testbeds using Open Source Platforms, *Communications - Scientific Letters of the University of Zilina*, vol. 14, No. 2, pp. 55-62, ISSN 1335-4205, 2012.

[8] DZINDZALIETA, R.: *SIP Protocol as a Communication Bus to Control Embedded Devices*, Local Proceedings and Materials of Doctoral Consortium of the Tenth International Baltic Conference on Databases and Information Systems, vol. 924, pp. 229-234, Vilnius, July 2012.

[9] GEBHARDT, J., MASSOTH, M., WEBER, S., WIENS, T.: *Ubiquitous Smart Home Control on a Raspberry Pi Embedded System*, The eighth Intern. Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, pp. 172-177, 2014.

[10] FIELDING, RESCHKE, J.: *Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing*, IETF, RFC 7230, June 2014.

[11] SAINT-ANDRE, P.: *Extensible Messaging and Presence Protocol (XMPP): Core*, IETF, RFC 6120, March 2011.

[12] OASIS Standards: MQTT Version 3.1.1, 2014.

[13] STANDFORD-CLARK, A., TRUONG, H. L.: MQTT For Sensor Networks (MQTT-SN), Protocol Specification, Version 1.2, 2013.

[14] REZAC, F., et al.: Bruteforce Attacks Blocking Solution on Embedded SIP Communication Server, *Communications - Scientific Letters of the University of Zilina*, vol. 15, No. 2A, pp. 180-184, ISSN 1335-4205, 2013.

[15] UPTON, E.: https://www.raspberrypi.org, 2015.

[16] CUBIE, T.: http://cubieboard.org, 2013.

[17] BeagleBoard Foundation, http://beagleboard.org/black, 2013.

Ivan Cimrak - Katarina Bachrata - Hynek Bachraty - Iveta Jancigova - Renata Tothova - Martin Busik
Martin Slavik - Markus Gusenbauer *

OBJECT-IN-FLUID FRAMEWORK IN MODELING OF BLOOD FLOW IN MICROFLUIDIC CHANNELS

We present a fully three-dimensional computational model of red blood cells and their flow in a fluid. This model includes all components necessary to capture important physical and biological aspects of the cell flow. It comprises descriptions of elasticity of the cell membrane, cell-cell interactions, two-way cell-fluid interaction, and adhesion of cell to surfaces.

Using this model, we analyze different processes involving flow of cells. We present the results of ongoing research concerning the development of model for cell adhesion, the analysis of microfluidic devices with periodic obstacle arrays, the optimization of microfluidic connectors and biological process of red blood cells formation from reticulocytes.

Keywords: *Numerical model, simulation, red blood cell, fluid flow.*

1. Introduction

Currently, blood flow modeling is very useful in various applications. It may be highly abstract (e.g. using electromechanical analogies to model arterial trees [1]) or very realistic (e.g. comparing simulations and experiments of passage of cells through a sensor that measures their deformability [2]). In this work, we describe one such model, which includes a homogeneous fluid - blood plasma - and moving objects immersed in it - red blood cells (RBCs). RBCs are very elastic and therefore, in modeling on single-cell level, it is very important to capture their elastic properties. The model further includes adhesion, cell-cell interactions and two-way cell-fluid interaction.

We use this model to simulate processes inside microfluidic devices. These lab-on-a-chip devices include various types of microchannels that may also contain arrays of obstacles. The blood flows through the micro-channels and it is possible to sort the cells or capture specific rare cells from the total cell populations. This kind of separation or capture is useful in diagnostics and monitoring of various diseases (e.g. cancer, sickle-cell anemia, malaria). More details on modeling blood flow in microfluidic devices can be found in [3].

Our goal is to optimize the design of such micro-channels. We investigate the connectors that are used for inflow into microfluidic devices. In general, the cells in these connectors undergo significant shear stress that may lead to their damage. We are working on deriving a modified blood damage index that would allow future optimization of such connectors. In case of channels that include periodic obstacle arrays (POA), we focus on the capture rate of rare cells, should the surface of obstacles be coated with rare-cell antibodies. We explain the adhesion model that can be used for modeling the capture. And finally, we look at a practical problem of modeling the formation of RBCs that are used for further simulations. The combined progress in these partial topics will bring us closer towards development of improved microfluidic devices.

2. Model

Our model consists of two main parts: fluid and immersed elastic objects. The fluid is modeled using lattice - Boltzmann method [4] and the objects using immersed boundary method and spring network model of the membrane. This boundary, i.e. object's surface, is covered with triangular mesh. Each mesh node is connected with several neighboring nodes and there are elastic forces corresponding to five elastic moduli, which are responsible for elastic behavior of the objects. They preserve the surface area (both locally and globally), volume and shape of the object. This model is implemented as Object-in-fluid framework in open-source

* [1,4]Ivan Cimrak, [1,4]Katarina Bachrata, [1,4]Hynek Bachraty, [2,4]Iveta Jancigova, [1,4]Renata Tothova, [1,4]Martin Busik, [1,4]Martin Slavik, [3,4]Markus Gusenbauer
[1]Department of Software Technologies, Faculty of Management Science and Informatics, University of Zilina, Slovakia,
[2]Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Zilina, Slovakia
[3]Center for Integrated Sensor Systems, Danube University Krems, Austria
[4]Cell-in-fluid Research Group, Slovakia
E-mail: ivan.cimrak@fri.uniza.sk

scientific package ESPResSo [5]. More details on the model can be found in [6].

The model of red blood cell was calibrated by comparing the simulation results with results from stretching experiments of biological red blood cells [7]. We have also investigated proper mass distribution in the immersed boundary points and seeding of dense suspensions of cells, which are necessary for simulations with high hematocrit [8].

3. Applications of RBC model

3.1. Optimization of device inlets

Cell sensitive applications like dynamic cell culturing rely on a predefined shear stress. Deviations from this optimal state lead to unwanted cell activation and even cell destruction. Microfluidic channels usually lack the information about cell activation or destruction. Complex geometries inside microfluidic devices increase the deviation from a blood cells' native condition. Recent efforts to quantify the blood cell damage concerned conventional blood pumping devices and a blood damage index (BDI). The BDI sums up the local shear stress along streamlines inside a microfluidic channel to estimate possible mechanical stress, leading to activation and/or hemolysis of RBCs. A conventional BDI calculation uses finite element solvers for Navier-Stokes equations to estimate the shear stress distribution of the whole blood volume. The blood damage index calculation typically uses $BDI = \sum_{inlet}^{outlet} K \cdot \Delta t^a \cdot \tau^b$ with time *t*, shear stress τ and empirical constants *K*, *a* and *b* (e.g. [9]).

Publications [10] and [11] summarize several attempts to relate the blood damage index to hemolysis of cells (destruction of cell membrane). However BDI often fails to reproduce experimental hemolysis [11].

A certain flow geometry can be optimized in such a way that BDI is minimized. Using the Navier-Stokes equations, the blood is considered as a continuum, there are no mutual interactions of cells or cells with fluid. Since in our model we have implemented these types of interactions, we introduce a cell damage index (CDI), in which the single cell behavior is included. There are several possibilities, how to specify the CDI. The RBC membrane can withstand a finite strain, beyond which it ruptures. The strain is related to both shape and area change of the given object. In our model, we know the current deviation of global area of each cell at any given time. We use this for an approximation of total strain. Moreover, the structural damage of the membrane is influenced not only by the strain magnitude, but also by the duration of strain. In [12], the authors claim that the exposure time of areal strain plays a significant role in the cell rupture, therefore we consider cumulative value of global area deviation during the time, when the cell passes a fluidic channel. Since the values for any individual cell can vary significantly compared to other cells in the same simulation, we consider average values from several simulations with different number of cells.

Inlets and outlets of microfluidic devices, where the flow is redirected, may be a major source of shear stress which acts on the cells. The relative deviation of global area of a cell is computed as $\Delta S/S_0$, where S_0 is relaxed area of the cell. This value is cumulated over the time the cell passes the channel and we define the cumulative deviation of global area as

$$CDG(cell) = \sum_{time} \frac{\Delta S}{S_0} \Delta t$$

We randomly seeded 50 cells inside such a typical inlet of a commercially available analysis chamber (Fig. 1). We run 10 simulations, each with a different seeding of cells. For each cell we computed relative deviation of global area at any given time. Fig. 2 shows a histogram of *CDG* for 50 cells in 10 simulations. The average of these values describes the average damage for the cells and it is a good candidate for CDI of the channel. Therefore we define

$$CDI = \frac{1}{n_{cells}} \sum_{i=1}^{n_{cells}} CDG(cell_i)$$

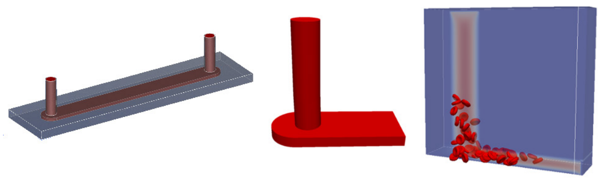


Fig. 1 Scheme of a micro-channel (left), design of simulation geometry (center), simulation with 50 cells (right)

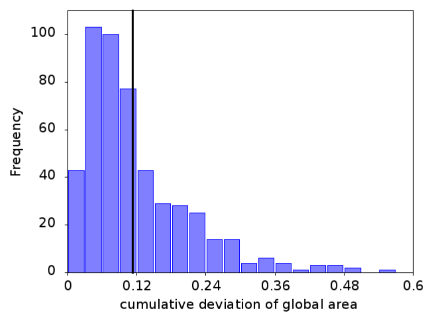


Fig. 2 Histogram of cumulative deviations of global area, values from 10 simulations, 50 cells each. Black line represents CDI.

The black line in Fig. 2 denotes CDI. We found a certain amount of cells with low damage, plenty of cells with mid-range damage and a few cells with high damage. Such histograms can indicate the quality of a microfluidic channel. Less damage is indicated by shift of the black line to the left. Further work will include different inlet geometries and we will compare simulation results of the standard continuum model (Navier-Stokes equations with BDI) and our Object-in-fluid model with newly developed CDI.

3.2. Estimating capture rates using simulations of periodic obstacle arrays

There are several applications in biomedicine that rely on isolation of rare cells (e.g. circulating tumor cells - CTCs) from large populations of other cells. This isolation can be achieved using periodic microfluidic obstacle arrays, which use various physical or biological properties of cells: some of them work on the principle of transverse displacement, others use antibody coated surfaces that capture the rare cells if these come into contact with the coated surface for sufficiently long time.

One of the assumptions sometimes used in modeling of periodic obstacle arrays (POAs) is that it is sufficient to know the rare cell's radius and its offset from obstacles that identifies the future trajectory, in order to determine the cell's collision mode and collision rate, e.g. using Lagrangian tracers in [13].

We studied the interplay of column radius and hematocrit and their combined influence on the number and duration of the rare cell - column contacts. We found that the interactions of rare cell with other cells may have impact on the contacts with obstacles, especially in dense suspensions, where the interactions may result in displacement of rare cell from the average expected trajectories [14].

We simulated a chamber periodic in $x-$ and $y-$ directions with cylindrical obstacles, Fig. 3. Due to the large computational complexity of the problem, we simulated one passage of the rare cell through the chamber over the range of parameters (varied column radius and number of RBCs, different initial CTC positions at the left side of the chamber, different random seedings of RBCs). We traced the distance of one CTC from the nearest obstacle and the duration of contact(s), if any. We then analyzed the CTC trajectories an determined that regardless their initial position, in RBC-sparse simulations they tend to fall into a small number of typical trajectories, Fig. 5. While we observe the same general pattern in simulations with large number of RBCs, Fig. 6, the trajectories are perturbed by the CTC-RBC interactions, which may influence the capture rate, as evident in Fig. 4. The capture rate was calculated using

$$P(\tau, A_c) = a \cdot A_c \cdot \exp\left(-b\frac{\tau}{A_c}\right),$$

where is the cell-obstacle contact area, is shear stress with respect to the closest obstacle. Constants *a* and *b* are lumped parameters determined using [15]. They include several biological quantities such as ligand and receptor densities, association constants, characteristic length of the ligand–receptor bond and others.

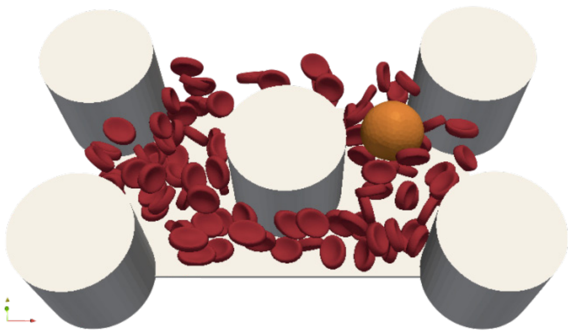


Fig. 3 Periodic chamber with cylindrical obstacles, red blood cells and a rare cell.

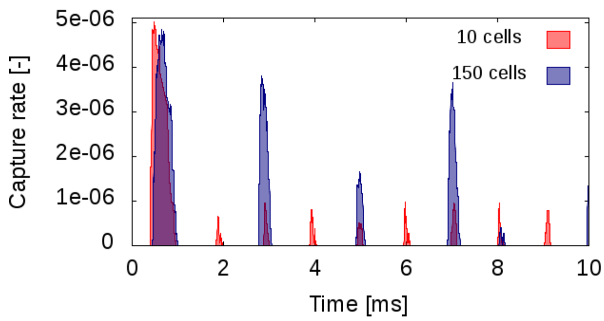


Fig. 4 Capture rate is significantly different in simulations with small and large number of RBCs.

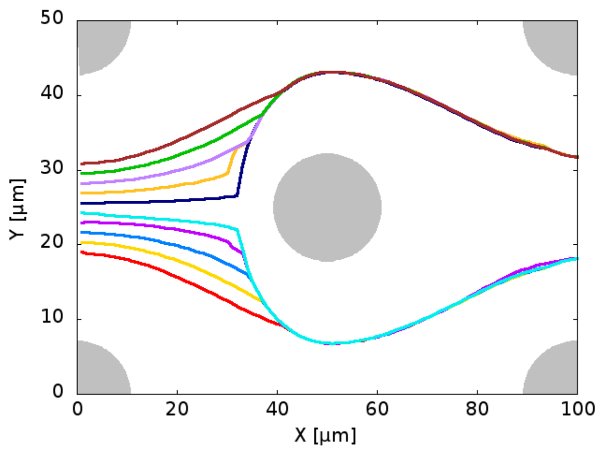


Fig. 5 In simulations with small number of RBCs, rare cells fall into predicted trajectories. Each line corresponds to the trajectory of center of rare cell in a different simulation.

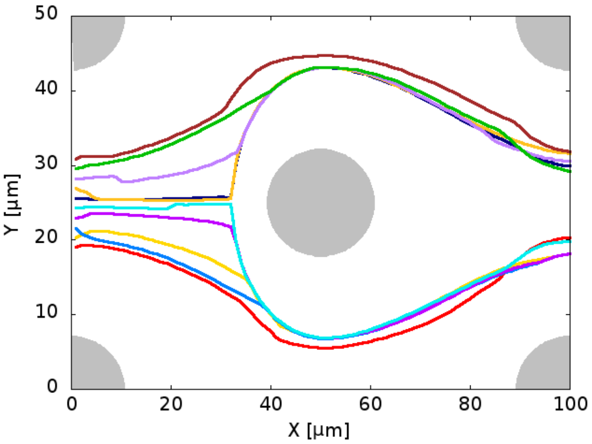


Fig. 6 In simulations with large number of RBCs, rare cells follow the same general behavior, but the trajectories are perturbed by the nearby RBCs. In this case, the capture rate may significantly differ.

This leads us to conclude that it is important to include actual cells and their cell-cell interactions when computing the capture rates in microfluidic channels with periodic obstacle arrays. In the future, this approach will be verified by comparisons with model that includes adhesion as described in the following section.

3.3. Simplified adhesion model

Modeling of the adhesion and rolling of biological cells on functionalized surfaces gives valuable insights into rare cell isolation. We study cells moving in shear flow above a wall to which they can adhere via specific receptor-ligand bonds based on receptors from selectin as well as integrin family. The adhesion mechanism is modeled by adhesive bonds including stochastic rules for their formation and rupture. Once the cell comes close enough to such a wall, receptors and ligands start forming bonds, slowing the motion of the cell down [16]. We explore a simplified model with dissociation rate independent of the length of the bonds.

Before a receptor and a ligand can form a bond, their physical locations need to be brought in close vicinity and a bond can be created with certain probability. Reversely, any bond can rupture, again with certain probability. Finally, whenever receptor and ligand positions move apart more than a certain threshold, the bond is broken with probability one, see [17].

The computational model of a cell described in previous sections can be readily extended with any other phenomenon based on forces acting on the boundary mesh points of the elastic object. The adhesion mechanism is based on bonds between the receptor site on a cell and the ligand site on a wall. This bond can be modeled as a harmonic spring. Once the stiffness of such spring is known, the bond exerts repulsive or attractive force $F_{aff} = \kappa \cdot l$ on the corresponding mesh points.

Dissociation and association dynamics are represented by probability of bond formation and bond breaking. At the moments when the receptor and ligand are close enough, the bond can be created with a specific rate k_{on} and the probability that this happens during time interval δt is given by P_{on}. On the other hand, any bond can rupture with the rate k_{off} with dissociation probability P_{off} over time.

$$P_{on} = 1 - \exp(-k_{on} \cdot \delta t),\ \ P_{off} = 1 - \exp(-k_{off} \cdot \delta t)$$

In [18], we have introduced the simplification of previously known adhesion model from [17]. The model from [17] includes the dependence of k_{off} on the magnitude of the force exerted on the bond. It is a natural assumption, however, the dependence of k_{off} on the bond prolongation is quite complicated, includes parameters such as bond detachment force and reactive compliance that are quite difficult to measure, and are not available for many receptor-ligand couples. Our simplification is based on the fact that k_{off} is a fixed constant.

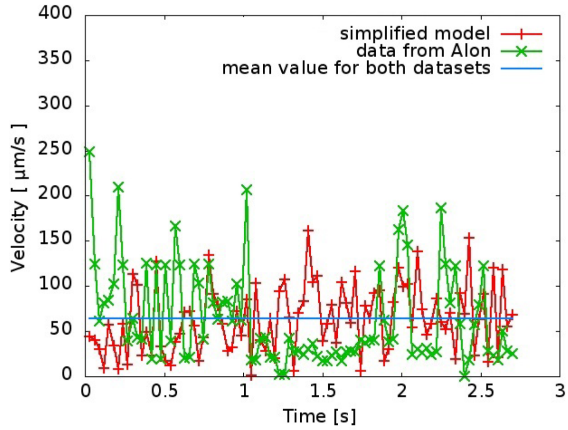


Fig. 7 Velocity reconstruction over the time period of 3s. The mean value of experimental and computational data is the same.

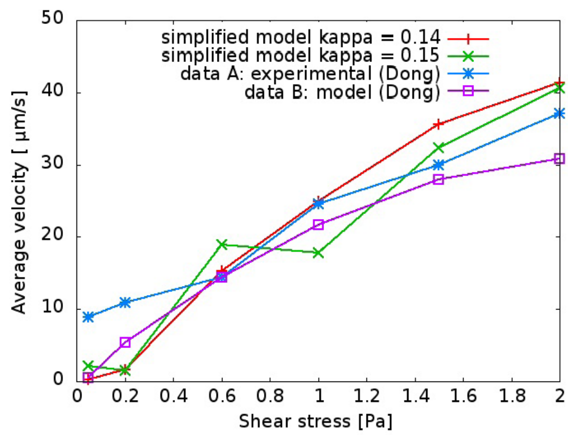


Fig. 8 Averaged velocity of the cell over a certain time period depends on the wall shear stress. Data A (experimental) and data B (model) are taken from [19].

Verification of the simplified model

To show that this model is capable of capturing the mesoscopic properties, such as velocity of rolling cells, we perform two computational experiments. In the first experiment, we test whether our model can capture statistical properties of velocity profile. We track the velocity of individual cells during the adhesion movement. The movement of adherent cell is much slower than that of non-adherent cell due to bonds between ligands and receptors. Since the bonds on a cell in the downstream brake and bonds in the upstream are continuously created, the cell rolls on the surface. The velocity of the cell itself randomly oscillates according to how often bonds brake and how often they are created again. Because of the stochastic nature of bond formation, every simulation run gives different velocity profile. However, one can look at the mean value and the standard deviation and compare the simulated movement with the measured data for real cells.

The second experiment reveals that our model can capture the dependence of cell velocity on the different fluid shear stresses. Namely, the question is whether one set of parameters gives the same values of averaged velocity of the cell for different fluid shear stresses. Detailed description of the computational experiments is given in [18]. We present the result in Figs. 7 and 8. In [20], the authors performed laboratory experiments and obtained the rolling velocity of the cells. Our model is capable of reconstructing these velocities. In Fig. 7, we can see the measured velocities depicted by green symbol cross and the velocities obtained from our model depicted with red symbols plus. Note that we were able to set the model parameters so that both velocity profiles have the same mean value.

In [19], the authors investigated the mechanics of leukocyte adhesion to endothelial cells using in vitro side-view flow assay. The authors showed how the averaged velocity of the cell depends on shear stress. We demonstrate that we are able to capture this dependence. The experiment was performed for shear stresses ranging from 0 to 2 Pa. For each value of stress, we tracked the movement of the cell and we computed the averaged velocity over the time 1ms. The results are presented in Fig. 8.

3.4. Modelling blood cell formation

Overall, RBCs represent up to 95% of the solid blood component volume. Therefore, correct modeling of RBC behavior is the key to the correct blood flow simulations. A natural validation of the RBC model is based on its specific properties associated with RBC formation process. It is known that the shape of RBC with a given volume maximizes its surface. We also know that RBC is formed from the spherical shape of the proerytroblast or reticulocyte, with the aim to maximize the surface-to-volume ratio.

The model we use, allows us to simulate and explore the properties of the process, using both the inflation and deflation scenarios. We based the first one on an elastic model of a sphere with the volume of a real RBC and with a geometrically corresponding, smaller surface. Using the global area conservation forces, we relaxed this triangulation into a shape with the surface of RBC. With the second scenario, we deflated a sphere with the RBC surface and larger volume to smaller RBC volume, using the global volume elastic modulus. In both cases, we achieved shapes that visually correspond to real RBCs (Figs. 9 and 10).

The simulation accuracy can be firstly verified by looking at basic characteristics such as surface and volume. Shapes obtained by inflation and deflation scenarios intrinsically have the surface and volume consistent with biological RBC. Another accuracy check of the obtained cell is verification of its curvature. Since the cell surface is discretized by a triangular mesh, there is a natural opportunity to characterize its curvature using the distribution of inner angles formed by the incidental triangles of the triangulation. We looked more closely at statistics of this distributions.

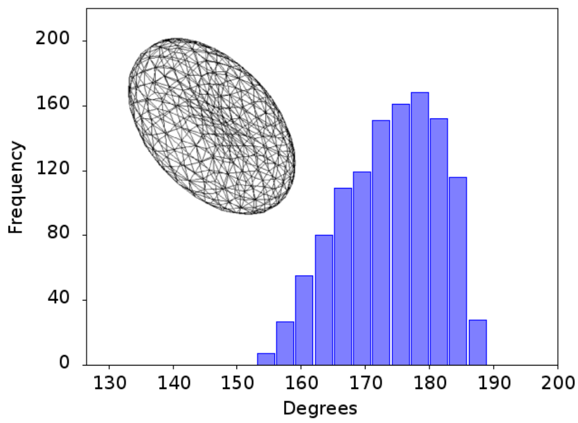


Fig. 9 Inflate scenario, shape and histogram of angles distribution

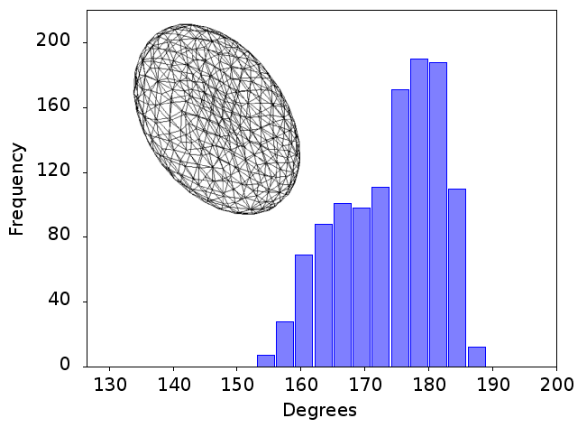


Fig. 10 Deflate scenario, shape and histogram of angles distribution

Figures 9 and 10 show the histograms of these angles for the inflation and deflation scenarios, and Table 1 shows the maximum, minimum, average and standard deviation of those angles. Note that this concerns the interior angles: the values above 180° correspond to the concave parts of the cell surface and the values below 180° correspond to the convex parts of the cell surface. The obtained data indicate that the two processes, where different elastic moduli were dominant, result in remarkably similar distributions. This result strengthens the confidence in consistency of the model.

Statistical values of angle distribution Table 1

	min	max	avg	stdev
Inflate scenario	153.60	188.86	174.03	7.68
Deflate scenario	153.14	188.08	173.98	7.52
Deformation	148.69	197.18	173.91	8.38
Triangulated RBC	140.79	191.67	173.66	11.17

With regard to the stochastic nature of the simulations, we also confirm this result by the statistical tests of the selection from identical empirical distribution. The necessary precondition of the results comparison is using comparable regular meshes with almost the same number of nodes and adjacent triangles. Since in our model we can easily use various regular meshes, we do not have to deal with the complicated problem of normalizing the obtained curvature spectrum for significantly different triangulations.

We tested the consistency of the RBC curvature characterization by looking at the angles distribution statistics in two more ways. The first one involved a simulation experiment and examined the RBC deformation during its flow through a narrow opening (Fig. 11, Table 1). As expected, the values of the convex and concave angles in the central deformed part of the RBC have larger range. The histogram shape and the increase of the standard deviation correspond to that.

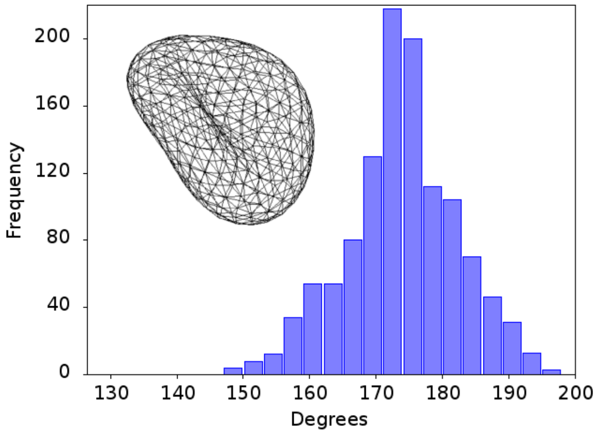


Fig. 11 Deformation, shape and histogram of angles distribution

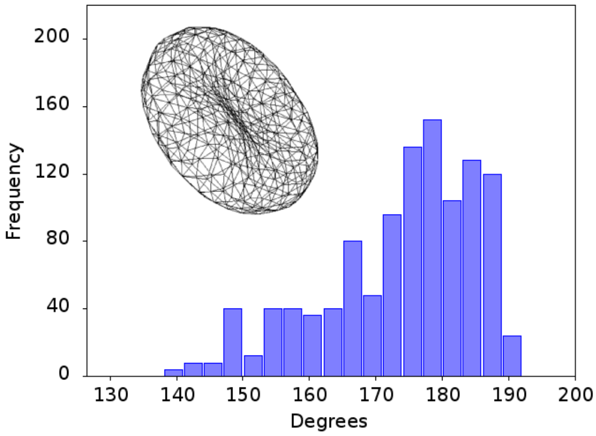


Fig. 12 Triangulated RBC, shape and histogram of angles distribution

The second was a comparison of the curvature histogram of the RBC created from sphere to its known biological shape. However, here we are confronted with the problem of the existing triangulations of this surface. They exhibit considerable regularity, which is reflected in the histograms and they are difficult to compare to our significantly stochastic triangulations (Fig. 12,

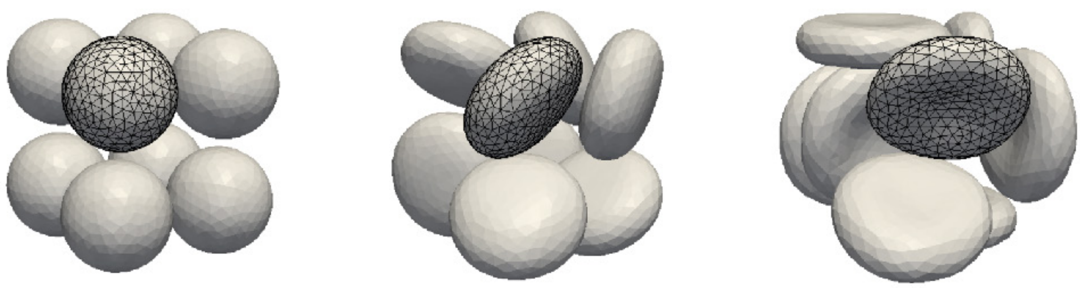


Fig. 13 RBC seeding

Table 1). For the time being, obtaining a sufficiently irregular and comparable RBC mesh remains an open task.

The RBC formation process may solve the problem of RBC seeding for large-scale simulations, where long warm-up time means inefficient usage of the calculation time. Generating random RBC positions, especially while keeping a realistic 45% hematocrit, is complicated when we want to have non-overlapping distribution of partly deformed cells. A solution to this is using the RBC inflation. At the beginning of a simulation, non-overlapping spheres of smaller volumes are placed randomly in the domain. During the warm-up, the spheres are inflated into RBC volume and shape. Here, the required hematocrit and, due their collisions, stochastic distribution of cell positions and shapes are obtained. At the same time, thanks to the repelling cell-cell forces, the RBCs do not overlap. An example of such seeding is shown in Fig. 13.

4. Conclusion

Our computational model of a red blood cell has proved to be a strong tool for investigation of cell behavior. Once the confidence in the model has been established by its calibration and verification, new conclusions may be drawn by utilizing this model without laboratory experiments. We have shown several examples of such approach: the capture rates can be computed for channels with periodic obstacle arrays, the shape of microfluidic connectors analyzed in terms of cell deterioration.

Further computational experiments from section 3.4 confirmed the consistency of the model. Using two different scenarios for formation of a red blood cell, we have obtain objects with the same statistical properties. Thus inflation and deflation scenario give comparable results. This justifies the use of inflation scenario to seed cells into dense suspensions: we have successfully created dense suspension by starting with smaller objects and gradually increasing their surface to real values.

Acknowledgement

I. Cimrak, K. Bachrata, H. Bachraty, I. Jancigova, R. Tothova and M. Busik were supported by the Slovak Research and Development Agency under the contract No. APVV-0441-11.

The research of I. Cimrak was supported by Marie-Curie grant No. PCIG10-GA-2011-303580.

The research of M. Gusenbauer was supported by NFB Austria through Life Science Calls, grant No. LSC13-024.

References

[1] BORIK, S., CAP., I., BABUSIAK, B.: Resonant Frequencies of Small Arterial Segments as Determining Factors for Estimation of Terminal Segments in Electromechanical Arterial Tree Model, *Communications - Scientific Letters of the University of Zilina*, vol. 17, No. 1A, pp. 66-70, 2015.

[2] TATSUMI, K., KATSUMOTO, Y., FUJIWARA, R., NAKABE, K.: Numerical and Experimental Study on the Development of Electric Sensor as For Measurement of Red Blood Cell Deformability in Microchannels, *Sensors*, vol. 12, 2012.

[3] BACHRATÝ, H., BACHRATÁ, K.: *On Modeling Blood Flow in Microfluidic Devices,* 10th Intern. Conference on Elektro, pp. 518-521, 2014.

[4] LADD, A. J. C.: Numerical Simulations of Particulate Suspensions Via a Discretized Boltzmann Equation, Part 2 Numerical Results, *J. of Fluid Mechanics*, vol. 271, pp. 311-339, 6 1994.

[5] ARNOLD, A., LENZ, O., KESSELHEIM, S., WEEBER, R. FAHRENBERGER, F., ROEHM, D., KOSOVAN, P., HOLM, C.: ESPResSo 3.1 – Molecular Dynamics Software for Coarse-Grained Models, *Meshfree Methods for Partial Differential Equations VI* , vol. 89 of *Lecture Notes in Computational Science and Engineering*, pp 1-23, Editors: M. Griebel and M. A. Schweitzer, 2013. Springer.

[6] CIMRAK, I., JANCIGOVA, I., TOTHOVA, R., GUSENBAUER, M.: Mesh-Based Modeling of Individual Cells and their Dynamics in Biological Fluids, *Applications of Computational Intelligence in Biomedical Technology,* ser. Studies in Computational Intelligence, Eds. Springer International Publishing, 2015, vol. 606, pp. 1-28.

[7] TOTHOVA, R., JANCIGOVA, I., BUSIK, M.: *Calibration of Elastic Coefficients for Spring-Network Model of Red Blood Cell,* Intern. Conference on Information and Digital Technologies (IDT), July 2015, pp. 376-380.

[8] JANČIGOVÁ, I.: Modeling Elastic Objects in Fluid Flow with Biomedical Applications, Ph.D. dissertation, University of Zilina, 2015.

[9] LI, H., RUAN, X. QIAN, W., FU, X.: Numerical Estimation of Hemolysis from the Point of View of Signal and System, *Artificial Organs*, vol. 38, No. 12, pp. 1065-1075, 2014.

[10] TASKIN, M. E., FRASER, K. H., ZHANG, T., WU, C., GRIFFITH, B. P., WU, Z. J.: Evaluation of Eulerian and Lagrangian Models for Hemolysis Estimation, *Asaio J.*, vol. 58, No. 4, pp. 363-372, 2012.

[11] CHAN, W., WONG, Y., DING, Y., CHUA, L., YU, S.: Numerical Investigation of the Effect of Blade Geometry on Blood Trauma in a Centrifugal Blood Pump, *Artificial Organs*, vol. 26, No. 9, pp. 785-793, 2002.

[12] LEE, S., AHN, K., LEE, S., SUN, K., GOEDHART, P., HARDEMAN, M.: Shear Induced Damage of Red Blood Cells Monitored by the Decrease of their Deformability, *Korea - Australia Rheology J.*, vol. 16, No. 3, pp. 141-146, 2004.

[13] SMITH, J., LANNIN, T., SYED, Y., SANTANA, S., KIRBY, B.: Parametric Control of Collision Rates and Capture Rates in Geometrically Enhanced Differential Immunocapture (Gedi) Microfluidic Devices for Rare Cell Capture, *Biomedical Microdevices*, vol. 16, No. 1, pp. 143-151, 2014.

[14] GLEGHORN, J. P., SMITH, J. P. , KIRBY, B. J.: Transport and Collision Dynamics in Periodic Asymmetric Obstacle Arrays: Rational Design of Microfluidic Rare-Cell Immunocapture Devices, *Phys. Rev. E*, vol. 88, p. 032136, 2013.

[15] DECUZZI, P., FERRARI, M.: The Adhesive Strength of Non-Spherical Particles Mediated by Specific Interactions, *Biomaterials*, vol. 27, No. 30, pp. 5307-5314, 2006.

[16] SPRINGER, T. A.: Traffic Signals for Lymphocyte Recirculation and Leukocyte Emigration: the Multistep Paradigm, *Cell*, vol. 76, No. 2, pp. 301-314, 1994.

[17] KORN C. B., SCHWARZ, U. S.: Dynamic States of Cells Adhering in Shear Flow: From Slipping to Rolling, *Phys. Rev. E*, vol. 77, Apr 2008.

[18] CIMRÁK, I.: A Simplified Model for Dynamics of Cell Rolling and Cell-Surface Adhesion, AIP Conference Proceedings, vol. 1648, No. 1, 2015, Art. ID 210005.

[19] DONG, C., LEI, X. X.: Biomechanics of Cell Rolling: Shear Flow, Cell-surface Adhesion, and Cell Deformability, *J. of Biomechanics*, vol. 33, No. 1, pp. 35-43, 2000.

[20] ALON, R., CHEN, S., PURI, K., FINGER, E., SPRINGER, T.: The Kinetics of L-selectin Tethers and the Mechanics of Selectin-Mediated Rolling, *The J. of Cell Biology*, vol. 138, No. 5, pp. 1169-1180, 1997.

Jaroslav Janacek - Peter Marton - Matyas Koniorczyk *

THE COLUMN GENERATION AND TRAIN CREW SCHEDULING

Better productivity and efficiency is more and more required in the railway operation. The train crew management is one of the several problems that could be solved using mathematical methods. Crew management is a problem that is well-known in Operations Research. We compare two approaches for solving the train crew scheduling problem. The first approach consists of solving the original problem by single model. The second approach corresponds to the step-by-step column generation. This technique was originally based on Dantzig-Wolfe decomposition. The benchmarks used for comparison of both approaches originate in real problems from railway systems in Slovakia and Hungary.

Keywords: *Train crew scheduling, Dantzig-Wolfe decomposition, column generation.*

1. Introduction

The main aim of crew management is to build the shift schedules of crews for covering a timetable that is planned. This is an activity which concerns the medium-term use of the available resources. It is broadly known as tactical planning. Crew management is problem that is well-known in Operations Research. In last decades, after regulations of the European Commission for railways (so called "railway packages"), railway applications of crew management come on the scene. Main results of the "railway packages" are deregulation and privatization of the rail industry in Europe. Both factors are increasingly pervading the rail industry. Better productivity and efficiency are strongly required. Due to these facts the railway operators are more interested in using effective techniques as before.

The crew-scheduling problem can be solved using mathematical programming techniques. These techniques are based on the application of a set covering model with a number of additional constrains. In the airline industry such models are popular for solving crew scheduling problems [1, 2, 3 and 4]. However, in the railway industry the sizes of the crew scheduling instances are a magnitude larger than in the airline industry, which prohibited the application of these models in the railway industry until recently. However, developments in hard- and software nowadays enable the application of such models in the railway industry as well [5, 6, 7 and 8]. Abbink [9], for instance, describes the successful application of a model and the corresponding solution techniques for scheduling more than 6,500 drivers and conductors of the Dutch railway operator.

Solving of the train crew scheduling problem by means of mathematical programming faces large dimensionality of the models describing the associated large instances of the problem. Especially in the case when the problem is described by one complete model, which covers all characteristics of the crew scheduling problem of a company, may lead to a large linear programming problem formulation, which can attack the limits of a common IP-solver. To overcome this potential threat, column generation technique can be used. This technique was originally based on Dantzig-Wolfe decomposition [10], when only small portion of sub-problem extremal points was considered and the original problem was reduced to a smaller one, which took a form the partitioning problem. As the solving process of the reduced problem usually provided a solution far from the optimal one, the process was completed by another process, which enables to recognize whether there is some other extremal point of the sub-problems, which can improve the current reduced problem solution. The process is called column generation method, and it enables not only to recognize existence of the improving solution, but even to find it. The column generation approach has been discussed and applied in connection with a broad spectrum of optimization problems from practice [4, 11 and 12].

This paper focuses on exploring the possibility of the train crew schedule designing by means of mathematical programming, including the comparison of two approaches: the first one consists of solving the original problem described by a model and the second approach corresponds to the step-by-step column generation. The benchmarks used for the comparison originate

* [1]**Jaroslav Janacek,** [1]**Peter Marton,** [2]**Matyas Koniorczyk**
[1]Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Zilina, Slovakia
[2]Department of Applied Mathematics, Faculty of Natural Sciences, University of Pecs, Hungary
E-mail: jaroslav.janacek@fri.uniza.sk

in real problems from crew scheduling in railway transportation systems [13].

The rest of the paper is organized as follows. The second Section describes the train crew scheduling problem by a complete mathematical programming model and an analysis of the specific model structure, facilitating the formulation of individual sub-problems. The third Section introduces the Dantzig-Wolfe decomposition and defines so called master problem corresponding to the reduced problem. The fourth Section is devoted to an explanation of column generating principle in connection with the crew scheduling problem. The fifth Section contains the description of benchmarks used for numerical experiments including the discussion of associated results. The obtained findings and conclusions are summarized in Section 6.

2. Train crew scheduling problem

Let J denote set of jobs, which must be covered by scheduled shifts of individual crews so that each job is covered by exactly one work shift of a crew. The jobs are representing the trains (or working trips of a crew member with a given train) in the real operation. Each job (train) is determined by the beginning and ending times and network nodes. Two successive jobs (trains) in a shift of one crew must enable transit of the crew from the ending node of the preceding job to the starting node of the successive job in the associated time interval. A crew shift begins at a determined time and network node and it must also finish at possibly another network node and at some higher ending time. These beginnings and endings of a shift are called auxiliary jobs and they form a so called frame of the shift. Let K denote set of all possible shift frames. Each shift with frame $k \in K$ has to start with an auxiliary job $b(k)$ and to finish with an auxiliary job $e(k)$. Let $J(k) \subseteq J$ be the set of jobs, which can possibly be covered by a shift within the frame k, i.e. each job $j \in J(k)$ can be reached from the auxiliary job $b(k)$ and the auxiliary job $e(k)$ can follow the job j. Each frame $k \in K$ is accompanied by set $R(k)$ of $m(k)$ ordered pairs (p, s), where p and s are auxiliary or original jobs, s can be performed right after p, and they both may succeed within a shift of frame k , see Fig. 1 for better understanding the used denotation. The crew transit from p to s is priced by value $c^k_{(p,s)}$. The price of a shift is sum of prices of all pairs forming the shift. The train crew scheduling problem consists of finding such collection of crew shifts that each job $j \in J$ belongs to exactly one crew shift and the total price of the collection is minimal. (Alternatively, one may consider the case when each job $j \in J$ belongs to at least one crew shift, thus allowing for additional auxiliary travels of the crew. Though we do not consider this case here, we shall treat auxiliary travels in a different way, our model can be easily modified to describe this case.)

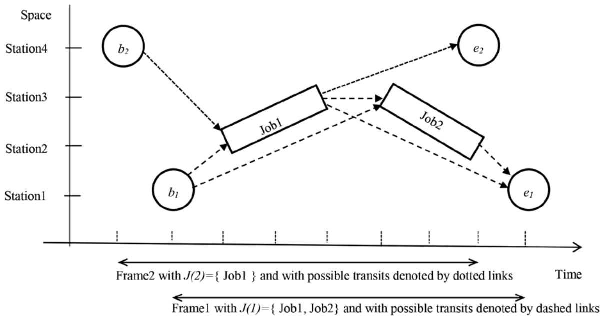


Fig. 1 A simple example consisting of two jobs and two frames and the associated two systems of possible transits within the frames. It holds $R(1)=[(b_1, Job1), (b_1, Job2), (Job1, Job2), (Job1, e_1), (Job2, e_1)]$ and $R(2)=[(b_2, Job1), (Job1, e_2)]$.

Let us introduce the variables $y^k_{(p,s)} \in \{0, 1\}$ for each $k \in K$ and $(p, s) \in R(k)$. The variable takes the value of one if and only if the job p is directly succeeded by job s in the crew shift of the frame k. Then our model can be formulated as follows.

$$Minimize \sum_{k \in K} \sum_{(p,s) \in R(k)} c^k_{(p,s)} y^k_{(p,s)} \quad (1)$$

$$Subject\ to \sum_{k \in K} \sum_{(p,j) \in R(k)} y^k_{(p,j)} = 1 \ for \ j \in J \quad (2)$$

$$\sum_{(p,j) \in R(k)} y^k_{(p,j)} = \sum_{(j,s) \in R(k)} y^k_{(j,s)} \quad for \quad k \in K, j \in J(k) \quad (3)$$

$$\sum_{(b(k),s) \in R(k)} y^k_{(b(k),s)} \leq 1 \quad for \quad k \in K \quad (4)$$

$$y^k_{(p,s)} \in \{0,1\} \quad for \quad k \in K, (p,s) \in R(k) \quad (5)$$

Constraints (2) of the model (1)-(5) ensure that each original job belongs to exactly one of the suggested shifts. The conservative constraints (3) assure that the number of jobs preceding job j is the same as the number of succeeding jobs in the shifts of the frame k. The constraints (4) allow only one shift in the frame k.

In order to formulate the model (1)-(5) in matrix form, we introduce several incidental matrices.

Let a matrix $\boldsymbol{A}^k$ with $|J|$ rows and $m(k)$ columns consist of binary coefficients $a^k_{j,(p,s)}$, defined for each $j \in J$ and $(p, s) \in R(k)$, where the coefficient is equal to one if and only if $j=s$ and it equals zero otherwise.

Let a matrix $\boldsymbol{D}^k$ with $|J(k)|$ rows and $m(k)$ columns consist of coefficients $d^k_{j,(p,s)}$, defined for each $j \in J$ and $(p, s) \in R(k)$, where the coefficient is equal to one if and only if $j=s$, it equals minus one if $j=p$ and it equals to zero otherwise.

Let a vector $\boldsymbol{E}^k$ with $m(k)$ components consist of binary coefficients $e^k_{(p,s)}$, defined for each $(p, s) \in R(k)$, where the coefficient is equal to one if and only if $b(k)=p$ and it equals zero otherwise.

Let the vector $\boldsymbol{C}^k$ with $m(k)$ components denote vector of costs coefficients $c^k_{(p,s)}$. If we decompose the vector $\boldsymbol{y}$ of the decision variables $y^k_{(p,s)}$ to set of sub-vectors $\boldsymbol{y}^k$, which elements correspond to sets *R(k)* for $k \in K$, we can write the model (1)-(5) in the following form.

$$Minimize \ \sum_{k \in K} \boldsymbol{C}^k \boldsymbol{y}^k \tag{6}$$

$$Subject\ to \ \sum_{k \in K} A^k \boldsymbol{y}^k = 1 \tag{7}$$

$$\boldsymbol{D}^k \boldsymbol{y}^k = 0 \quad for \quad k \in K \tag{8}$$

$$\boldsymbol{E}^k \boldsymbol{y}^k \leq 1 \quad for \quad k \in K \tag{9}$$

$$\boldsymbol{y}^k \in \{0, 1\}^{m(k)} \quad for \quad k \in K \tag{10}$$

Let us consider the *k*-th sub-problem for a general set of the objective function coefficients in the following form.

$$Minimize \ \boldsymbol{C}^k \boldsymbol{y}^k \tag{11}$$

Subject to (8)-(10)

The sub-problem corresponds to the task of finding the cheapest path through the sub-network determined by the arcs of *R(k)* on the nodes corresponding to auxiliary and original jobs. The path has to start at *b(k)* and finish at *e(k)* due to the conservative constraints and the fact that the network is acyclic. The sub-problem obeys the integrity property, that is, all the extreme points of the feasible solution set of LP-relaxation are integers. Furthermore, we can claim that the set of feasible solutions of the LP-relaxation is non-empty (it contains zero-solution $\boldsymbol{Y}^k_0$) and it is bounded which implies that the set of feasible solutions is a polytope. It follows that each feasible solution $\boldsymbol{y}$ can be expressed as a convex combination of all extreme points $\boldsymbol{Y}^k_r$ r=0, ...,$n(k)$ using nonnegative multipliers λ_{kr}.

$$Thus,\ \boldsymbol{y}^k = \sum_{r=0}^{n(k)} \lambda_{kr} \boldsymbol{Y}^k_r = \sum_{r=1}^{n(k)} \lambda_{kr} \boldsymbol{Y}^k_r,\ where \sum_{r=1}^{n(k)} \lambda_{kr} \leq 1 \ and\ \lambda_{kr} \geq 0\ for\ r = 1, ..., n(k) \tag{12}$$

3. The Dantzig-Wolfe decomposition

Let us substitute (12) for $\boldsymbol{y}^k$ in the LP-relaxation of the model (1)-(2). We obtain the LP problem, where the multipliers λ_{kr} are hereafter the decision variables.

$$Minimize \ \sum_{k \in K} \sum_{r=1}^{n(k)} (\boldsymbol{C}^k \boldsymbol{Y}^k_r) \lambda_{kr} \tag{13}$$

$$Subject\ to \ \sum_{k \in K} \sum_{r=1}^{n(k)} (\boldsymbol{A}^k \boldsymbol{Y}^k_r) \lambda_{kr} = \mathbf{1} \tag{14}$$

$$\sum_{r=1}^{n(k)} \lambda_{kr} \leq 1 \ \ for\ k \in K \tag{15}$$

$$\lambda_{kr} \geq 0 \ \ for\ \ k \in K, r = 1, ..., n(k) \tag{16}$$

This problem formulation itself is of a little use for direct solution of the original problem due to necessity of enumerating all extreme points of each sub-problem. Nevertheless, if a small collection zero-one solutions $\boldsymbol{Y}^k_r$ for $k \in K$, r=0, ..., $P(k)$is known and the problem (13)-(16) has a feasible solution, then we can formulate the column generation process, which informs us at each step whether the optimal solution of (13)-(16) can or cannot be improved by another solution of a sub-problem *k*.

4. Column generation

Let us consider the current problem in the following form:

$$Minimize \ \sum_{k \in K} \sum_{r=1}^{P(k)} (\boldsymbol{C}^k \boldsymbol{Y}^k_r) \lambda_{kr} \tag{17}$$

$$Subject\ to \ \sum_{k \in K} \sum_{r=1}^{P(k)} (\boldsymbol{A}^k \boldsymbol{Y}^k_r) \lambda_{kr} = \mathbf{1} \tag{18}$$

$$\sum_{r=1}^{P(k)} \lambda_{kr} \leq 1 \ \ for\ k \in K \tag{19}$$

$$\lambda_{kr} \geq 0 \ \ for\ \ k \in K, r = 1, ..., P(k) \tag{20}$$

We assume that the individual equalities of (18) are subscribed by $j \in J$. Let us denote by $\boldsymbol{v}^*$ the row vector of shadow costs of the constraints (18) and w_k^* the shadow cost of the constraint (19) corresponding to the shift frame *k*. Then the reduced cost of a variable λ_{kr} in the optimal solution of the problem (17)-(20) can be expressed as $\boldsymbol{C}^k\boldsymbol{Y}^k_r$ - $(\boldsymbol{v}^*(\boldsymbol{A}^k\boldsymbol{Y}^k_r)+w_k^*)$, which must be nonnegative due to the optimality of the solution. If there is any sub-problem solution $\boldsymbol{Y}^k$ which improves the optimal solution of (17)-(20), the inequality (21) must hold:

$$0 > \boldsymbol{C}^k \boldsymbol{Y}^k - \boldsymbol{v}^* (\boldsymbol{A}^k \boldsymbol{Y}^k) - w_k^* = (\boldsymbol{C}^k - \boldsymbol{v}^* \boldsymbol{A}^k) \boldsymbol{Y}^k - w_k^*. \tag{21}$$

Thus an attempt to obtain an improving solution $\boldsymbol{Y}^k$ can be performed by solving the following modified sub-problem:

$$Minimize \ (\boldsymbol{C}^k - \boldsymbol{v}^* \boldsymbol{A}^k) \boldsymbol{y}^k \quad Subject\ to\ (8)-(10) \tag{22}$$

If the optimal value of objective function (22) is less than w_k^*, then the obtained solution is a candidate for the improving solution. If no improving solution for any $k \in K$ exists, then the current collection is complete.

5. Case studies

In this section, we describe benchmarks and experiments realized for demonstration of our approach. We used two sets of benchmarks. The first one uses timetable data and technological data from freight railway operation on one single-track line in Slovakia. The second one is derived from timetable data from passenger railway operation on three lines, mostly the double-track lines, in Hungary. We realized several experiments with the two solving techniques and with several versions of input data from both sets of benchmarks.

5.1 Benchmarks

In order to demonstrate our method in a close-to-realistic situation, we addressed simplified examples, which are, however, based on real train paths gathered from the public freight and passenger timetable data of the Slovak Railways and the Hungarian Railways.

In the first set of benchmarks, we consider freight railway traffic connected mainly with operation of the container terminal in Dunajska Streda. The container trains to and from Dunajska Streda need to transit Bratislava Nove Mesto station. Electric locomotives are exchanged by diesel locomotives in this station.75 trains (jobs) run in single-track line from Bratislava Nove Mesto to Dunajska Streda in both directions during one week. The crews originate from two depots: Bratislava vychod and Dunajska Streda.

The train paths were scraped from the freight trains timetable valid in spring 2011. Majority of container trains run from Bratislava Nove Mesto to Dunajska Streda (four trains) or in opposite direction (five trains). There is one pair of feeder and pick-up trains to Bratislava vychod marshalling yard. There is one pair of direct trains from Bratislava vychod to Dunajska Streda. Some trains run every day in week. Some trains run only during the working days. Some trains run only in one specific day. We assume that each train (job) requires a crew of one person (locomotive driver).

We do not consider real number of crew in the depots. We created several shift frames for each experiment. The number of shift frames is parameter of the associated instance.

We prepared four instances with timetable data from the freight railway benchmark from Slovakia. The distances are denoted as D1...D4 respectively. The first instance is based on eleven jobs (trains) of one day of operation. 28 frames are defined, in one hour period. The second instance is based on 24 jobs of two days of operation. 30 frames are defined, in three hour period. The third instance is based on 36 jobs of three days of operation. Finally, four days of operation are replicated in the fourth instance where 48 jobs are considered. 46 or 62 frames are defined for the third and fourth instances, using three hour inter-shift time.

We had exact technological data at hand. We used these data as follows. Each train with start from Dunajska Streda needs 100 minutes for brake test. The brake test has to be realized before the planned departure time. Times of 40 or 20 minutes are included in the experiments' data too, to consider time needed for take over and hand over of the locomotive or exchange of the locomotive between two successive trains. The starting and ending times of each job are modified by taking necessary technological activities into account.

Having specified the jobs which can be served in a given frame, it was necessary to define the job pairs (p, s) for each frame. The calculation of the cost coefficients referring to the job pairs was necessary too, to use them in the objective function. As for the objective function coefficients, we consider two components – the waiting times calculated for each job pair and transition costs for each job pair. Each waiting time for the job pair was obtained as difference between the ending time of p and the starting time of s. In some cases, traveling of crew without locomotive (e.g. by passenger train or taxi) between Bratislava Nove Mesto, Dunajska Streda and other stations is considered. Time constants of 60 and 120 minutes were used as the transition costs for the job pairs. Zero costs were used in case that transition between p and s is realized in same station.

In the second set of benchmarks, we consider a segment of the suburban traffic around Budapest: the Budapest-Szekesfeherar the Budapest-Pusztaszabolcs, and the Budapest-Dunaujvaros passenger trains. In this case we consider the duties of the passenger railway company, that is, conductors. The considered trains have a common terminal, Budapest Deli Palyaudvar (Budapest-Deli in what follows). Pusztaszabolcs station is between Budapest-Deli and Dunaujvaros terminals, some of the trains return in Pusztaszabolcs, while some of them commute between the two terminals. Budapest-Deli-Szekesfehervar is another important suburban line. The passenger trains commuting on these routes have an almost periodic timetable, but we do not take the advantage of this structure. We assume that the set of trains, that is, those commuting between the listed stations form a closed set, they are served by the same set of crew members and this group of employees do not serve trains available during their trip but running on non-considered routes (such as, e.g. Budapest-Szekesfehervar-Nagykanizsa trains). This assumption is close to reality. The crews originate from three depots: Budapest-Deli, Dunaujvaros and Szekesfehervar. Note that there is no depot in Pusztaszabolcs.

The train paths were scraped from the public timetable (http://www.elvira.hu), based on the query for a particular day, Tuesday, 27 October 2015. The trains running on that day represent a typical workday situation, for which the basic rosters are designed in practice. We assume that each train needs a crew of one person (or at least, a given group of crews compatible with any train), and a given trip between the terminals of the train is not interrupted, they are served by the same crews. This is a simplification, as the required number of crew members depends on the train composition and it is allowed in practice for a crew member to switch to another train on the way if necessary. It would be trivial to overcome these simplifications by introducing the jobs appropriately, but it is not necessary for the demonstration of the method, and we do not have train composition data at hand. So we consider 83 train paths, and thus we have 83 "real" (that is, related to a train) jobs to serve. They all are specified by their starting and ending stations and times.

Since in the Hungarian passenger railway transportation the conductors are scheduled according to periodic rosters, we consider a single day with periodic boundary conditions in time: after midnight we return to the start of the day. Each designed daily shift will be realized each day in the later phases of crew scheduling based on the here designed roster. This enables us to include night duties. All the conditions are evaluated based on this assumption, e.g. a job ending at 23:20 p.m. can be followed by one starting at 0:30 a.m.

The benefit of the present approach is that if once the number of the crews at the depots is fixed, one could define a realistic set of shift frames, and thereby some conditions influencing relations between daily shifts, such as the required time between two shifts could also be taken into account in a realistic situation. On the other hand, the present approach requires to have the shift frames to be defined in advance, so it is not directly capable of optimizing the frames themselves, to contribute to, e. g. the establishment of a recruitment plan. Setting up proper shift frames is a sensitive task requiring rostering experience. An improper setting may result in infeasibility or suboptimal solutions. The shift frames in our present instances are defined in an ad-hoc manner, based on the previous experience in rostering of one of the authors. We consider 32 shift frames, therefore we shall introduce 64 auxiliary jobs. All the frames start and end at the same station. Though the model could support a more general approach, in the considered case the crews need to start and end up at their home depot.

In our consideration, we allow for auxiliary trips serving only the change of the location of the crews. These trips are assumed to be direct: the staff cannot change train during such a trip. The implementation of a proper route planner would enable us to lift this latter assumption, yielding possibly better optima. However, we found that this simplified approach produces bona fide shifts, perfectly satisfactory for our proof of principle demonstration.

As for the determination of the set of jobs feasible in a shift frame, we examined two approaches. The associated instances are denoted as K5 and K6 respectively. In the rigorous setting we assume that the starting and ending auxiliary job of the frame should be accessible (possibly via direct auxiliary trips) from each job of a frame. This results in a relatively lower number of jobs in a frame, which is desirable for the optimization algorithm. Due to our restriction to have direct auxiliary trips, however, the rigorous approach limits the accessibility of the jobs which are remote to the home depot of the given crew. Hence, we also tried a less rigorous approach, in which all the jobs within the time duration of the frame are considered as feasible. This resulted in a larger, but still tractable amount of jobs per frame. On the other hand, in the studied setting, it eventually did not improve the optimum.

Having specified the jobs which can be served in a given frame, one needs to collect the job pairs (p,s) which can follow each other, and calculate their coefficients in the objective function. Between two such adjacent jobs there should be a given technological time depending on the train kind, train composition, etc. As we do not have these exact technological data at hand, we simply assume that there should be at least 20 minutes between the ending time of p and the starting time of s. Of course, when one were to integrate this model to technological software, it could trivially be made accurate. Also, we assume that the ending station of p should be either the starting station of s, or at least this latter should be accessible via auxiliary trips.

As for the objective function coefficients, we simply consider the waiting time between p and s. This includes the non-variable cost due to the required technological time, too, but since it is just an offset, it will not influence the optimum. In a practical situation one should consider a more elaborate cost function taking into account the paid waiting times, auxiliary travels, etc. The present model can support a variety of such considerations flawlessly.

Finally we remark that our simplified model contains some additional simplifications, which would need further considerations. There are short breaks which the crews should be given, and the rules of designing them is rather complex. They could be handled via special jobs with flexible starting and ending times. The tasks before and after a trip should also rather be treated as special tasks. Their data could be easily obtained from actual technological information, thereby turning the present model to a realistic one. We also ignore the minimum required length of a duty. Finally, one should prefer job pairs within a given shift in which the crews commute to and from the terminals of trains. This is done in practice in order to increase the stability of the rosters against non-planned events. This could be modeled by special weighting of the cost function.

Altogether, the instances we elaborated here contain a lot of simplifications but it would be possible to eliminate them using technological data if one were to put the model into real practice. In their present form, they describe a problem of sufficient complexity to illustrate the practical applicability of the method.

5.2 Numerical experiments

The following numerical experiments were performed with the goal to find a suitable informatics tool for crew schedule design. We suggested two solving techniques for this purpose. The first technique called shift construction (SC) is based on the direct solving of the problem (1)-(5) by a common IP-solver. The second technique referred to as shift generation (SG) consists of implementation of column generation process described in Sections 3 and 4 in the optimization environment XPRESS-IVE. The reported implementation of the column generation process searches for the improving sub-problem solution over whole set *K* of frames of the solved instance at each step of the iteration process.

To test the above-mentioned techniques, six instances D1 ... D4 and K5 and K6 were at disposal. Entries of Table 1 report on characteristics of the instances and on the sizes of associated starting linear programming models of the techniques SC and SG. In the table, the symbol noJob denotes the number of jobs contained in an instance, while noK denotes the number of frames. In table sections SC and SG symbol noVar denotes the number of decision variables used in the particular model for the given instance and noCon denotes the associated number of structural constraints.

Parameters of instances and starting sizes of processed models

Table 1

	Instance		SC		SG	
	noJob	noK	noVar	noCon	noVar	noCon
D1	11	28	338	151	28	39
D2	24	30	564	206	30	54
D3	36	46	880	324	46	82
D4	48	62	1192	438	62	110
K5	83	32	6871	874	32	115
K6	83	32	19552	1506	32	115

To solve the problem instances described in the previous Sections, the optimization software FICO Xpress 7.7 (64-bit, release 2014) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 4790 processor with the parameters: 3.6 GHz and 8 GB RAM. The obtained results are reported in Table 2. Following denotation is used in Table 2; Tct denotes the total computational time of the associated technique used for solving a given instance, Tcg denotes the total computational time consumed by solving the sub-problems in one run of the column generation process and Ofv denotes resulting objective function value obtained by the given solving technique, when the given instance is solved. The computational times are given in seconds.

It can be noticed that the technique SC considerably outperformed the SG technique as concerned the computational time, when the SC technique is in several orders of magnitude faster that the SG technique. Furthermore, the SG technique did not reach the optimum when used for solving the instances K5 and K6.

Computational times and resulting objective function values obtained by SC and SG solving techniques.

Table 2

	SC		SG		
	Tct [s]	Ofv	Tct[s]	Tcg[s]	Ofv
D1	0.03	2118	0.81	0.71	2118
D2	0.03	5182	2.55	2.16	5182
D3	0.05	7807	5.93	5.13	7807
D4	0.04	10423	10.41	8.78	10423
K5	0.16	9022	148.08	46.19	9742
K6	0.42	9022	359.25	96.55	11682

Possible causes of the differences can be found in the sparsity of the model matrix on side of the SC technique and in long duration of the iteration process on side of SG technique. Nevertheless, the reported computational times are still acceptable from the point of practical use. That is why; none of the techniques can be excluded from the future research, which should provide the decision supporting tool for the crew schedule design under practical conditions.

Concerning the results obtained for K5 and K6, we can state that the calculated optimal rosters, apart from certain deficiencies due to our simplifying assumptions, resemble the daily duties utilized in practice. The solution points out that certain shift frames are redundant, they do not contain real jobs in the optimal solution. To compare the efficiency of the work within a day, we adopted the following approximation. We consider the time elapsing between the start of the first trip till the end of the last one, plus a time of 30 minutes as the length of the duty. As for productive time, we consider the sum of the duration of the trips in a duty, plus 30 minutes for auxiliary activities. This is clearly an underestimate. Even so, the duty productivity, the typical efficiency of the duties ranges typically from 40% (usually for night duties) up to 85% (very well chosen frames), which agrees with the current practical experience, that is, practical rosters designed manually, or semi-automatically using the algorithm defined in [14]. As we would expect, this suggests that the present algorithm could possibly improve to the possible extent the actual roster designs, at least if the improvement is feasible at all. After the formulation of the practical considerations, the introduced method would be a good candidate in the considered problem as well as many others alike. As it can be made highly parallel, it could be very efficient in practice.

6. Conclusions

We have suggested and implemented two approaches for solving the train crew scheduling problem. We used two sets of benchmarks for evaluation of our approach capabilities. For the first set of benchmarks, the experiments resulted in optimal solutions regardless the used technique. This is probably due to the fact that all jobs (trains) are running on same single-track line, in contrast to the instances from the second set of benchmarks. Both approaches proved to be able to solve large instances of the crew scheduling problems. Even if the SC approach proves to be much faster than the current implementation of the SG technique, the SG technique must not be excluded the future research due to the outperforming SC technique may reach sooner the limits of commercial IP-solvers due to extremal increase of the numbers of decision variables and constraints. Nevertheless, the future research should be focused on improving the SG technique performance. There are several points, which should be considered. The better determination of the starting collection of sub-problem solutions may considerably accelerate the iterative process of the SG technique. An attention should be paid to the process of column generation. The current implementation is based on application of the best-admissible strategy, which is time demanding and which could be replaced by some faster strategy with the same effect. Also the way of the sub-problem solving deserves some research focused on finding a faster algorithm for the cheapest path determination in an acyclic graph.

Acknowledgment

This work was supported by the research grants VEGA 1/0518/15 "Resilient rescue systems with uncertain accessibility of service". We would like to thank to the Research and Development Operational Programme funded by the ERDF for built up the used infrastructure in the frame of project "Centre of excellence for systems and services of intelligent transport II" ITMS 26220120050.

References

[1] BARNHART, C., JOHNSON, E. L., NEMHAUSER,G. L., SAVELSBERGH, M. W. P., VANCE, P. H.: Branch-and-Price: Column generation for Solving Huge Integer programs. J. R. Birge and K. G. Murty (eds.), *Mathematical Programming: State of the Art 1994*, University of Michigan, 186-207, 1994.

[2] DESROSIERS, J., DUMAS, Y., SOLOMON, M. M., SOUMIS, F.: Time Constrained Routing and Scheduling. M. O. Ballet al. (eds.), *Handbooks in OR&MS*, vol. 8, Elsevier Science Publishers, 35-139, 1995.

[3] HOFFMANN, K. L., PADBERG, M. W.: Solving Airline Crew Scheduling Problems by Branch-and-cut. *Management Science*, 39:657-682, 1993.

[4] RIOS, J., ROSS, K.: Massively Parallel Dantzig-Wolfe Decomposition Applied to Traffic Flow Scheduling. *Proc. of the AIAA Guidance, Navigation and Control Conference*, 10-13 August, Chicago Illinois, pp 1-14, 2009.

[5] CAPRARA, A., FISCHETTI, M., TOTH, P., VIGO, D., GUIDA,P. L.: Algorithms for Railway Crew Management. *Mathematical Programming*, 79:125-141, 1997.

[6] KOHL, N.: *Solving the World's Largest Crews Scheduling Problem.* ORbit Extra, Danish Operations Research Society, August 2003.

[7] KROON, L. G., FISCHETTI, M.: Scheduling Drivers and Guards: The Dutch Noord-Oostcase. *Proc. of 34th HICSS conference*, Hawaii, 2000.

[8] WILSON, N. H. M. (ed.): *Computer Aided Transit Scheduling*, Springer-Verlag, 1999.

[9] ABBINK, E., FISCHETTI, M., KROON, L., TIMMER, G., VROMANS, M.: Reinventing Crew Scheduling at Netherlands Railways. *Interfaces*, 35.5:393-401, 2005.

[10] GASS, S. I.: *Linear Programming-Methods and Applications*, McGraw-Hill Book Co.: New York, 1969, Slovak translation Alfa Bratislava, 1972, p. 397.

[11] BORNDORFER, R., GROTSCHEL, M., PFETSCH, M. E.: A Column-generation Approach to Line Planning in Public Transport. *Transportation Science*, 41(1), pp. 123-132, 2007.

[12] JANACEK, J., SIRC, J.: The Column Generation Technique for Public Transport Line Planning by IP-solver. *Communications-Scientific Letters of the University of Zilina*, 12(3A), pp 60-64, 2010.

[13] KONIORCZIK, M., TALAS, B., GEDEON, F.: Preconditioning in the Backtracking Duty Generation of Passenger Rail Crew Scheduling: A Case Study, *Communications - Scientific Letters of the University of Zilina*, 17(2), pp. 23-29, 2015.

[14] WEDELIN, D.: Analgorithm for Largescale 0-1 Integer Programming with Application to Airline Crew Scheduling. *Annals of Operations Research*, 57:283–301, 1995.

Martina Blaskova - Rudolf Blasko - Stanislaw Borkowski - Joanna Rosak-Szyrocka *

SEARCHING CORRELATIONS BETWEEN COMMUNICATION AND MOTIVATION

The aim of the paper is to examine various contents of the relations between communications (including informativeness), and motivation, i.e. consider the benefits of effective communication as well the risks of ineffective communication acting on motivation, perceive feedback as sustainable and accelerative element of any communication system, reflect on communication and informativeness as motivational tools, consider the need to motivate employees for quality communication and raising awareness in the organization and so on. The empirical part presents the most significant results of Slovak-Polish research in motivation regarding the linkages of informativeness, communication and motivation. Conclusion of the paper stresses the need to achieve a state in which effective communication will actually lead to permanent harmonizing all the motivations in the organization, i.e. all kinds of motives (individual, group and organizational) will be mutually harmonized.

Keywords: *Communication, informativeness, motivation, survey, motives behind harmonization.*

1. Introduction

Although people (in organizations) are communicating creatures, the language we use is not simply a means of communicating ideas. People express themselves to each other in symbolic form, and through verbal and non-verbal communication they create and shape relationships [1]. The importance of communication in organization is really great, e.g. from the viewpoint of building creativity [2 and 3], effectivity of education process [4], improving cooperation [5], affecting work behavior [6], increasing level of understandability [7], etc. Effective communication is the lifeblood of any organization [8] and the amount, nature, and uses of communication in the excellent organizations were remarkably different from those of their non-excellent peers [9].

Many studies present there is a big difference between work expectations and needs that employees actually feel and how the managers foresee and interpret these needs of employees. Such a difference in perception may be due to just inoperable, impaired communication or incorrect interpretation of gained responses and requests from the employees and managers. For example, according to Tepper, most managers are convinced that their employees ranked first in their labor-motivational requirements particularly money and job security. However, many employees prefer a good job, recognition, a sense of belonging to the organization, and good working environment [10]. It follows that if organization management is based on misinterpretation of importance of their employees' motivational or other job preferences, i.e. if this one *incorrectly draws conclusions from ongoing communication and sharing information,* this one can adjust managerial mechanisms and activating documents in a false way. Incorrect knowledge bases on motivation and possible ways of motivating could potentially be dangerous, even fatal for the organization.

From mentioned reason, it is important to properly understand the motivation, its complexity, overall structure and, above all, its consequences for the organization. Motivation is the most dynamic, most changing, while most creative and destructive plethora of reasons, intentions, desires, and enhancers that characterize and elicit the behavior of individuals and groups. When defining it, for example, it is possible to focus attention on the positive and negative aspects. In this connection, Barret points out that people feel motivated, i.e. safe, secure, respected and happy when are able to meet their deficiency needs, but they feel anxious or fearful when they are prevented from meeting these needs or when the satisfaction of these needs is under threat [11]. Motivation has to be understood in all its complexity and varietal senses. According to Matuska, motivation is linked with biological standards of human organism, but also with spiritual and social side of human nature [12].

Communication in these experiential processes plays an important role, both in motivating and demotivating way. Based

* [1]**Martina Blaskova,** [1]**Rudolf Blasko,** [2]**Stanislaw Borkowski,** [2]**Joanna Rosak-Szyrocka**
[1]Faculty of Management Science and Informatics, University of Zilina, Slovakia
[2]Faculty of Management, Czestochowa University of Technology, Czestochowa, Poland
E-mail: blaskova@fri.uniza.sk

on this, *paper intent* in the theoretical part is to examine various forms of connecting communication with motivation. The methodological part will present the most significant results of Slovak-Polish research on interdependence of communication (including informativeness) and motivation. Conclusion will stress the need to achieve a state in which the quality of communication and information will actually lead to *permanent aligning all the motivations in the organization*.

2. Communication and informativeness in relation to motivation

Communication participates in the creation and regulation of a subject's behavior in the way it directs the allocation of information because people communicate in order to receive data, or to exchange and modify it and, using such data, they perceive and can persuade thus potentially forming social ties [13]. According to Byers, communication is a symbolic process in which individuals exchange perceptions and ultimately build a knowledge bank for themselves and for others, for the purpose of shaping future actions [7]. It serves as an important channel of transferring and sharing information [14] and aims to understand and reduce uncertainty on both sides of the communication [15].

The relationship between the communication and awareness must be mutual. Employees and managers should communicate with each other in order to be mutually informed on future tasks, objectives, expected risks, results achieved and lessons learned, and able to utilize them even better. Vice-versa, members of the organization should be informed on how formal communication links are operating. It is needed to instill them the principle the informal communication links must always support the organization's vision and mission.

It is important because of, unfortunately, "informal spontaneous communication comes at a cost: interruption. Modern communication technologies have made communication more convenient but have also increased sources of interruption. Empirical research demonstrates the costs associated with interruptions in the workplace," [16]. This leads to knowledge we have to put emphasis on the necessity to remove unwished barriers, and strengthen the *feedback in communication relations* [17]. It means we have to respect the communication is a two-way process – it is about saying something, listening to the response, responding accordingly [6], waiting for and listening to the opinions and new suggestions from the side of communication partner, expressing strong interest in partner and his attitudes, etc.

It is necessary to very carefully pay attention to the *nature of feedback*. Feedback should be *motivating, conducive*. According to Barrett, people feel happiness when they get positive feedback – when they are praised or acknowledged by someone they respect [11]. Feedback must not be hurtful because it could induce undesirable response in terms of decrease of performance, effort, belief, motivation, and willingness to engage in further processes. The unexpected benefit may be the feedback based on the de-encryption of nonverbal communication [18] or decoding meta-communication (hidden) messages [19, 20 and 21], which can detect malfunctioning communication, point to the specific errors in the communication systems of individuals and groups, and thus show that awareness in the organization is incomplete and fails to fulfill the role of necessary platform for the motivation and security.

In order to harmonize interconnection and potential benefits of informativeness, communication and high motivation, it is necessary to consider not only the abundance of quality information for employees but it is necessary to focus attention on the *content and the difficulty of communicating on motivation.* More specifically, the complexity of motivating lies in the fact the employees and managers are *often unable to accurately define and express* what constitutes their work motivation and what exactly they expect from the work – what will bring them the desired satisfaction. Content of primary and secondary motivational categories of each individual is unique and it is always a very special and personal and many individuals consider it to be sensible and do not like to present it to others. In addition, it is not easy to navigate absolutely perfectly in one's own motivation. Bedrnova, Nový et al. made similar comments: "Investigation of motivation is difficult. It is because man himself does not realize his motivation in all its complexity and complex structure and, therefore, often is unable to speak about it," [22].

The importance and seriousness of communication and informativeness in relation to the motivation must be considered also from the *managerial viewpoint*. "Communication serves to motivate; clarify duties, authority, and responsibilities; and permit the expression of feelings," [2]. Mayfield, Mayfield and Sharbrough [23], based on motivating language theory, point out that top leaders can construct and transmit strategic vision of communications for improving organizational performance. Motivating language has to include multilevel and external communication, and should be supported by a cybernetic feedback process.

When considering communication from the viewpoint of an organizational culture, culture encourages the cooperation and effective mutual communication across the organization [24] which strengthens the motivation. Through communication we are comforted and feel empathy from others; but on the other hand, through communication we are hurt, abused and trivialized by others [1]. It means, communication has both positive and negative impacts on the motivation, and thus, it must be carefully performed and mixed.

From the positive viewpoint, open and motivating communication by the management structure of the organization and thorough informativeness might be understood as the *effective motivational tools*. According to Figurska, reciprocal information is an essential factor of motivation. Employees must have truthful information regarding their duties and the content of their role in achieving the objectives and their impact on organizational success. They must therefore obtain information on the organization as well emerging chances, and threats [25]. It is also important to ensure not only enough information but also the sufficient diversity and objective credibility of it. The creative employees should have information derived from primary and secondary sources: books, journals, reports, specifications and norms, rules of procedures, laws and decrees, professional literature, internal reports, etc. [3].

When *connecting informativeness with the self-motivation*, an idea has to be also mentioned: The self-improvement motive leads people to "improve their traits, abilities, skills, health status, or well-being" [26]. This motivates seeking genuine improvement and personal growth. The motive is inferred from behaviors such as actively approaching and coping with problems, seeking information that enables improvement, practicing existing skills, and choosing to work on remedial tasks that reduce deficiencies [27].

However, in order for the communication-informational motivators to be able to discharge their potency in depth, it is necessary to allow employees and managers to be able to *communicate openly,* without fear or embarrassment when their ideas considered will not be perfect, or without fear of sanctions when the information concerned warnings on the inefficiencies in the organization. Frankel points out that comments and suggestions from anyone within and without the organization should be encouraged and given serious consideration no matter from which level or individual they originate and how critical they may be. ... Employee suggestions and involvement must be recognized and rewarded [8, p. 14].

Compared with easier understandable *verbal communicating*, the importance of *nonverbal communication* more and more increases at present. Non-verbal channels such as nodding head, encouraging glance and other gesture serve as an illustration to enhance people's speech or argument [28, 18 and 20]. Employees and managers use such illustrators to emphasize their message more effectively [14]. However, *great vigilance* must prevail in the use of nonverbal communication channels. Positive non-verbal communication (smile communication, closeness, encouraging touch, etc.) can significantly strengthen the motivation. It can enhance the veracity of the claims and overall cultivation of performed communication. On the other hand, negative nonverbal communication (aggressive voice, frown or haughty view, posture full of negative power, etc.) can be very harmful to motivation.

3. Method

Based on mentioned above assumptions, we decided to confirm assumption in our survey that there exists a *significant dependency of level/quality of communication (and awareness/ informativeness) and the level of motivation*. In addition, we assume the comparison of results between Polish and Slovak respondents – residents of the two Slavic countries – could generate more interesting inspirations for the more advanced motivating.

In terms of the survey efficiency undertaken in conditions of two countries, we decided to use the method of a sociological questioning, performed through a technique of questionnaire. In total, 3,682 respondents (1,946 = 52.85% from Slovakia and 1,736 = 47.15% from Poland) with an average age of 37.50 years participated in the survey. There were 1,845 (50.11%) males and 1,837 (49.89%) females. More detailed identifiers of respondents are indicated in Table 1. When processing and evaluating respondents expressions in questionnaires we used our own program, created on the platform of database systems.

Firstly, we investigated whether the respondents considered themselves to be *sufficiently informed on the objectives and problems of their workplace and the organization* and whether they *consider the communication performed towards them as an open and effective*. To evaluate these issues, Likert 5-point scale was chosen where respondents could choose from the following answers: 1 – not; 2 – mostly not; 3 – average; 4 – mostly yes; 5 – yes. To meet the needs of a significant evaluation, we structured these answers into the following three categories: no (level 1 + 2); sometimes

Identification of respondents in the Slovak Republic and Poland (total and percentage) Table 1

	Slovak Republic (1,946 = 100%)				**Poland** (1,736 = 100%)			
Sex/gender	**Males**		**Females**		**Males**		**Females**	
	839 43.11%		1.107 56.89%		1.006 57.96%		730 42.05%	
Age in years	**17-25**	**26-35**	**36-45**	**46-68**	**17-25**	**26-35**	**36-45**	**46-68**
	239 12.28%	585 30.06%	546 28.06%	576 29.60%	249 14.34%	647 37.27%	515 29.67%	325 18.72%
Practice in years	0-5	6-15	16-25	26-50	0-5	6-15	16-25	26-50
	439 22.56%	550 28.26%	515 26.46%	442 22.71%	620 35.71%	649 37.38%	293 16.88%	174 10.02%
Education	**Secondary**	**College**	**PhD.**		**Secondary**	**College**	**PhD.**	
	1.166 59.92%	757 38.90%	23 1.18%		1.061 61.12%	634 36.52%	41 2.36%	

Answers of respondents about their informativeness and communication openness Table 2

Answers scale	All (N = 3,682)			Slovak respondents (N = 1,946)			Polish respondents (N = 1,736)		
	All	Male	Female	All	Male	Female	All	Male	Female
	3,682 100.00%	1,845 100.00%	1,837 100.00%	1,946 100.00%	839 100.00%	1,107 100.00%	1,736 100.00%	1,006 100.00%	730 100.00%
	Level of informativeness								
No [1-2]	7.44%	6.29%	8.60%	8.17%	5.36%	10.30%	6.62%	7.06%	6.03%
Average [3]	18.52%	17.67%	19.38%	19.48%	17.88%	20.69%	17.45%	17.50%	17.40%
Yes [4-5]	74.04%	76.04%	72.02%	72.35%	76.76%	69.02%	75.92%	75.45%	76.58%
Mean	3.94	3.99	3.90	3.89	4.00	3.81	4.00	3.98	4.03
	Openness and efficiency of communication								
No [1-2]	7.17%	6.29%	8.06%	8.07%	6.56%	9.21%	6.16%	6.06%	6.30%
Average [3]	23.09%	22.60%	23.57%	24.05%	22.05%	25.56%	22.00%	23.06%	20.55%
Yes [4-5]	69.74%	71.11%	68.37%	67.88%	71.39%	65.22%	71.83%	70.87%	73.15%
Mean	3.86	3.89	3.83	3.82	3.87	3.78	3.90	3.90	3.91

Answers of respondents about the motivation level to basic areas of work behavior Table 3

Answers scale	All (N = 3,682)			Slovak respondents (N = 1,946)			Polish respondents (N = 1,736)		
	All	Male	Female	All	Male	Female	All	Male	Female
	3,682 100.00%	1,845 100.00%	1,837 100.00%	1,946 100.00%	839 100.00%	1,107 100.00%	1,736 100.00%	1,006 100.00%	730 100.00%
	Level of motivation								
	a) to quality of work								
Low [1-2]	4.10%	4.01%	4.19%	4.32%	4.77%	3.97%	3.86%	3.38%	4.52%
Average [3]	22.32%	22.60%	22.05%	21.17%	19.43%	22.49%	23.62%	25.25%	21.37%
High [4-5]	73.57%	73.39%	73.76%	74.51%	75.80%	73.53%	72.52%	71.37%	74.11%
Mean	3.91	3.92	3.91	3.92	3.95	3.90	3.91	3.89	3.93
	b) to the increase of knowledge and skills								
Low [1-2]	6.49%	6.18%	6.80%	7.55%	7.99%	7.23%	5.30%	4.67%	6.16%
Average [3]	29.11%	29.05%	29.18%	30.06%	27.41%	32.07%	28.05%	30.42%	24.79%
High [4-5]	64.39%	64.77%	64.02%	62.38%	64.60%	60.70%	66.65%	64.91%	69.04%
Mean	3.73	3.73	3.73	3.69	3.72	3.66	3.78	3.75	3.83
	c) to new suggestions								
Low [1-2]	9.70%	9.43%	9.96%	11.25%	11.08%	11.38%	7.95%	8.05%	7.81%
Average [3]	37.24%	35.50%	38.98%	41.78%	37.54%	44.99%	32.14%	33.80%	29.86%
High [4-5]	53.07%	55.07%	51.06%	46.97%	51.37%	43.63%	59.91%	58.15%	62.33%
Mean	3.53	3.56	3.49	3.42	3.51	3.36	3.65	3.61	3.70
	d) to cooperation with superior								
Low [1-2]	10.92%	11.06%	10.78%	11.10%	11.20%	11.02%	10.71%	10.93%	10.41%
Average [3]	36.61%	35.61%	37.62%	37.51%	34.68%	39.66%	35.60%	36.38%	34.52%
High [4-5]	52.47%	53.33%	51.61%	51.39%	54.11%	49.32%	53.69%	52.68%	55.07%
Mean	3.50	3.51	3.49	3.48	3.53	3.44	3.52	3.50	3.56

(level 3); yes (level 4 + 5). Table 2 shows the percentage of relative evaluation of individual response categories, as well as average value (mean) of achieved response level (on a scale 1 - 5). Almost a third of respondents consider their degree of awareness and quality of communication only as average or even insufficient. When examining communication openness, answers of Slovak women were on average lower than the value of Slovak men (in Poland it was almost balanced). When comparing between the two countries, the results of Slovak respondents were slightly worse than Polish. Medians for all groups were 4.00; upper quantiles were 4.00 or 5.00.

In a next question we examined the *level (strength) of motivation*. Respondents were asked to tag the strength of their motivation: a) to quality of work; b) to increase their knowledge and skills; c) to submission of new ideas; d) to cooperate with superiors. We also used a Likert five-level scale, with following specific content of the replies: 1 – low motivation; 2 – rather lower; 3 – average; 4 – rather higher; 5 – high. We subsequently merged responses into these three categories: low (level 1 + 2); average (level 3); high (level 4 + 5). The results in Table 3 show the respondents feel the highest motivation to the quality of their work.

The average values of motivations were again higher in the case of Polish respondents. This implies that, like the area

Results in searching dependency between level of informativeness and level of motivation Table 4

Motivation level/ Information Level	**All** (N = 3,682)			**Slovak respondents** (N = 1,946)			**Polish respondents** (N = 1,736)		
	Yes [4+5]	**Average [3]**	**No [1+2]**	**Yes [4+5]**	**Average [3]**	**No [1+2]**	**Yes [4+5]**	**Average [3]**	**No [1+2]**
	2,726 100.00%	682 100.00%	274 100.00%	1,408 100.00%	379 100.00%	159 100.00%	1,318 100.00%	303 100.00%	115 100.00%
	a) Motivation to quality of work								
Low [1-2]	2.16%	6.30%	17.88%	2.63%	6.60%	13.84%	1.67%	5.94%	23.48%
Average [3]	17.57%	37.24%	32.48%	16.48%	34.56%	30.82%	18.74%	40.59%	34.78%
High [4-5]	80.26%	56.45%	49.64%	80.89%	58.84%	55.35%	79.59%	53.47%	41.74%
High π	77.26%	49.45%	38.64%	76.89%	49.84%	41.35%	64.59%	43.47%	26.74%
Mean	4.05	3.60	3.39	4.04	3.65	3.51	4.05	3.55	3.23
Median	4.00	4.00	3.00	4.00	4.00	4.00	4.00	4.00	3.00
Std deviation	0.75	0.81	1.14	0.75	0.83	1.08	0.75	0.79	1.19
	b) Motivation to the increase of knowledge and skills								
Low [1-2]	3.85%	11.73%	19.71%	4.47%	14.51%	18.24%	3.19%	8.25%	21.74%
Average [3]	25.02%	43.11%	35.04%	26.70%	39.05%	38.36%	23.22%	48.18%	30.43%
High [4-5]	71.13%	45.16%	45.26%	68.82%	46.44%	43.40%	73.60%	43.56%	47.83%
High π	67.13%	38.16%	34.26%	63.82%	37.44%	30.40%	68.60%	33.56%	31.83%
Mean	3.85	3.41	3.35	3.80	3.40	3.33	3.90	3.42	3.37
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.78	0.88	1.14	0.79	0.93	1.06	0.76	0.82	1.23
	c) Motivation to new suggestions								
Low [1-2]	6.31%	17.01%	25.18%	7.17%	20.84%	24.53%	5.39%	12.21%	26.09%
Average [3]	34.19%	48.39%	39.78%	38.85%	51.19%	45.28%	29.21%	44.88%	32.17%
High [4-5]	59.50%	34.60%	35.04%	53.98%	27.97%	30.19%	65.40%	42.90%	41.74%
High π	55.50%	27.60%	25.04%	48.98%	19.97%	19.19%	60.40%	32.90%	26.74%
Mean	3.65	3.20	3.11	3.56	3.06	3.04	3.75	3.37	3.21
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.83	0.89	1.10	0.82	0.87	1.03	0.82	0.87	1.19
	d) Motivation to cooperation with superior								
Low [1-2]	7.45%	17.16%	29.93%	7.46%	16.36%	30.82%	7.44%	18.15%	28.70%
Average [3]	32.72%	49.27%	43.80%	32.53%	51.19%	49.06%	32.93%	46.86%	36.52%
High [4-5]	59.83%	33.58%	26.28%	60.01%	32.45%	20.13%	59.64%	34.98%	34.78%
High π	55.83%	27.58%	17.28%	55.01	26.45%	11.13%	54.64%	25.98%	20.78%
Mean	3.64	3.17	2.90	3.64	3.16	2.77	3.64	3.18	3.08
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.84	0.88	1.07	0.84	0.89	0.96	0.85	0.88	1.19

of information and communication quality, also the area of motivation in Slovakia has more significant reserves than in Poland. In terms of gender differences, Slovak women expressed higher motivation versus Slovak men for all the motivational orientations. This leads to recommendation to much more improve and intensify the motivational efforts against the women in Slovakia. In contrast, the average values of Polish women's motivation were higher than those of Polish men's.

A key concern of our survey was to determine whether the level of information and quality of communication can be considered as *dependent in relation to the strength of motivation*. As appears from Tables 4 and 5, our assumption was confirmed: all the motivations are declining with the decreasing awareness and communication. This means the lower the quality of informativeness and communication, the lower the motivation. If respondents are informed or communication is open, their motivation is approximately 25% higher. This changes neither in

Results in searching dependency between openness of communication and level of motivation — Table 5

Motivation level/ Communication level	**All** (N = 3,682)			**Slovak respondents** (N = 1,946)			**Polish respondents** (N = 1,736)		
	Yes [4+5]	**Average [3]**	**No [1+2]**	**Yes [4+5]**	**Average [3]**	**No [1+2]**	**Yes [4+5]**	**Average [3]**	**No [1+2]**
	2,568 100.00%	850 100.00%	264 100.00%	1,321 100.00%	468 100.00%	157 100.00%	1,247 100.00%	382 100.00%	107 100.00%
	a) Motivation to quality of work								
Low [1-2]	1.48%	4.82%	27.27%	1.97%	6.20%	18.47%	0.96%	3.14%	40.19%
Average [3]	16.74%	36.94%	29.55%	17.18%	30.56%	26.75%	16.28%	44.76%	33.64%
High [4-5]	81.78%	58.24%	43.18%	80.85%	63.25%	54.78%	82.76%	52.09%	26.17%
High π	78.78%	52.24%	32.18%	75.85%	54.25%	40.78%	77.76%	43.09%	14.17%
Mean	4.08	3.65	3.18	4.05	3.72	3.46	4.11	3.55	2.77
Median	4.00	4.00	3.00	4.00	4.00	4.00	4.00	4.00	3.00
Std deviation	0.72	0.80	1.19	0.73	0.85	1.13	0.71	0.73	1.16
	b) Motivation to the increase of knowledge and skills								
Low [1-2]	3.62%	9.76%	23.86%	4.92%	11.97%	16.56%	2.25%	7.07%	34.58%
Average [3]	23.99%	42.82%	34.85%	26.42%	38.46%	35.67%	21.41%	48.17%	33.64%
High [4-5]	72.39%	47.41%	41.29%	68.66%	49.57%	47.77%	76.34%	44.76%	31.78%
High π	68.39%	41.41%	31.29%	63.66%	41.57%	33.77%	71.34%	35.76%	17.78%
Mean	3.88	3.46	3.17	3.80	3.47	3.34	3.96	3.44	2.93
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.78	0.85	1.11	0.81	0.90	1.00	0.74	0.78	1.21
	c) Motivation to new suggestions								
Low [1-2]	5.72%	14.35%	33.33%	6.66%	18.16%	29.30%	4.73%	9.69%	39.25%
Average [3]	32.79%	51.06%	35.98%	39.06%	50.64%	38.22%	26.14%	51.57%	32.71%
High [4-5]	61.49%	34.59%	30.68%	54.28%	31.20%	32.48%	69.13%	38.74%	28.04%
High π	57.49%	28.59%	21.68%	49.28%	24.20%	20.48%	64.13%	29.74%	15.04%
Mean	3.69	3.22	2.90	3.58	3.14	2.96	3.82	3.32	2.81
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.81	0.83	1.18	0.81	0.87	1.12	0.79	0.77	1.26
	d) Motivation to cooperation with superior								
Low [1-2]	5.72%	17.18%	41.29%	5.00%	19.23%	38.22%	6.50%	14.66%	45.79%
Average [3]	31.31%	52.12%	38.26%	32.02%	51.71%	41.40%	30.55%	52.62%	33.64%
High [4-5]	62.97%	30.71%	20.45%	62.98%	29.06%	20.38%	62.95%	32.72%	20.56%
High π	59.97%	24.71%	12.45%	57.98%	22.06%	11.38%	57.95%	24.72%	9.56%
Mean	3.70	3.14	2.67	3.71	3.09	2.67	3.70	3.20	2.66
Median	4.00	3.00	3.00	4.00	3.00	3.00	4.00	3.00	3.00
Std deviation	0.81	0.85	1.13	0.79	0.84	1.10	0.83	0.86	1.17

terms of gender, nor age, nor experience, nor education (although it is true that the higher the education, the higher motivation).

We searched the null hypothesis H0: "The probability, that the informativeness, respectively communication level in a given level of motivation is high, is equal to π" and an alternative hypothesis H1: "The considered probability is less than the π." Test with parameter π of an alternative distribution was chosen for testing, where an unknown probability of π is estimated through using the relative frequency p of the phenomenon occurrence (informativeness/communication is high) – i.e. their number is divided by the number of choice. Tested criterion has a format , where π is the estimated probability, p is the relative frequency of respondents who said their informativeness/communication level is high (i.e. grades 4–5, that is rather higher and high). The value of π (High π in Table 4 and 5) is the lowest just such that level of significance α, at which the hypothesis H0 is not rejected, was $\alpha < 0.05$.

When considering the *communication quality,* the differences are even greater and the average values decrease. Major differences are visible in the case of Polish respondents.

4. Conclusion

As follows from our survey, communication and informativeness have undeniable and extremely strong ties to motivation. Just their accuracy, frequency, clarity and willingness can have a major impact on the quality, orientation, and sustainability of the organization members' motivation. The problem of *sophistication of the communication* arises forcefully: communication must be decent, carried out perfectly and directed positively. It should be focused on inducing positive impressions, voluntary admission of high responsibility and enthusiasm for the joint metes and challenges. An *active togetherness of verbal and nonverbal communication* appears to have possible contribution. Verbal communication puts emphasis, accent, and necessity. Smile communication, identical physical and habit emblems, squeezing a mutual distance, empathy, etc., give depth, a sense of reciprocity, support, dignity, and mutual esteem. Just such communication convinces the motivation.

In addition, the content and possibilities of *motivational program* of organization should be intentionally communicated to all employees and managers. Based on understanding the organizational motivation program, the managers can create appropriate sectional and individualized programs for their employees. Employees must know what their growth opportunities and prospects in organization are, what profits could be earned in return for extra work. They must know they are *highly valued* for their manager.

Based on our survey results, it is appropriate to ***activate interdependence between quality of communication and quality of motivation***. This means communication must also be about the need to *constantly update and dynamize all types of motivation programs* within the organization. The necessity of proactive changes and culture of high performance should be implanted into the behavior of the organization members. This necessity has to be built on a high level of understanding, fellowship, and sharing views on processes and efficiency within the organization. In this situation, it is needed to obtain as much as possible inspirations and suggestions from the employees and managers on how the new motivation program should be aimed, what elements should become the parts of this program, in what way the organizational program should be transformed into the group/sectional and individualized programs, in what way these programs should be evaluated, etc. This implies a huge amount of communicated facts related to the quantitative and qualitative content, events, or elements. It is necessary to *motivate employees and managers to be willing to communicate* in this area and thus provide expected inspirations – communicate not only about the work but also about the opportunities that can improve its course and increase the feeling of joy from work done.

A key aim of such a broad conception of the link between communication and motivation is to achieve the state in which the quality of communication and informativeness will reasonably be resulted in *permanent harmonizing all the motivations in the organization*. Based on the cultivated communication, these motives *must be harmonized* inside the organization: motives of the employees with motives of the managers; motives of the individuals with motives of the group; motives of the groups with motives of the organization; contemporary motives with future motives; egoistic motives with social, altruistic motives; accelerative motives with dampening motives; short-term motives with long-term motives, and so on. In other words, communication should be a means of connection of many motivations of many individuals and groups. Communication should be motivating; communication and its improvement should become the motivation. And, motivation and its contents must be communicated to others so that the individual motivations might become the common motivations.

Acknowledgement

Paper was conducted within the projects: VEGA 1/0890/14 Stochastic Modelling of Decision Making Processes in Motivating Human Potential and VEGA 1/0064/15 Competences Optimization in Correlative Relations with Specificities of Type Positions in Private Security.

References

[1] BRATTON, J.: *Introduction to Work and Organizational Behaviour,* 3rd ed., New York : Palgrave, 2015.
[2] ALDAG, R. J., KUZUHARA, L. W.: *Creating High Performance Teams,* New York : Routledge, 2015.
[3] SOKOL, A.: Management of a Creative Organization. Sokol, A. (ed.). *Managing Diversity in the Organization,* London : Sciencee Publishing, 2015, pp. 9-48.
[4] BLASKOVA, M., BLASKO, R., JANKALOVA, M., JANKAL, R.: Key Personality Competences of University Teacher: Comparison of Requirements Defined by Teachers and/versus Defined by Students. *Procedia - Social and Behavioral Science,* vol. 114, 2014, pp. 466-475. doi: 10.1016/j.sbspro.2013.12.731.
[5] VODAK, J., SOVIAR, J., LENDEL, V.: Identification of the Main Aspects of Cooperation Management and the Problems Arising from Their Misunderstanding. *Communications - Scientific Letters of the University of Zilina,* vol. 16, No. 3, 2014, pp. 93-101.
[6] KING, D., LAWLEY, S.: *Organizational Behaviour,* Oxford University Press, 2013.
[7] BYERS, P. Y.: *Organizational Communication: Theory and Behavior,* Boston : Allyn & Bacon, 1997.
[8] FRANKEL, E. G.: Quality Decision Management - the Heart of Effective Futures-Oriented Management. *Topics in Safety, Risk, Reliability and Quality, 14,* 2008.
[9] PETERS, T. R., WATERMAN, R. H.: *In Search of Excellence: Lessons from America's Best-Run Companies,* New York : Harper and Row, 1982.
[10] TEPPER, B. B.: *Managerial Knowledge and Skills (in Czech),* Praha : Grada, 1996.
[11] BARRETT, R.: *The Values-Driven Organization*. New York: Routledge, 2014.
[12] MATUSKA, E.: *Human Resources Management in a Modern Company,* Gdynia : Wyzsza Szkola Adminsitracji i Biznesu im. Eugeniusza Kwiatkowskiego, 2014.
[13] URBAN, P.: Communication from the Expression Qualities System Perspective. *Communications – Scientific Letters of the University of Zilina,* vol. 16, No 3, 2014, pp. 51-55.
[14] VANATHAS, CH., JAMBULINGAMA, M., MATIAH, J.: Non-Verbal Communication and Satisfaction on the Banking Industry. *International Business Management,* vol. 10, No. 4, 2016, pp. 324-328. doi: 10.3923/ibm.2016.324.328.
[15] HOLA, J.: *Internal Communication in a Company (in Czech),* Brno : Computer, 2006.
[16] DABBISH, L., KRAUT, R. E.: Controlling Interruptions: Awareness Displays and Social Motivation for Coordination. *Proc. of CSCW ´04,* New York, 2004.
[17] OWEN, J.: *How to Manage,* London : Pearson Education Ltd., 2006.
[18] TEGZE, O.: *Non-verbal Communication (in Czech),* Praha : Computer Press, 2003.
[19] MIKULASTIK, M.: *Communication Skills in Practice,* Praha : Grada Publishing, 2003.
[20] BUDA, B.: *Empathy. Psychology of Empathizing and Taking on the Other (in Slovak),* Nove Zamky : Psychoprof, 1994.
[21] TUBBS, S. L., MOSS, S.: *Human Communication,* 6th ed., McGraw-Hill, 1991.
[22] BEDRNOVA, E., NOVY, I. et al.: *Psychology and Sociology of Management (in Czech).* 2nd ed., Praha : Management Press, 2004.
[23] MAYFIELD, J., MAYFIELD, M., SHARBROUGH, W. C.: Strategic Vision and Values in Top Leaders' Communications: Motivating Language at a Higher Level. *Intern. J. of Business Communication,* vol. 52, No. 1, 2015, pp. 97-121.
[24] KACHANAKOVA, A., STACHOVA, K., STACHO, Z.: Organisational Culture Affects All Formalised Activities of Human Resources Management. *Human Resources Management and Ergonomics,* vol. 8, No. 2, 2014, pp. 61-71.
[25] FIGURSKA, I.: *Knowledge Management in an Organization,* Slupsk : HHS, 2012.
[26] SEDIKIDES, C., STRUBE, M. J.: Self-evaluation: To Thine Own Self Be Good, to Thine Own Self Be Sure to Thine Own Self Be True, and to Thine Own Self Be Better. *Advances in Experimental Social Psychology,* vol. 29, 1997, pp. 206-269.
[27] SILVIA, P. J., DUVAL, T. S.: Self-Awareness, Self-Motives, and Self-Motivation. Wright, R. A., Greenberg, J., Brehm, S. S. (eds.). *Motivation and Emotion in Social Contexts: Jack Brehm's Diverse Influence on Contemporary Psychological Thought,* Mahwah, NJ: Erlbaum, 2004. doi: 10.4324/9781410610089.
[28] EKMAN, P.: Emotional and Conversational Non-verbal Signals. Larrazabal, J. M., Miranda, L. (eds.). *Language, Knowledge and Representation*. Springer, 2004, pp. 39-50.

Michal Varmus - Viliam Lendel - Jakub Soviar - Josef Vodak - Milan Kubina *

SPORTS SPONSORING – PART OF CORPORATE STRATEGY

Sponsoring is closely connected with sports and is often incorrectly seen as mere paid advertising. In this article we focused on sponsoring and how it is understood and perceived by Slovak marketing managers and the business community. Based on a survey, interviews and literature analysis we highlighted the strategic aspect of sponsoring in relation to the companies and sports clubs. We focused only on the main factors related to sponsoring. This article also represents a basis for further research in order to better define approaches and motivational aspects of sponsoring.

Keywords: *Sponsoring, management, cooperation, strategy, company, sport.*

1. Introduction

The use of sponsoring as a strategic marketing tool has been considered and evaluated by the business community already for some time. However, perhaps due to the unsuitable legislative environment in Slovakia, sponsoring has been generally seen as mere paid advertising. Even the opinions of international experts differ. Authors Speed & Thompson [1] suggested in one of their reviews that the research on sponsoring [2] has not yielded any theoretical framework which could be used to guide the analysis of customers' reactions to sponsoring. They also extensively discuss the methodology of such research. Hastings [3] specified in his work that sponsoring has two primary promotional goals: 1) to build awareness and 2) to strengthen positive communication about the product and the company. It is now year 2015 and the situation has not changed much. Becker-Olson and Hill [4] pointed out the strategic importance of the relationship between the brand of the sponsor and the sponsored organization. They suggested in their research that the more these two are intertwined, the more positive the reaction of the customers is and vice versa. However, their research was limited to studying a relationship between one sponsor and one sponsored subject. For better illustration, we included in Table 1 an overview of various definitions of sponsorship as used by different authors.

When researching the relationship between a sponsor and a sponsored organization, it is necessary to understand the requirements and criteria that led them to engage in this strategic partnership. Mullin et al. [6] defined the needs of a sponsor company and Robinson [11] defined what a sponsored organization such as a sports club could offer to a sponsor. Comparison of these two opinions is in Table 2.

Definitions of sponsorship Table 1

Author(s)	Definition
Meenaghan [5]	Provision of assistance either financial or in kind to an activity by a commercial organization for the purpose of achieving commercial objectives.
Mullin, Hardy, Sutton [6]	A cash or in-kind fee paid to a property in return for access to the exploitable commercial potential associated with the property.
Simmons, Becker-Olsen [7]	The sponsorship involves an exchange between a sponsor and the event property; the property sells the right to associate with the event to the sponsor, thereby providing leverage opportunity to the sponsor to exploit its communications to consumers.
Cornwell et al. [8]	Sponsorship-linked marketing is a subset of event-related marketing and is distinguished by the presence of sponsorship contracts that authorize certain entities to associate with the event in an official way.
Keller [9]	Sponsorship, through an association with an event, builds brand equity more effectively than traditional marketing communications such as advertising.
Chien et al. [10]	The collection of brand and/or company sponsorships comprising sequential and/or simultaneous involvement with events, activities and individuals (usually in sport, art and charity) utilized to communicate with various audiences.

* **Michal Varmus, Viliam Lendel, Jakub Soviar, Josef Vodak, Milan Kubina**
Department of Management Theories, Faculty of Management Science and Informatics, University of Zilina, Slovakia
E-mail: michal.varmus@fri.uniza.sk

Comparison of needs and offers Table 2

Partner (company) needs by Mullin et al. [6]	**Offers for company** by Robinson [11]
– Brand awareness – Image enhancement – Product trials or sales – Increases sales – Access to potential customers – Entertainment – Employee morale – Return on investment (ROI)	– Facility – Events – Signage – Publication, Public Relations Materials, and the Web Site – Coupons, Special Benefits, and Affinity Programs – Stands or Tables at Events – Contests and Sweepstakes – Use of the Facility – Community Relations Programs – Synergistic Packages

From this we can conclude that the strategic management in the area of sponsoring may also have a cooperative character. Cooperation management is one of the tools that may help to compete on the sports markets [12]. A prerequisite is the creation of a cooperation relationship. Figure 1 represents the elementary steps that are necessary in order to create a cooperative organization. The organizations that exist on the market separately (in a competitive environment) may deal with the potential challenges by cooperation (cooperative environment). This is however a dynamic process. The cooperation continues as long as it is beneficial for the cooperating subjects and as long as it helps to solve their challenges. In case the environment changes and the existing benefits are no longer sufficient, the cooperation changes or is brought to an end.

Belzer [14] suggested that the survival of brands that use sponsoring as their primary tool depends on the ability of the companies to create dynamic and capturing campaigns. Such campaigns enable the brands to become literally part of the subject which is being used as the channel. In addition, it is important to say that the alignment of the values and the messages between the brand and the subject has very high importance.

According to De Pelsmacker et al. [15], the purpose of sponsoring is to connect the sponsored brand or company with the sponsored event. For this reason, the research in this area should focus more on the perceived relationship between the brand image and the event, as seen by the target group. Consequently, the efficiency of sponsoring can be expressed using the following equation of the convincing impact, which is based on several important factors through which the sponsorship influences the level of communication.

Convincing impact = *strength of the connection x duration of the connection x (thankfulness as a result of the connection + felt change as a consequences of the connection)*

The more members of the target group become aware of the connection between the sponsor and the sponsored organization, the longer the duration of this connection will be and the bigger the impact of the sponsorship will be. Similarly, the bigger the thankfulness which is felt as a result of the sponsorship is and the more pronounced the perceived changes are, the bigger the impact of the sponsorship will be, particularly on the image of the sponsor.

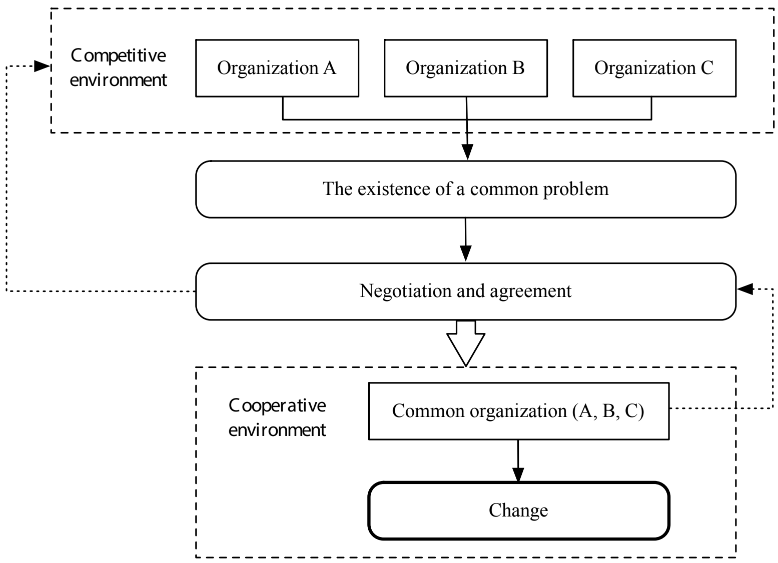


Fig. 1 The process of creating a cooperative organization [13]

2. Research in Slovakia – understanding the role of sponsorship

TNS Slovakia and the magazine Strategy published an exclusive survey of the opinions, attitudes and trends of the marketing managers regarding sponsoring. Results of this survey were published in an article Marketers: top sport for sponsoring is hockey. The survey was conducted among 67 marketing managers and one of the main findings (Fig. 2) was that the managers see sponsoring as something that develops the emotional attachment to a brand or a company.

73 % of marketing managers indicated that the role of sponsoring is to build the awareness and brand image.

The most suitable target sectors for sponsorship as indicated by the marketing managers were sport, culture, art and education (Fig. 3). In sport, the two main sports mentioned were football and ice hockey (Fig. 4).

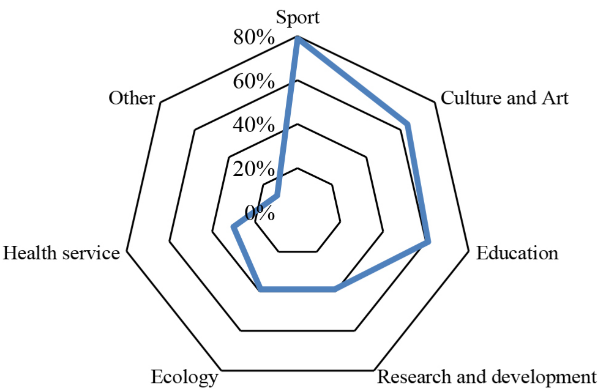


Fig. 3 The best industry for sponsorship [16]

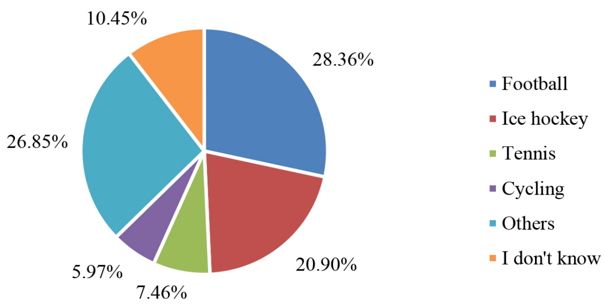


Fig. 4 The best sport for sponsorship [16]

When asked about the position of sponsorship in marketing communication, 40 % of marketing managers indicated that sponsorship is a stable part of marketing communication (Fig. 5). 34 % confirmed that sponsorship should be regarded as an independent marketing tool. Only 3 % of managers indicated that sponsorship is a waste of money.

In Fig. 6 we see that only 26 % (7 % and 19 %) of the surveyed managers indicated that their companies have a department responsible for sponsoring. 45 % of marketing managers confirmed that it is the marketing department which is responsible for sponsoring. 22 % of marketing managers said that they do not plan to establish a department which would focus on sponsoring.

When asked what should be improved in the area of sponsoring, the marketing managers said that it is necessary that the sports clubs or sponsored individuals work and communicate more actively with their fans (e.g. communication via social networks and internet, special events for fans). Among the other areas mentioned were additional publicity of sports events that could improve broader and more relevant reach of the target group, and more specific instructions from the sponsors as well

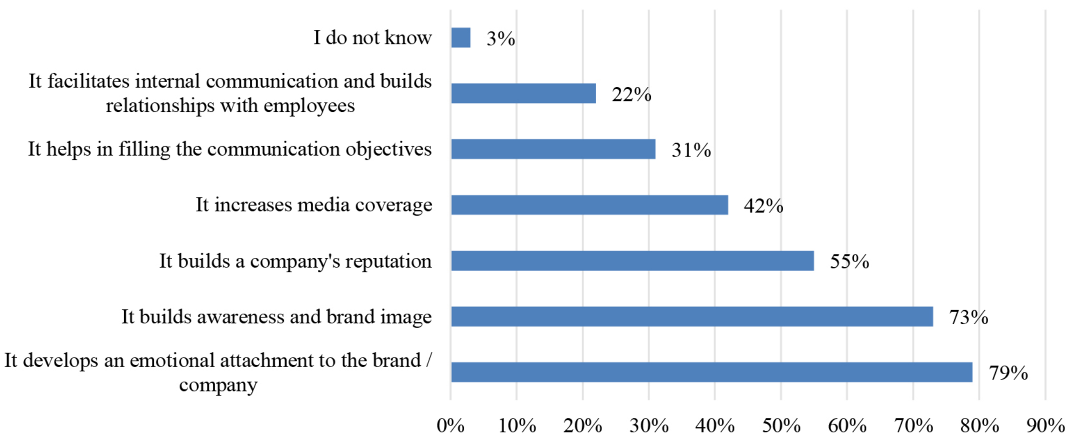


Fig. 2 The role of the sponsorship [16]

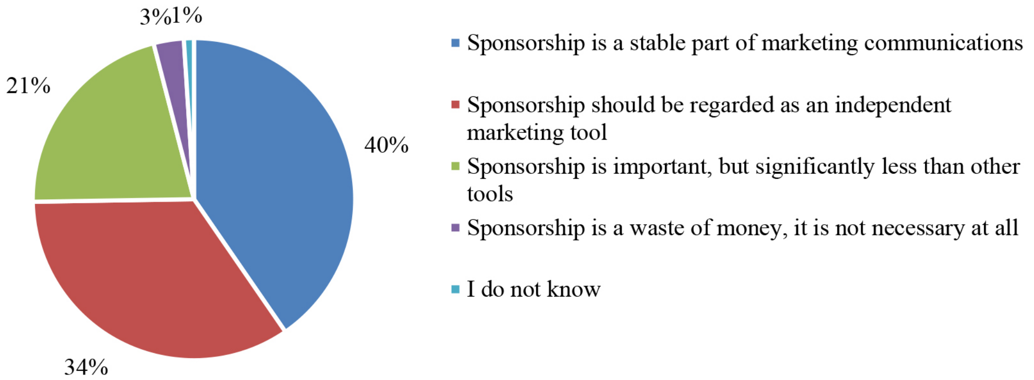


Fig. 5 Position of sponsorship in marketing communication [16]

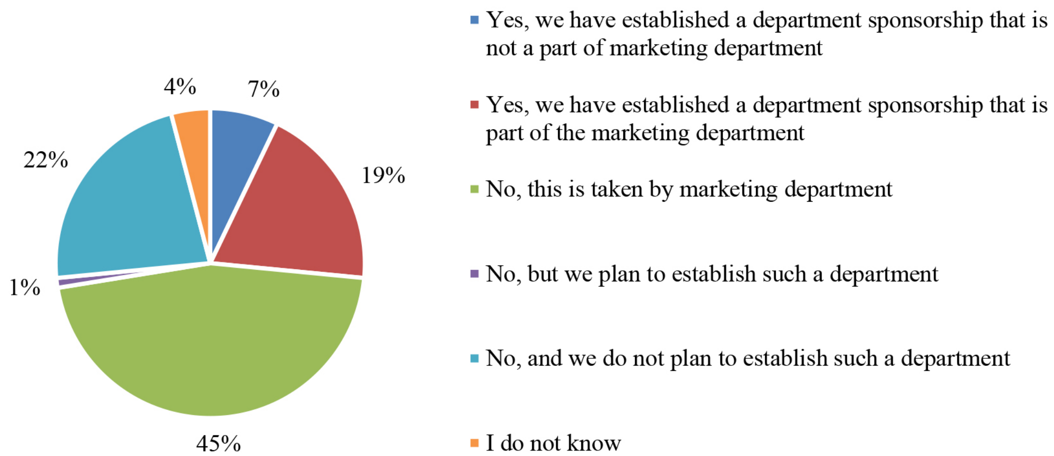


Fig. 6 Management of sponsorship activities [16]

as the sponsored organizations with respect to e.g. their activities, publicity, marketing, budgeting.

3. Discussion

Every potential sponsor faces the following fundamental question: What are the main advantages of sponsoring a certain organization? The answer needs to very clearly specify the main advantages for the sponsor company. As was mentioned in the survey results of TNS, commercial companies expect that sponsoring will help to build an emotional bond to their brand/brands/company, improve brand awareness, brand image, company reputation and publicity. Mullin, Hardy a Sutton [6] see sponsorship as a partnership which is supposed to bring more than just business advantages.

According to the survey as well as other authors [17], [18], [19] and [15], the main goals of sponsorship can be understood as is summarized in Table 3.

Robert Hager [20], senior client partner in TNS, believes that the main reason why the sponsors do not provide money to the clubs could be that the sponsors do not know and understand the detailed concept of the sponsored organization, its added value, publicity coverage and primary audience. In other words, sponsors would like to see real numbers. However, while the numbers are important, not all benefits can be measured in this way. Sponsorship needs to be more than just a support of one event, especially when the company sees sponsorship as part of its marketing strategy.

Thwaites et al. [17] performed a research in Canada where they focused on the question who in the companies is mostly responsible for sponsoring. They found out that the marketing managers play the key role in the sponsorship-related decisions. They also showed that the marketing managers have twice the decision making power in sponsoring compared to the top management. Today and in relation to other research we can say that sponsoring is part of cooperation management. According to Vodak et al. [21] management of cooperation initiatives is closely

Goals of sponsorship Table 3

Communication goals of the company	Public	• Improved awareness of the public about the company • Improved image of the company
	Business relationships and their subjects	• Good reputation • Support of business relationships
	Employees	• Employee motivation • Improved health of the employees
	Key opinion leaders	• Increased attention of the media • PR • Individual goals of the top management
Goals of marketing communication	Building awareness	• Improved brand awareness
	Brand image	• Changed/improved brand image
	Market share	• Increased market share

Offers and requirements of participating parties Table 4

	Offers	Requirements
Company	Money Products Partnership Brand	Branding PR Improved reputation
Sports club	Partnership Branding Events Place for employees Services	Money Partnership

related to the management of the company itself. Cooperation management offers valuable information that represent inputs for strategic decision making. Cooperation management is in the first place a process that includes complete change of the company culture and its values. For this purpose, it is necessary to apply the principles of strategic management.

Based on the research we summarized in Table 4 the main offers and requirements of the companies and sports clubs involved in the sponsoring relationship.

4. Conclusions

Majority of the international authors who study the topic of sponsoring place great emphasis on the cooperation between companies and sports clubs. Based on the research performed in the Slovak business environment we concluded that significant differences exist in the perception of sponsoring. On the other hand, we noticed there is a desire for a change, particularly in the perception of sponsorship, management of the sponsorship-related relationships and in the search for better sponsorship approaches and models. Continuous development of competition and increasing pressure in developing market lead managers to introduce new, progressive forms in management of companies [22]. We concluded that sponsorship as part of company's cooperative strategy may significantly positively contribute to the development of the company and its better image among its employees and the public. According to Blaskova [23], it is necessary for managers, experts and organization employees to continuously reinforce their own motivation levels. This article aims to bring attention to the current situation in Slovakia with respect to sponsoring and the related opinions of the business community. It is clear that the use of sponsoring will be further studied and evaluated as there are many questions that need to be answered.

Acknowledgement

This paper was partially supported by the Slovak scientific grant VEGA 1/0621/14, and VEGA 1/0363/14.

References

[1] SPEED, R., & THOMPSON, P. (2000). Determinants of sports sponsorship response. *J. of the Academy of Marketing Science*, 28(2), 226-238.

[2] COMWELL, T. B., MAIGNAN, I.: An International Review of Sponsorship Research, *J. of Advertising*, 27(1): 1-22, 1998.

[3] HASTINGS, G. B.: Sponsorship Works Differently from Advertising. *Intern. J. of Advertising*, 3(2), 1984, 171-178.

[4] BECKER-OLSEN, K. L, HILL, R. P.: The Impact of Sponsor Fit on Brand Equity. The Case of Nonprofit Service Providers. *J. of Serv Res*, 2006 August; 9(1):73-83.

[5] MEENAGHAN, T.: The Role of Sponsorship in the Marketing Communication Mix. *Intern. J. of Advertising*, 10(1), 1991, 35-48.

[6] MULLIN, B., HARDY, S., SUTTON, W.: Sport Marketing (3rd ed.). Champaign, IL: Human Kinetics, 2007.

[7] SIMMONS, C. J., BECKER-OLSEN, K. L.: Achieving Marketing Objectives through Social Sponsorships. *J. of Marketing*, 70(4), 2006, 154-169.
[8] CORNWELL, T. B., HUMPHREYS, M. S., MAGUIRE, A. M., WEEKS, C. S., TELLEGEN, C. L.: Sponsorship-linked Marketing: The Role of Articulation in Memory. *J. of Consumer Research*, 33, 2006, 312-321.
[9] KELLER, K. L.: Brand Synthesis: The Multidimensionality of Brand Knowledge. *J. of Consumer Research*, 29(4), 2003, 595-600.
[10] CHIEN, P. M., CORNWELL, T. B., PAPPU, R.: Sponsorship Portfolio as Brand Image Creation Strategy. *J. Bus. Res.*, 64(2): 142-9, 2011.
[11] ROBINSON, M. J.: *Sport Club Management*. Champaign: IL Human Kinetics, 2010, ISBN 0-7360-7596-8.
[12] VARMUS, M.: *Management of Cooperative Activities in Sporting Organizations within Alliance Networks. Management of Network Organizations: Theoretical Problems and the Dilemmas in Practice.* Cham: Springer Intern. Publishing, 2015, 215-231, ISBN 978-3-319-17346-7.
[13] SOVIAR, J., LENDEL, V.: *Model of Cooperative Management. Cooperation Management: Effective Approaches to Gain Competitive Advantage.* University of Zilina, 2013, 115-148.
[14] BELZER, J.: The (R)evolution of Sports Sponsorships 22nd July 2015 retrieved from http://www.forbes.com/sites/jasonbelzer/2013/04/22/the-revolution-of-sport-sponsorship/
[15] DE PELSMACKER, P., GEUENS, M., VAN DEN BERG, J.: *Marketing Communications (in Czech)*. Praha: Grada Publishing, a.s., 2003.
[16] Marketers: Top Sports for Sponsoring is Hockey. 20th July 2015 retrieved from http://strategie.hnonline.sk/spravy/marketing/marketeri-top-sportom-pre-sponzoring-je-hokej, 2015.
[17] THWAITES, D., ANGUILAR-MANJARREZ, R., KIDD, C.: Sport Sponsorship Development in Leading Canadian Companies: Issues and Trends. *Intern. J. of Advertising*, 17 (1), 1998, 29-50.
[18] ABRART, R. CLAYTON, B. C., PIRT, L. F.: Cooperate Objectives in Sports Sponsorship. *Intern. J. of Advertising*, 8(2), 1987, 299-311.
[19] HANSEN, F., SCOTWIN, L.: An Experimental Inquiry into Sponsorship: What Effects can be Measured? *Marketing and Research Today,* 23(3), 1995, 173-81.
[20] HAGER, R.: Sports Sponsorship - A Waste of Money or Effective form of Communication? 20th July 2015 retrieved from http://www.tns-global.sk/informacie-pre-vas/blogy/sportovy-sponzoring-vyhodene-peniaze-alebo-ucinna-forma-komunikacie, 2015.
[21] VODAK, J., SOVIAR, J., LENDEL, V.: Identification of the Main Aspects of Cooperation Management and The Problems Arising from Their Misunderstanding, *Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3, 2014. ISSN 1335-4205.
[22] KUCHARCIKOVA, A.: The Importance of Identification and Analysis of Educational Needs for Investment in Human Capital, *Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3, 2014. ISSN 1335-4205.
[23] BLASKOVA, M.: Theoretical Framework of Vega 1/1243/04 Research: Interpersonal Interviews and Motivation. *Communications - Scientific Letters of the University of Zilina*, vol. 7, No. 2, 2005. ISSN 1335-4205.

Viliam Lendel - Stefan Hittmar - Wlodzimierz Sroka - Eva Siantova *

IDENTIFICATION OF THE MAIN ASPECTS OF INNOVATION MANAGEMENT AND THE PROBLEMS ARISING FROM THEIR MISUNDERSTANDING

Goal of the article is to identify main aspects of innovation management using detailed analysis of available scientific literature together with outcomes of performed empirical research. With the identified main aspects, we were able to identify problematic areas related to the use of innovation management and to propose a number of recommendations in order to correct these issues. The article thus brings to managers a handy tool in a form of a set of recommendations (instructions) that are meant to support problem-free use of innovation management within a company.

Keywords: *Innovation management, project management, strategic management, change management, process management, management of human resources.*

1. Introduction

The issue of innovation management is currently highly topical. Innovations are an important tool for increasing competitiveness of companies. Companies do not develop their innovation activities on the basis of "impressions" or "intuition", but on the basis of knowledge obtained from the opinion survey of customers, employees and partners. They collect the necessary information and innovative ideas, reveal innovative opportunities and make decisions about the need to innovate. The aim of the companies is to realize their full potential for innovation. However, in order to be successful, it is necessary to effectively manage these activities and to be able to quickly and flexibly respond to developments in the market. Right there is a place for the identification of major problems and draft of appropriate recommendations to ensure the effective use of market opportunities through innovation.

2. Objective and methodology

The main goal of the article is to collect novel information and knowledge in the area of innovation management, specifically focusing on defining innovation management within the area of management, and to highlight managerial tools and techniques available for use in innovation management in a company. Identification of the main aspects of innovation management could significantly help with the search for problematic areas related to innovation management. These areas require additional attention of the company and represent room for further development and improvement. Subsequent recommendations are meant as practical tools in a toolkit of company managers, ready to be used during implementation of innovation management of the company.

In order to address the points in question, as set by this article, it was necessary to use several methods, depending on and fitting to the character of the individual parts of the solution. In order to accumulate necessary data, we used the method of document analysis (for analysis of current as well as historical data about the topic), a questionnaire method and a method of semi-structured interview (gathering data in an empirical research) and a method of observation (used during visits of selected companies).

For processing the data, we mainly used a method of quantitative evaluation (statistical methods and tools were applied) and a method of comparison (for comparing data gathered by empirical research and data from the analysis of secondary information sources). The following methods were used for approaching and solving the research goals: induction, deduction, synthesis (identification of the main aspects of the innovation management, formulating recommendations for minimal chance of problems), abstraction and model building.

* [1]Viliam Lendel, [1]Stefan Hittmar, [2]Wlodzimierz Sroka, [1]Eva Siantova
[1]Department of Management Theories, Faculty of Management Science and Informatics, University of Zilina, Slovakia
[2]Academy of Business in Dabrowa Gornicza, Faculty of Management, Social Sciences and Informatics, Poland
E-mail: viliam.lendel@fri.uniza.sk

3. The current state of dealing with the issue

Innovation management is a key activity of a company and the synergistic mechanism between technological and non-technological elements, including strategy and culture [1]. In the scientific literature is also a mention about the new paradigm called a total innovation management, which represents the management of the innovation value chain. The innovation value chain dynamically integrates the concept, strategy, technology, business process, culture, and people at all levels of the company. The aim of a total innovation management is to enhance the innovative ability of a company, create value for stakeholders and keep a competitive advantage in a company [2].

Definition of innovation management can be understood as a process of solving problems within the area of innovation in the company [3] or as all activities relating to the effective implementation of new ideas into effective solutions for the market or company [4]. Preble and Hoffman (2012) identified three basic organizational dimensions which affect innovation management: organizational structure, management systems, corporate values and culture [5]. Johannesen et al. (1999) claim that ideas and knowledge are needed to create new products, services and added value, while they emphasize the importance of knowledge and learning of employees [6]. Ortt and Duin (2008) perceive innovation management as an administration and organization of innovation processes [7]. Niosi (1999) pointed out the content development of definition of innovation management at the time [8]. It distinguishes four basic generations of innovation management. The first generation focused on corporate R & D department and management of research activities of the R & D department. The second generation adopts and applies methods and tools of project management. The third generation focuses on processes and strategic management in the form of internal cooperation between different departments of the company. The latest generation represents opening up of the company externally while it includes knowledge of users and competitors in the management of innovation processes.

Birkinshaw and Mol (2009) bring an interesting view of the innovation management. These authors claim that innovation management is the implementation of new business management practices to increase its performance [9]. A similar view is shared by several authors who emphasize the need for regular investing into innovative projects and implementing the necessary organizational changes [10, 11 and 12]. Buchanan and Badham (1999) highlight the specific role of the innovation management, which is the ability to perform rapid and radical changes in the behaviour of the company, if necessary (ensuring the availability of resources, changing work habits ...) [13].

Havlíček (2011) considers the system approach to implementation of the changes as the basis of innovation management [14]. Changes should aim to improve products, processes or position of the company. According to Trommsdorff and Steinhoff (2009) innovation management means integration of functions and areas involved in the innovation process [15]. Turekova and Micieta (2010) perceive innovation management as a comprehensive tool for effective management of innovation processes in the company [16]. Skokan (2004) highlights the growing importance of regional innovation systems, which enable easier sharing of tacit knowledge and increase the capacity for localized learning [17]. Several authors perceived the innovation management as a process in which manager through scientific knowledge and especially practical recommendations, performs basic management activities while using available resources to identify and achieve innovation goals of the company. Innovation management creates a framework for the innovation processes management in order to improve company performance while effectively using innovation potential.

4. Situation in Slovak enterprises – results of the empirical research

Between October 2012 and January 2014 we conducted a research, whose primary goal was to gather and interpret information about the level of the use of innovation management in the environment of Slovak enterprises. In total, 321 managers of small, medium and large enterprises from companies active in the Slovak Republic took part in the research. The calculated recommended sample size was 384 respondents. The survey covered 321 respondents. After the re-analysis, the real sampling error was at the level of 5.46% [18].

In terms of Slovak companies, customers and their identified needs are the most frequently used source of innovative ideas (in 228 companies). The most frequently used sources of innovative ideas also include employees of the company (in 194 companies), analysis of competitive products and services (in 187 companies), the Internet (in 175 companies), exhibitions, conferences, trade shows (in 166 companies), journals and publications (in 159 companies). As a significant source of innovative ideas can also be considered own research activities (115 companies) and the research of the partners of the company (in 98 companies).

Managers of Slovak companies indicate customer needs analysis as a source of the most successful innovative ideas, or ideas with the greatest potential. Customer focus should also be implemented in the management of innovation processes. As resources for successful innovative ideas the respondents considered also their own research activities and their own employees. In this case it is necessary to finance and support their own research activities, create a favorable environment for their workforce and care for their educational development.

The importance of customers confirmed the fact that only 257 companies (80.06%) create innovative ideas based on the input from customers and their requirements. Another stimulus for generating innovative ideas is finding a specific problem (186 companies, 57.94%). This means that these companies prefer the opposite strategy, it means they do not apply a proactive approach, companies create ideas only when a problem arises. Other less used incentives include business needs (81 companies, 25.23%) and new technologies (72 companies, 22.43%). As a positive can be considered that only 9.03% of companies generate the innovative ideas at random and only 5.92% of Slovak companies do not generate them at all.

Companies in Slovakia do not use an information system to work with innovative ideas (recording, sorting, distribution...). This was confirmed by 61.28% of managers surveyed. Only 28.96% of respondents have their innovation process supported by IT solutions, while the most used information system is a system provided by the company Salesforce.com (16.72% companies). Other companies use their own IT solution. In the survey 9.76% of the respondents were not able to express their opinion about the use of the information system to work with innovative ideas.

As the main criteria for deciding on further elaboration of innovative ideas into innovative opportunities, the managers of Slovak companies identified the availability of funds (it influences up to 25.81% of decision-making process).Other important decision criteria are the reality of demand, technological options, available knowledge in the issue and the availability of human resources. In terms of priorities for decision-making time and the physical space in the development phase are less involved. The mostly used methods in the deployment phase of innovation are the techniques of development of the creativity (42.99%). The quite frequently used methods are conceptual methodological tools (24.61%), forecasting methods (19.63%) and pragmatic methodological tools (19.31%). Techniques of knowledge management (14.64%) and innovative graphs (7.79%) are used on a small scale as well.

As major problems hindering the effective innovation management by the managers are considered: the lack of the necessary financial resources to ensure the innovation process (210 companies), distrust of the company managers to the possible outcomes arising from a lack of innovation (187 companies) and missing information ensuring the innovation process (168 companies).

5. Identification of the main aspects of innovation management

Management represents the main element on the connecting line between a company and its customers. In order to work, management requires information from the IT system, mainly information accumulated by the departments involved in the innovation activities. A mutual connection exists between management and innovation management. Innovation management offers immediate overview of the overall development of the key areas of the business as well as a perspective on individual innovation activities performed by employees. Management manifests itself in the innovation management mainly through the following elements (Fig. 1):

- strategic management,
- change management,
- project management,
- process management,
- management of human resources.

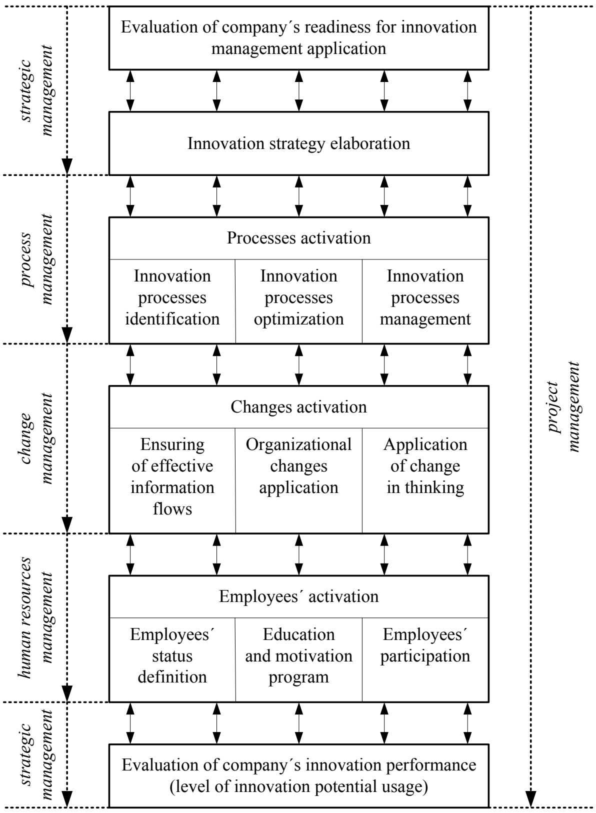


Fig. 1 Identification of the main aspects of innovation management [19]

Strategic management and innovation management

Strategic management is used in the area of innovation management mainly for formulating innovation strategy which is preceded by a detailed analysis of current situation within the company. Innovation strategy is innovative direction of business approach to the choice of objectives, methods and ways to fully utilize and develop the innovative potential of the business [20]. Kovac (2007) sees innovation strategy as determining long-term fundamental business goals and determines the activities and resources for achieving these goals [21]. Orientation of goals is focused on timely response to changes in signalling of need of innovations. According to Malek Alekakhlagh et al. (2013), innovation strategy must be based on knowledge and facts, supplemented with learning and creativity to cope with the concepts of change and the ability to expand an institution's creative capacity [22]. Zgrzywa-Ziemak & Walecka-Jankowska (2013) show that the location of the innovation strategy in the corporate strategy initiates the innovation process from the determination and configuration of future product parameters based on the product and organizational environment analysis [23].

Change management and innovation management

During the application of innovation management in a company, the company strategy is revised and modified so that it reflects the plans of top management regarding management of innovation activities. However, such a change can end up influencing the roles of multiple employees. Depending on the character and number of realized innovation projects, it is necessary to revise the currently used company organizational structure and to adapt it to the current situation.

Given great variability of innovation projects, it is possible to use multiple types of organizational structures for their organization. The general rule is that the organizational structure adapts to the innovation project (content, complexity, extent, time needs) and not vice versa.

Organization remains of key importance in the process of managing innovation activities, especially in today's turbulent environment. Innovation management aims to ensure competitiveness of the company in such environment. However, for this to happen it is needed that the company is capable to dynamically react to the arising changes.

Here is room for using dynamic innovation organization structures that offer immediate reaction and consequent change in configuration of employees and processes, as necessary. Dynamic innovation organization structures (champions, purpose teams, project teams, project centers...) are characterized by the following properties:

- ability to rapidly react to changes,
- decentralized management,
- use of the creative approach,
- flexibility in content and activities of the groups and individuals,
- acceptance of higher degree of uncertainty and risk in management,
- direct evaluation and testing of new ideas,
- focus on results,
- adequate number of management levels,
- administratively undemanding methods of management,
- high added value,
- informal team work,
- lower number of organizational elements and connections,
- lower requirements on the management system.

Project management and innovation management

Planning in the innovation management represents a significant group of activities which are used to set innovation goals and by which resources and ways for achieving them are determined. Complexity and demanding character of this process increases with size of the company, with the increasing hierarchical level on which it is performed, with the length of the time horizon and the number of involved parties (partners). Planning of innovation activities in a company requires:

- to anticipate future development of external and internal environment and the changes that occur in them (development of customers, change in segments, new communication tools, development of customers' demands, development of factors that influence customers' purchasing decisions...),
- to take into consideration interests of the various involved parties who take part in the innovation processes (employees, top management, suppliers, banks, partners, surroundings); interests of individuals, groups and society,
- to consider economic as well as social conditions and their criteria,
- to arrange in hierarchy goals and tasks of innovation processes, which create conditions for internal harmony of relationships and processes and occurrence of synergistic effects; also to identify relationships and processes related to future innovation and to arrange them in hierarchy,
- to consider limitations of resources, their suitable allocation and efficient use for supporting innovation activities,
- to choose suitable methods and techniques that enable creation of innovation ideas and their evaluation and selection.

Process management and innovation management

Basis for functioning of any company are its processes. Process management is established on the principle of activity integration into comprehensive processes [24]. The process approach is based on the assumption that the causes of unacceptable economic results are poorly set innovation processes. For this reason it is necessary to make all innovation processes more efficient and to eliminate those that do not bring the desired outcomes.

The main goal of process management in relation to innovation management in a company is the effectiveness of innovation processes. Process management takes a comprehensive view of all company activities and integrates them into individual processes. Information exists and is available for each activity, e.g. who is realizing the activity, how is the realization going, in what way, what are the limitations for realization, what is the input for the activity, what are the outputs, relationship of the internal customers, spread of overhead costs to individual activities within the process. Every activity has a defined metric of performance. Another strong point for process management is implementation of any changes resulting from realization of innovation initiatives in the process of managing a company.

Key requirement for successful use of innovation management in a company is thorough mapping of current company processes. Above all, the attention should be namely paid to innovation processes. It is necessary that they are identified and consequently optimized.

Management of human resources and innovation management

A key role in innovation management is definitely played by people, the employees of the company. The success of innovation initiatives will depend on their performance and attitude. People and their motivation, knowledge, skills, capabilities, creativity and flexibility become the most important strategic resource for successful realization of innovation activities in a company. Employees prepare an analysis of the situation, together with the management they set innovation goals, formulate innovation strategy, action plans as well as a system to control effectiveness and efficiency of innovation activities. Education and training representing an important part of the activities of personnel management of the company contribute to the development of the necessary competencies of company staff [25]. By using a suitable motivation program, ensuring possibility for further education and establishing suitable work environment the company can achieve smooth implementation and consequential usage of the innovation management and eventually successful realization of innovation activities.

6. Discussion

The use of innovation management in a company is a complex process that requires thorough understanding of the company environment. Causes of failure in this process can be multiple and may have a different character. For example, there may be a lack of innovation expertise, failure to secure information flow in the company, lack of education and motivation of employees and so on. Reasons of failure in the management of innovation activities are caused by several actors. Firstly, they are managers, in the case of lack of support to innovative activities in the company, employees of the company, in the case of passive participation in the innovation process and customers, in the case of their indifference towards receiving the value added. In the next part we identify possible risks together with recommendations that are meant to help to reduce these risks. A necessary requirement for successful functioning of innovation management in a company is also prevention.

As a problem in ensuring the information security in innovation processes arising in this field can be considered the inefficient information flows in the company. This is the reason of misunderstandings and employees cannot carry out their innovative tasks entirely. It is recommended to the managers to ensure the efficient work with information related to innovation in the company. This means that it is necessary to ensure the access to information to all interested parts in the innovation process, and it is necessary to collect all valuable information in one database and apply the principles of effective work with information.

Another problem is the lack of information systems to work with innovative ideas, opportunities and innovations. In many cases there is no evidence of implemented innovations and innovative ideas generated in the company. This often means that the potential innovative opportunities are left unused. To solve this problem, managers should make a comprehensive record of innovative ideas and innovations in the company. Every innovative idea in the company has to be recorded and assessed. Of the same importance is the record of the currently unusable innovative ideas which have potential value for the future.

In the area of ensuring the innovation expertise an insufficient usage of creative thinking in the development of ideas to solve the problem can be seen. The result is a low number of ideas that do not allow to make a decision to resolve the problem. It can be recommended to the managers to develop human creative potential in the company. It is necessary to create conditions for application of creative thinking selecting appropriate exercises to develop creative skills and abilities of employees. It would also be appropriate to create innovative teams, including representatives of the young and old. Innovative teams should include representatives of several departments of the company (production, sales, logistics, trade, service).

Insufficient implementation of knowledge management belongs to the common problems in the field of innovation management. On the one hand it is reflected as a lack of knowledge of the employees, on the other hand, much of the knowledge created in the innovation process is forgotten or lost. The result is unnecessary, repeated creation of new knowledge which is already created in the innovation process. In the first case it can be recommended to the managers to ensure the management training courses designed to supplement the necessary knowledge to employees. In the latter case it is necessary to provide recording and archiving of acquired knowledge in the innovation process through the implementation of appropriate IT solutions.

Problem in the application of management elements in the innovation process is the lack of a coherent methodology for the management of innovation in the company. The managers can often exchange the innovation process by using a simple type of creative brainstorming techniques. In many cases, they are in time and work stress and they do not pay attention to the support of innovation and integrating innovations into long-term strategic plans of the company. The managers may be encouraged to pay more attention to innovative activities of the company, turn them into long-term business objectives and incorporate them into innovative business strategy.

Another problem is the lack of development of the innovation program. In many cases, managers do not have sufficient information about available resources and means when they plan innovative activities. The result is the increased probability of failure of implementation of the innovation project. It is recommended to overcome this probability, it is necessary to carry out a detailed analysis of the current state of innovation potential and the application of methods and techniques of project management.

Another problem is the absence of remuneration for innovative ideas and appropriate motivation program. The result is the passivity of employees who are not motivated to bring new innovative ideas and get involved in innovative task solutions beyond their tasks and responsibilities. As a solution, the managers should establish a fair system of remuneration for innovative ideas. It is necessary to develop appropriate motivation program, which involves employees in innovation activities of the company. Employees will be informed of the expected changes and motivation program will encourage open communication within the company. One of the most important factors of employees´ motivation is to show an interest in them, employees need to feel valued and important for the company [26].

Failure of the management of innovation processes can be caused by unsuitable organizational structure, which does not allow open communication among the stakeholders and does not support new innovations. It is a role of the managers to rethink the current organizational structure, create and implement flexible organizational structure that will have the ability to respond to changes in business environment and allow fast exchange of information, organize meetings and activities of innovative teams.

A common problem in the measurement of innovation performance is non-evaluation of effects and benefits of implemented innovations. The result is that the company has no feedback on the adoption of innovations by the customers, cannot measure their performance and take measures leading to continuous improvement of the management of the innovation process. We suggest that the managers should create an evaluation system focusing on the diagnosis of the results and contributions created and on the market launched new products. Based on the results it is also recommended to formulate measures to improve the management of the innovation process.

A common problem is that the first and the second phase of the innovation process are carried out insufficiently. In many cases it can be observed that the development of the first innovative idea pops up without detailed analysis. The results are the frequent changes and increased costs of implemented innovative projects. We propose that the managers should carry out a detailed analysis of all the innovative ideas and suggestions from internal and external environment. It is necessary to pay attention to the first two phases of the innovation process because they affect its success. Only the identification of valuable innovative ideas can bring successful innovation.

7. Conclusions

The management of innovation activities cannot be carried out only intuitively based on the development of the situation. Company managers should be aware that the implementation of innovation management also brings certain risks. Otherwise this initiative would be destined to fail. In order to succeed in this area, it is a key to identify risk areas early and to take corresponding measures in order to increase the probability of the success of implementing innovation management in a company.

Acknowledgement

This paper was partially supported by the Slovak scientific grant VEGA 1/0363/14 Innovation management – processes, strategy and performance.

References

[1] ZHIRONG, Y. et al.: *Total Innovation Management: A New Emerging Paradigm of Innovation Management*: Engineering Management Conference. IEMC ,03. Managing Technologically Driven Organizations: The Human Side of Innovation and Change, 2003, ISBN 0-7803-8150-5.

[2] XU, Q. R. et al.: Total Innovation Management: A Novelparadigm of Innovation Management in the 21st century, *J. of Technology Transfer*, vol. 32, No. 1-2, 2007, 9-25.

[3] ROGERS, E. M.: *Diffusion of Innovations*. 4th ed. NY : Free Press, 1995, ISBN 0-02-926671-8.

[4] DREJER, A.: Towards a Model for Contingency of Management of Technology. *Technovation*, vol. 22, No. 6, 2002, 363-70.

[5] PREBLE, J. F., HOFFMAN, R. C.: General Management of Innovation: Lessons from the Shaker Community, *J. of Management History*, vol. 18, No. 1, 2012, 24-45, ISSN 1751-1348.

[6] JOHANNESSEN, J. A. et al.: Managing and Organizing Innovation in the Knowledge Economy, *European J. of Innovation Management,* vol. 2, No. 3, 1999, 116-128, ISSN 1460-1060.

[7] ORTT, J. R., VAN DER DUIN, P. A.: The Evolution of Innovation Management Towards Contextual Innovation, *European J. of Innovation Management*, vol. 11, No. 4, 2008, ISSN 1460-1060.
[8] NIOSI, J.: Fourth-generation R&D: From Linear Models to Flexible Innovation, *J. of Business Research,* vol. 45, No. 2, 1999, 111-7.
[9] BIRKINSHAW, J., MOL, M. J.: The Sources of Management Innovation: When Firms Introduce New Management Practices, *J. of Business Research*, vol. 62, No. 12, 2009, 1269-1280.
[10] BURNES, B.: *Managing Change: A Strategic Approach to Organizational Dynamics*. Prentice Hall, 2004, ISBN 978-0-273-71174-2.
[11] CALDWELL, R.: Models of Change Agency: A Fourfold Classification, *British J. of Management,* vol. 14, No. 2, 2003, 131-142.
[12] STENSAKER, I., LANGLEY, A.: Change Management Choices and Change Trajectories in a Multidivisional Firm, *British Journal of Management*, vol. 21, No. 1, 2010, 7-27.
[13] BUCHANAN, D., BADHAM, R.: *Power, Politics, and Organisational Change: Winning the Turf Game.* SAGE, 1999, ISBN 978-1-4129-2834-2.
[14] HAVLICEK, K.: *Management and Controlling of Small and Medium Businesses (in Czech),* Praha: Eupress, 2011, ISBN 9788074080562.
[15] TROMMSDORFF, V., STEINHOFF, F.: *Marketing Innovation (in Czech)*, Praha: C. H. Beck, 2009. ISBN 978-80-7400-092-8.
[16] TUREKOVA, H., MICIETA, B.: *Innovation Management - The Promotion of Innovation (in Slovak),* University of Zilina: EDIS, 2010, ISBN 978-80-89478-02-6.
[17] SKOKAN, K.: *Competitiveness, innovation and clusters in regional development (in Czech)*, Ostrava: Repronis, 2004. ISBN 80-7329-059-6.
[18] LENDEL, V., HITTMAR, S., SIANTOVA, E.: *Identification of the Main Problems in the Management of Innovation Processes and the Draft of Appropriate Recommendations,* Knowledge Management in Organizations: 10th intern. conference, KMO 2015: Maribor, Cham: Springer, 2015, 221-232, ISBN 978-3-319-21008-7.
[19] LENDEL, V., VARMUS, M.: Creation and Implementation of the Innovation Strategy in the Enterprise, *Economics & Management*, No. 16, 2011, 819-825, ISSN 1822-6515.
[20] LENDEL, V.: *Management of Innovation Processes in Company (in Slovak).* Habilitation thesis, University of Zilina, 2014, 130 p.
[21] KOVAC, M.: *Creation and Innovation Management (in Slovak),* Kosice: Technicka univerzita v Kosiciach, 2007, 121 p. ISBN 80-8073-690-1.
[22] MALEK AKHLAGH, E. et al.: Innovation Strategies, Performance Diversity and Development: An Empirical Analysis in Iran Construction and Housing Industry. *Iranian J. Management Studies,* vol. 6, No. 1, 2013, 31-60, ISSN 2008-7055.
[23] ZGRZYWA-ZIEMAK, A., WALECKA-JANKOWSKA, K.: Strategy and Innovation Relationship, *Przedsiebiorstwo przyszlosci,* No. 1, 2013, 45-57, ISSN 2080-8461.
[24] VODAK, J., SOVIAR, J., LENDEL, V.: Identification of the Main Aspects of Cooperation Management and the Problems Arising from their Misunderstanding, *Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3, 2014, 93-101, ISSN 1335-4205.
[25] KUCHARCIKOVA, A.: The Importance of Identification and Analysis of Educational Needs for Investment in Human Capital, *Communications - Scientific Letters of the University of Zilina,* vol. 16, No. 3, 86-92, ISSN 1335-4205.
[26] KAMPF, R., HITKA, M., POTKANY, M.: Interannual Differences in Employee Motivation in Manufacturing Enterprises in Slovakia, *Communications - Scientific Letters of the University of Zilina,* vol. 16, No. 4, 98-102, ISSN 1335-4205.

Alzbeta Kucharcikova - Lubica Konusikova - Emese Tokarcikova *

APPROACHES TO THE QUANTIFICATION OF THE HUMAN CAPITAL EFFICIENCY IN ENTERPRISES

Human capital is one of the enterprise inputs. There are people who are the bearers of new knowledge, ideas, experiences and skills; they contribute to their own personal growth as well as to increasing the performance and competitiveness of the whole enterprise. Human capital management helps enterprises to achieve enterprise goals just by using the human capital efficiently. For the assessment of the results achieved and directing further development it is important to know how to measure the efficiency. The aim of the article is: based on the study of special and scientific foreign literature to identify and compare the different theoretical approaches to the measurement of the efficient use of human capital for possible implementation in enterprise practice.

Keywords: *Human capital, value, efficiency, measurement, management.*

1. Introduction

Organizational performance refers to an organization's results, including operating results (productivity, quality, efficiency, etc.), market results (sales, market share, customer satisfaction, etc.) and financial results (costs, revenues, profits, etc.). There is considerable evidence that achieving expected organizational performance is determined by achieving desired employee performance that refers to employees' working results and behaviour, determined by employees' abilities (knowledge and skills to perform agreed work) and motivation (willingness to perform agreed work), which enable an organization to achieve expected goals [1].

The prerequisite for increasing the performance and competitiveness of the enterprise is, in particular, the efficient use of production inputs. The efficiency is a qualitative criterion, which involves such ways of the use of production factors, which compare the inputs and the effects achieved. Economic theory introduces, and there are united methodologies for the quantification and assessment of the efficiency of non-current assets, current assets and work in enterprise practice. However, this does not apply to human capital.

Human capital is relatively a young production factor, and as it is closely related, and in some respects, it has similar features as a production factor of work, economic analyses and research works do not often make an explicit distinction among them. However, a different angle on the working activity of a man and currently increasing appreciation of the importance of quality human capital also require new, different types of indicators, which would express better the impact and mainly the benefits of the variable [2].

Qualitative nature of the human capital is difficult and many times does not allow easy quantification, an attempt to derive the partial relations and contexts, the modification of existing indicators, identification of new parameters as well as the proposal for a new methodology of new statistical surveys can contribute to a better understanding of the total impact of human capital on the company and the enterprise as well.

The aim of the article is to identify and compare the different theoretical approaches to the measurement of the effective use of human capital for possible implementation in enterprise practice.

2. Human capital and Efficiency

Efficiency is generally defined as the relationship between the outputs achieved and the inputs used. The efficiency of human capital can be calculated analogically as a percentage of the value of enterprise output and the value of the input (the human capital). However, when choosing output and input, there must be chosen such variables, which relate to each other and are crucial to the administration of the enterprise performance [3].

* [1]Alzbeta Kucharcikova, [2]Lubica Konusikova, [1]Emese Tokarcikova
[1]Department of Macro and Microeconomics, Faculty of Management Science and Informatics, University of Zilina, Slovakia
[2]Silesian university in Opava, School of Business Administration in Karvina, Czech Republic
E-mail: Alzbeta.Kucharcikova@fri.uniza.sk

Therefore, the efficiency represents the ability of the work value to create a product. Human capital is effective only if the product is greater than its value. It is being understood that the higher is the difference between the value and the product, the higher is the efficiency. **From a macroeconomic point of view, the efficiency of human capital** can be expressed as a percentage of the lifelong product HC (Y_{LK}) and its value (H_{LK}).

$$E_{LK} = \frac{Y_{LK}}{H_{LK}} \quad (1)$$

When starting working, human capital has a certain value (e.g., school graduate), which, however, constantly improves and its value rises in the context of the working process. This is due to the acquisition of new skills, knowledge, and habits. The indicator of the value of human capital is therefore not stable, but on the other hand, a dynamic variable, which constantly develops and changes by the influence of circumstances [4].

The value of human capital differs in each individual, as it is developed by the age and conditions the bearer has available. It is gradually increased from the birth until the old age, when it begins to decline. The value of human capital is increased, particularly through new experiences and knowledge acquired in education or other forms of development of human capabilities and creativity [5].

Calculating and monitoring the efficient use of human capital, it is therefore necessary to quantify the output produced by the human capital and the size / value of the human capital.

The output, when measuring the efficient of human capital, may be the degree of fulfilment of the enterprise goal (e.g.: increasing the customer satisfaction, increasing sales, reducing the error rate, shortening the time dealing with complaints, etc.), the quantification of the enterprise savings, the number of registered patents and inventions per one employee or worker of the development department, the volume of performance per one employee, etc.

Considering the measurement of the size of human capital, there are several known approaches, but no single methodology, as in other enterprise inputs, has been accepted so far. The problem is the quantification of the different components of human capital, i.e. knowledge, abilities, skills, motivation, talent, etc.

According to Mazouch and Fischer [6] there are three factors acting on human capital and its value. The initial factors are genetically inherited, inborn abilities of an individual. These can subsequently be developed by further activity of the environment and education. As a consequence of family, social, other factors of environment and education there are acquired and developed skills and knowledge. An important discovery that we must not forget is that the various components of human capital will affect each other. As the individual properties may affect the acquirement of the knowledge or the development of skills, the development of properties and skills depends largely on the environment in which the individual is born.

3. Indicators of the efficient use of human capital

The use of human capital metrics is an important part of the human capital management (HCM). If management knows how to properly determine the value of human capital it has available in the enterprise, it is important to find whether it is used effectively. If not, it is essential that enterprise management identifies problems and sets measures, which would ultimately lead to the improvement and enhancement of individual enterprise processes.

Quantifying the size of the value of human capital and its efficiency is very difficult, but in the selection and construction of whether direct or indirect (auxiliary) indicators in this area, it is important that the metrics should be linked to the key performance indicators of the organisation.

Conceptual model

Conceptual model points out to the interconnection of human capital and the overall performance of the organisation (Fig. 1). This approach puts the emphasis on the strengthening of human capital, which leads to the increase of performance of individual employees and ultimately it has a positive effect on the overall improvement of organisation performance.

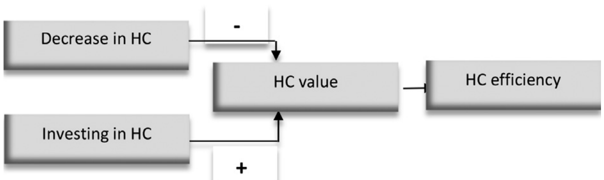


Fig. 1 Conceptual model for measuring the HC efficiency
Source: [7]

The structure of conceptual model involves a link of relations among the investments in human capital, its decrease, value, total efficiency and their measurement. The results of the measurements referred to four key areas in the enterprise can then be compared with other companies (benchmarking) [7].

Investment in human capital represents expenses the organisation invests to ensure training, education, development of knowledge and skills of its workers. The investment, together with the decrease, affects the value of human capital in the enterprise, and consequently its efficiency. Investments have a positive effect and contribute to increasing the human capital efficiency and, ultimately, lead to increasing the overall performance of the enterprise. It may have financial, but also non-financial forms (Fig. 2). *Financial performance* includes productivity, market share, enterprise profitability, while *non-financial performance* includes customer satisfaction, innovation, and improvement of workflow and the development of workers´ skills [8].

Fig. 2 A link of conceptual model with investments in HC, HC efficiency and performance of the organisation, (source: [8])

Human capital efficiency represents a dependent component within the conceptual model. Income factor is the basic measure of human capital efficiency that is calculated as a proportion of total income and total number of employees in the enterprise, while the factor of costs, profit and total return on investment to HC can be used as the other measures.

Decrease in human capital that may be caused by the departure of workers, has a negative impact on the HCM. The decrease of workers can be given, for example, by the percentage of individuals who have left the enterprise voluntarily, or the percentage of involuntary loss of HC (Fig. 1). For the calculation of the value of human capital, an enterprise may use measurements expressing the factor of income compensated, i.e. what percentage of the income from the sale goes to employees, or the factor of costs compensated, i.e. what percentage of operating expenses goes to employees [7].

Human Capital Efficiency - HCE

According to Andrienssen [9], the human capital efficiency (HCE) may be expressed as a percentage of the added value (VA) and human capital (HC).

$$HCE = \frac{VA}{HC} = \frac{\text{Operating profit} + HC}{HC} \quad (2)$$

The metrics of human capital efficiency according to 4 factors

One of the most commonly used measurements of the efficient use of human capital is based on the study of the authors Bontis and Fitz-enz [7]. The HC measurement is realised on the basis of four essential metrics, which are income factor, cost factor, profit factor and HC ROI. The main goal of any organisation is to increase income, or profit per one employee. Increasing human capital and its efficiency bring higher financial results per one employee. Its development is largely influenced just by raising the level of education together with the total satisfaction of the employees. For this reason, human capital has also a direct impact on the profitability of investment in the ROI enterprise. The FTE indicator in the metrics is the full-time equivalent (FTE), which includes full and partial employment contracts and volunteers. It shows how much time has been spent on generating income (Table 1).

Metrics of human capital efficiency according to 4 factors Table 1

Indicator	Formula / metric
COST FACTOR	$\frac{\text{operating costs}}{FTE}$
INCOME FACTOR	$\frac{\text{income}}{FTE}$
PROFIT FACTOR	$\frac{\text{income} - \text{operational costs}}{FTE}$
HC ROI	$\frac{\text{income} - (\text{operational costs} - \text{wages and benefits})}{FTE}$

Source: [7]

Human Economic Value Added – HEVA

The measurement of the economic value added of human capital (HEVA) is the possible way used to measure human capital efficiency. The indicator is based on the economic value added (EVA) from the Stern Stewart. The purpose of the calculation is to find whether the activities of the management have added value for the enterprise. The indicator EVA is also used in the measurement of human capital. In this case, it is completed in divisor of item FTE (full-time equivalent, which includes full-time, partial work load, or volunteers). It is one of the fundamental metrics designed to measure human productivity, that is, how much time people spend creating profits for the enterprise [10].

$$HEVA = \frac{(NOPAT - WACC.C)}{FTE} \quad (3)$$

where:

NOPAT net operating profit,
WACC weighted average cost of capital,
C the size of the capital invested.

Human Capital Cost Factor – HCCF

Human capital cost factor (HCCF) is another indicator taking into account the costs associated with wages and the calculation of the value. The basic principles of human capital are ranked into four categories. It is salary and costs for benefits for employees, volunteer wages, costs associated with the fluctuations and absence. Salary represents a real wage paid to the employees of the enterprise. Considering the benefits, there are calculated all benefits provided by the company for employees of the enterprise. The costs associated with fluctuations themselves involve several factors, such as severance payments, or costs the enterprise must spend on a new employee or his/her total training and enlistment in the collective of the enterprise [10].

HCCF = salary + benefit costs + volunteer wages + costs associated with fluctuation and absence (4)

Human Capital Value Added – HCVA

HCVA is an indicator of human capital value added, which is the sum of all operating costs, where the costs of employees' work and other associated benefits are not taken into account. The result is an operating profit per employee working full-time. In the majority of enterprises, the employees are just the most important asset. Therefore, human capital is the main prerequisite for success in the future. On the other hand, they are often associated with large costs, or in other words, investments. Due to this fact, it is very important to know what value added they can bring for the enterprise. Using this indicator, we can find the average profitability per employee of the enterprise [11].

$$HCMV = \frac{\text{sales} - (\text{cost of wages and benefits})}{\text{FTE}} \quad (5)$$

Human Capital Market Value – HCMV

Human capital market value (HCMV) is an interesting indicator for the calculation of human capital efficiency. It expresses the relationship between the value of the enterprise and its value based on the accounting. In terms of some literature, indicator is understood as a metric for measuring the HC value. In the context of a given indicator, there are also included, except HC, various forms of intellectual capital. Due to the fact that there is mainly taken into account the value of intangible assets, in addition to human capital, there may be also included capability of processes, public awareness about the brand, or setting up of marketing processes. The indicator is more interesting for economists, analysts, but it is not very useful for managers. It reflects the market value premium per employee [10].

$$HCMV = \frac{\text{enterprise market value} - \text{accounting value}}{\text{FTE}} \quad (6)$$

All of these metrics which are designed to calculate the efficiency of human capital differ from each other in the approach to the calculation of the upper part of the formula. The denominator of these metrics have always found the number of partial and full-time jobs. However, this cannot clearly define the value of human capital in a particular firm. For this reason, it is necessary to pay more attention to determining the value of human capital. In business, it is a challenge for human resource managers to find various indirect indicators that are able to assess the size of human capital with regard to the sectorial focus of the company.

There are a number of consulting companies, which have extensive databases, monitor and evaluate a number of indicators in the field of human capital. Involvement, however, is worth the money and does not contain a thorough analysis aimed at identifying key indicators of human capital for a particular company, but only the evaluations based on the internal data provided and its comparison with other companies operating in the same industry [12].

4. Principles of efficient management and measurement of human capital performance

Efficient performance management and human capital require new perspectives and new competencies of line managers and HR professionals. Human capital management considers people the wealth or asset and not as a cost item, as it was regarded in the concept of human resources management or personnel management. However, in connection with the HC management, the application of metrics at any price and without the advanced concept, can ultimately be counterproductive for the enterprise.

Becker, Huselid, Ulrich [13] suggested six principles for measuring performance and efficiency of human capital. The use of these principles can transform human capital into major driving engines of financial performance of any enterprise:

- Focus on the strategic impact of human capital.
- Avoid the alchemy of human capital.
- Measure not only the level of human capital, but the relationships as well.
- Recognise the limitations of benchmarking.
- Do not start with the measurement.
- Think in terms of the "architecture" of human capital.

In this dynamic era it is very important for enterprises to identify strategies how in current high-velocity environment sustaining its competitive advantages [14]. Globalization, increasing independence and complexity, new technologies and climate change create new challenges for enterprises [15]. The investments to the human capital present the most profitable form of improving the situation in firms from the long-term and perspective point of view [16]. To increase performance and efficiency of human capital contributes introduction of the motivation program. Employers should focus on the process of motivation through factors based on interpersonal relationships and job security, too. We can assume that motivation through these factors will increase and employees will consider them more important [17]. Also we can say that in spite of employee heterogeneity in terms of age, seniority and level of completed education it is possible to create a unified motivational programme for the analysed enterprise that will suit all employees regardless of their age, seniority or education. Its main items are following factors: basic salary, job security, good work team, further financial reward and fair appraisal system in different order according to preferences of specific group of employees [18]. Results indicated that innovation competency positively influences organisational performance and organisational knowledge management influences innovation competency positively, too [19].

5. Conclusion

For the efficient functioning of the enterprise on the open market, it is necessary to increase the efficient use of the enterprise production inputs. In terms of human capital management, it means to search for the connection of enterprise performance with the use of specific knowledge, skills and abilities of its employees in order to increase competitiveness. There were identified and compared some of the metrics published in the field of human capital efficiency. However, for the implementation in practice of a specific enterprise, it is necessary to adapt these indicators to the specific conditions of enterprise, in the context of enterprise strategy. With respect to the fact that there are no united metrics adopted and used in this area, it is precisely the role of the HR managers to take advantage of their knowledge, creativity, analytical thinking and design such metrics of efficient use of human capital for the conditions of their enterprises to be able to direct the care of human capital towards increasing individual performance and the performance and competitiveness of the whole enterprise. We can see a great space for further research in this area.

Acknowledgement

This article was created as part of application of project: Innovation and internationalization of Education – Instrument to increase the quality of the University of Zilina in the European educational area. Modern Education for the Knowledge Society ITMS 26110230079/Project is funded by EU; Grant VEGA 1/0421/13 and grant VEGA 1/0526/13.

References

[1] SIKYR, M.: *Best Practice Approach to Human Resource Management.* The 9th Intern. Days of Statistics and Economics, Prague, September, 2015, pp. 1405-1414, ISBN 978-80-87990-06.

[2] KUCHARCIKOVA, A., BABISOVA, R., BACINSKA, K., GRIVALSKA, I., KRISTANOVA, J., VADOVIC, T.: The Human Capital Efficiency in Enterprises. *Scientific J.,* Faculty of Social and Economic Relations: Alexander Dubcek University of Trencin, vol. 13, No. 1, 2015, pp. 58-65, ISSN 1336-3727.

[3] DURISOVA, M., JACKOVA, A., HOLKOVA, B.: *Business Economics*. EDIS: University of Zilina, 2015, 247 p., ISBN 978-80-554-0997-9.

[4] BARANIK, M., HABANIK, J.: *Introduction to Macroeconomics,* IRIS, Bratislava, 2002, p. 232, ISBN 80-89018-45-9.

[5] VOJTOVIC, S.: *The Conception of Personnel Management Human Resources*. Grada Publishing: Prague, 2012, p. 192.

[6] MAZOUCH, P., FISCHER, J. *Human Capital - Measurement, Context, Forecasts*. Praha: C. H. Beck, 2011, 116 p., ISBN 978-80-7400-380-6.

[7] BONTIS, N., FITZ-ENZ, J. Intellectual Capital ROI: A Causal Map of Human Capital Antecedents and Consequents. *J. of Intellectual Capital*, vol. 3, No. 3, 2002, pp. 223-247.

[8] MARIMUTHU, M., AROKIASAMY, L., ISMAIL, M.: Human Capital Development and its Impact on Firm Performance: Evidence from Developmental Economics. *The J. of Intern. Social Research*, vol. 2/8, summer 2009, pp. 265-272.

[9] ANDRIENSSEN, D.: *Making Sense of Intellectual Capital,* Taylor & Francis Ltd: New York, 2011, ISBN: 978-0-7506-7774-5, p. 456.

[10] FITZ-ENZ, J.: *The ROI of Human Capital: Measuring the Economic Value of Employee.* Amacom: New York, 2009, 310 p., ISBN 978-0-8144-1332-6.

[11] MARR. B.: Key Performance Indicators (KPI): The 75 Measures Every Manager Needs to Know. FT Press: Harlow, 2012, 376 p., ISBN 978-027375011.

[12] KUCHARCIKOVA, A., DURISOVA, M.: Human Capital Efficiency in Transport Enterprise. *Proc. of 18th Intern. Conference. Transport Means,* 2014, pp. 23-24,10.20104, Kaunas University of Technology, pp. 239-243. ISSN 2351-4604.

[13] BECKER, B. E., HUSELID, M. A., ULRICH, D.: *Six Key Principles for Measuring Human Capital Performance in Your Organization*. State University of New York: Buffalo, 2002, pp. 1-16.
[14] CIRJEVSKIS, A.: Sustainability in Higher Education: Discourse on Dynamic Capabilities of Privately Run Higher Educational Institutions (HEI) in Latvia. *J. of Security and Sustainability Issues,* vol. 5, No. 1, 2015, pp. 111-112, ISSN 2029-7071.
[15] KLUCKA, J.: Renewable Resources and Critical Infrastructure - the Framework. *Advanced Material Research*, vol. 1001, 2014, pp.475-483, ISSN 1662-8985.
[16] KONUSIKOVA, L., KUCHARCIKOVA, A.: Approaches to Active Labour Market Policy in the Slovak Republic, Czech Republic and Finland. *Communications – Scientific Letters of the University of Zilina*, vol. 17, No 2, 2015, pp. 73-79, ISSN 1335-4205.
[17] HITKA, M., STACHOVA, K., BALAZOVA, Z., STACHO, Z.: Differences in Employee Motivation at Slovak Primary Schools in Rural and Urban Areas. *Intern. Education Studies,* vol. 8, No. 5, 2015, pp. 33-42, ISSN 1913-9020 E-ISSN 1913-9039.
[18] HITKA, M., BALAZOVA, Z.: The Impact of Age, Education and Seniority on Motivation of Employees. *Business: Theory and Practice,* 16(1), 2015, pp. 113-120 ISSN 1648-0627/eISSN 1822-4202, doi:10.3846/btp.2015.433.
[19] JUNGES, F. M., GONCALO, C. R., GARRIDO, I. L., FIATES, G. G. S.: Knowledge Management, Innovation Competency and Organisational Performance: A Study of Knowledge-intensive Organisations in the IT Industry. *Intern. J. of Innovation and Learning,* vol. 18, No. 2, pp. 198-221, ISSN 1471-8197/eISSN 1741-8089.

Peter Czimmermann - Stefan Pesko - Jan Cerny *

UNIFORM WORKLOAD DISTRIBUTION PROBLEMS

In this paper we review common studies from the years 1984-2015 of problems occurring in uniform scheduling of workload distributions. These problems were formulated first by Pesko in dealing with the practical problem of regular scheduling of service vehicles. Given the nonnegative, real matrix with daily records of vehicles as columns, we need to minimize some irregularity measure of row sums (workloads) by permuting matrix columns. The problem has various practical modifications, such as the weighted rows of matrix, workload uncertainty variance, graph approach of exchange of elements and generalized inverse formulations. Some of them are presented in the paper.

Keywords: *Regular scheduling, matrix permutation, irregularity measure, NP-hard problem.*

1. Introduction

In this paper we study problems which occur in uniform scheduling of workload distribution. The uniformity is expressed by an irregularity measure and the goal is to minimize the irregularity measure of parts of the schedule. This problem is motivated by the following job scheduling problem:

Let n vehicles and m×n jobs be given. Every job has assigned to it the value ai,j which represents its quantity. These values form the m×n matrix A. We need to find the most regular m-days job schedule, which gives minimal difference between the sums of rows of the permuted version of the matrix A.

The problem is also called the Matrix Permutation Problem (MPP) which was first mentioned in [1]. Further investigation of MPP can be found in [2], where this problem was solved for garbage trucks.

The mathematical formulation of the problem is the following:

Let a nonnegative real $m \times n$ matrix $A=(a_{ij})$ be given. Let $I=\{1,2,...,m\}$, $J=\{1,2,...,n\}$ be sets of row and column indices. For each column $j \in J$ of matrix A we will use the notation π_j for the permutations of elements in that column. Let $\pi=(\pi_1,\pi_2,...,\pi_n)$ denote the vector of permutations of all columns of A and let A^π denote the permuted matrix. Let P_{mn} be the set of all such permutation vectors. Let $S^\pi=(s_1^\pi, s_2^\pi,...,s_m^\pi)$ denote the vector of row sums of permuted matrix A^π. Let f be a function called the irregularity measure. The following optimization problem will be called the uniform workload distribution problem (UWDP) $min\{f(S^\pi); \pi \in P_{mn}\}$.

The irregularity measure $f: J^m \to \langle 0,\infty)$ is any Schur-convex function, where $J=(a,b)$ is an interval. More on irregularity measures and Schur-convex functions can be found in [3]. The most used irregularity measures are:

- $f_{sqr}^{\delta}(x_1,x_2,...,x_m) = (x_1-\delta)^2 + (x_2-\delta)^2 + ... + (x_m-\delta)^2$
- $f_{abs}^{\delta}(x_1,x_2,...,x_m) = |x_1-\delta| + |x_2-\delta| + ... + |x_m-\delta|$
- $f_{dif}(x_1,x_2,...,x_m) = \max(x_1,x_2,...,x_m) - \min(x_1,x_2,...,x_m)$
- $f_{\max}^{\delta}(x_1,x_2,...,x_m) = \max(x_1,x_2,...,x_m) - \delta$
- $f_{min}^{\delta}(x_1,x_2,...,x_m) = \delta - \min(x_1,x_2,...,x_m)$

$(\text{where } \delta = \min(x_1+x_2+...+x_m)/m)$

2. Computational complexity of UWDP

In [4] it was proved that UWDP (MPP) is an NP-hard problem. In [2] it was shown that even a 2-row version of UWDP is NP-hard. The proof of this fact was established using transformation from the Set partition problem (definition of SPP can be found in [5]).

* [1]Peter Czimmermann, [1]Stefan Pesko, [2]Jan Cerny
[1]Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Zilina, Slovakia
[2]Department of Exact Methods, Faculty of Management, University of Economics, Prague, Czech Republic
E-mail: peter.czimmermann@fri.uniza.sk

On the other hand, the two-column case is solvable in polynomial time. A simple polynomial algorithm can be found in [4]. It is enough to order the elements of the first column in a descending order and the elements of the second column in an ascending order. It is possible to show that we obtain the optimal solution for any irregularity measure mentioned above. The complexity of the algorithm based on this approach is $O(m \log m)$.

3. Solutions of UWDP

In this section we describe two heuristics for the solution of general UWDP. Both algorithms are based on the fact that the two-column case is solvable in polynomial time. We also consider the representation of UWDP by two models of mathematical programming.

3.1. Decomposition method

In the paper [2] the following heuristic was introduced:

Input: Matrix $A=A_1$ of the type $m \times n$.
For $i=1$ to $n-1$ do:
Create submatrix $B_{m\times 2}$ from the first pair of columns of A_i.
Solve the UWDP for matrix $B_{m\times 2}$. (The row-sum vector of the solution is denoted by S^π.)
Create matrix A_{i+1} by replacing the first two columns of A_i with column S^π.
Output: Solution for the matrix A_{n-1}.

Tests showed that this method is not very successful since the set of possible column permutations is very restricted.

3.2. Stochastic decomposition method

This method was introduced and studied in [6 and 7]. Let J_1 be the nonempty proper subset of column indices J and $J_2=J-J_1$. The set $\{J_1,J_2\}$ will be called the admissible decomposition of the columns of matrix A.

S1: Let $\pi=(\pi_1,\pi_2,\ldots,\pi_n)$ be arbitrary permutations of index set I.
S2: Choose randomly $\{J_1,J_2\}$ – admissible decomposition of the columns of A.
S3: Solve the two-column UWDP with aggregated matrix $B=(b_{ij})_{(m\times 2)}$

where $b_{i1}=\sum_{j\in J_1} a_{\pi_j(i),j},\ b_{i2}=\sum_{j\in J_2} a_{\pi_j(i),j}, i \in I$

We get optimal column permutations φ_1, φ_2, satisfying conditions

$$b_{\varphi_1(1),1} \le b_{\varphi_1(2),1} \le \cdots \le b_{\varphi_1(m),1},$$
$$b_{\varphi_1(1),2} \le b_{\varphi_2(2),2} \le \cdots \le b_{\varphi_2(m),2}.$$

S4: Update π by applying permutation φ_1 (resp. φ_2) to π_j for $j \in J_1$ (or $j \in J_2$).
S5: If no stopping criterion is satisfied GOTO S2, else END, $A^\pi = \left(a_{\pi_j(i),j}\right)$.

Tests on real data give promising results. The authors Š. Peško and M. Kaukič state conjecture that this algorithm can also be (with a suitable iteration count and sufficient number of restarts) the exact algorithm (at least for irregularity measure f_{dif}).

3.3 Model of linear programming

In the mentioned work [7] the following model of mixed, integer, linear programming (MILP) was introduced. Objective function is f_{dif}.

$$\begin{aligned}
& z_U - z_L \to \min \\
\text{s.t.}\ & \sum_{k\in I} x_{ijk} = 1 && \forall (i,j) \in I \times J, \\
& \sum_{i\in I} x_{ijk} = 1 && \forall (k,j) \in I \times J, \\
& z_i = \sum_{(j,k)\in J\times I} a_{kj} x_{ijk} && \forall i \in I, \\
& x_{ijk} \in \{0,1\} && \forall (i,j,k) \in I \times J \times I \\
& z_L \le z_i \le z_U && \forall i \in I, \\
& S_L \le z_L \le z_U \le S_U.
\end{aligned}$$

The value of x_{ijk} is equal to one if $\pi_j(i)=k$ and otherwise zero. The real variables $z_i\ i\in I$ are the i-th row sums of permuted matrix A^π and the variables z_L, z_U are variables for lower and upper bounds S_L, S_U of row sums.

3.4 Model of quadratic programming

In [7] the model of mixed, integer, quadratic programming (MIQP) was introduced. Objective function is f^δ_{sqr}.

$$\begin{aligned}
& \sum_{i\in I} (z_i - \delta)^2 \to \min \\
\text{s.t.}\ & \sum_{k\in I} x_{ijk} = 1 && \forall (i,j) \in I \times J, \\
& \sum_{i\in I} x_{ijk} = 1 && \forall (k,j) \in I \times J,
\end{aligned}$$

$$z_i = \sum_{(j,k)\in J\times I} a_{kj}x_{ijk} \qquad \forall i \in I,$$
$$x_{ijk} \in \{0,1\} \qquad \forall (i,j,k) \in I \times J \times I$$
$$S_L \le z_i \le S_U \qquad \forall i \in I.$$

Variables x_{ijk} and z_i mean the same as in the MILP model. Tests show that the MILP model is more effective than the MIQP model, but the mathematical programming solvers (even of such quality as Gurobi) have sometimes difficulties with the solving of not very large UWDP instances in reasonable time (since the model has large number of bivalent variables).

4. Restricted sets of permutations

In [8] the generalization of UWDP was suggested in which the set of permutations is restricted. The permitted permutations are represented by graphs. Paper [9] deals with conditions which allow to represent the set of permitted permutations by graphs. The two-column case of this generalization is studied in [10]. It was shown that this problem can be solved as the most regular perfect matching in a graph. In the mentioned paper [10], there are introduced exact polynomial algorithms for finding the most regular perfect matching in graphs for irregularity measures $f_{dif}, f^{\delta}_{\max}$ and $f^{\delta}_{\min}$. The existence of exact polynomial algorithms for measures f^{δ}_{sqr}, f^{δ}_{abs} was an open problem.

We will show that there is also polynomial algorithm for irregularity measure f^{δ}_{sqr}, but the weights of edges must satisfy a special property (which is fulfilled by graphs that arise from the mentioned generalization of UWDP). Let an edge-weighted graph $G=(V,E,w)$ be given. Let $V=\{v_1,v_2,\ldots,v_{2m}\}$. Let $w : E \to R_0^+$ be a function that represents the weights of edges. If there exists a mapping $x : V \to R_0^+$ such that

$$\forall e \in E \;\; e = \{v_i, v_j\} \;\; w(e) = x(v_1) + x(v_j),$$

we say that w is induced by x. There exist mappings w that cannot be induced by any x. For example, we can consider complete graph on 4 vertices v_1,v_2,v_3,v_4 with edge weights

edge	$\{v_1,v_2\}$	$\{v_1,v_3\}$	$\{v_1,v_4\}$	$\{v_2,v_3\}$	$\{v_3,v_4\}$	$\{v_3,v_4\}$
weight	2	2	2	4	2	4

This leads to the unsolvable system of linear equations:

$x_1 + x_2 = 2$
$x_1 + x_3 = 2$
$x_1 + x_4 = 2$
$x_2 + x_3 = 4$
$x_2 + x_4 = 2$
$x_3 + x_4 = 4$

where x_1,x_2,x_3,x_4 are possible values of vertices. Hence w cannot be induced by any x.

Let $G=(V,E,w,x)$ be a graph. Let $V=\{v_1,v_2,\ldots,v_{2m}\}$ and w be a mapping on edge set induced by mapping x on vertex set. Then the problem of finding the most regular perfect matching is the problem to find such perfect matching in G, for which the function f^{δ}_{sqr} is minimal. The weight of edge e_i will be denoted by w_i. If the perfect matching contains the edges $e_1,e_2,\ldots,e_m$, then we have

$$f^{\delta}_{sqr}(w_1, w_2, \ldots, w_m) = \sum_{i=1}^{m} (w_i - \delta)^2$$

where $\delta=(w_1+w_2+\cdots+w_m)/m$. Since the matching is perfect (it contains all vertices of G) and the edge weights are induced by vertex values, we obtain:

$$\sum_{i=1}^{m} w_i = \sum_{j=1}^{2m} x(v_j) = m\delta$$

The last sum is taken over all vertices of G and hence it is a constant. The consequence is that δ is a constant for given graph G and mapping x. Then

$$\sum_{i=1}^{m} (w_i - \delta)^2 = \sum_{i=1}^{m} (w_i^2 - 2\delta w_i + \delta^2) =$$
$$\sum_{i=1}^{m} w_i^2 + m\delta^2 - 2\delta \sum_{i=1}^{m} w_i = \sum_{i=1}^{m} w_i^2 + m\delta^2 - 2m\delta^2 =$$
$$\sum_{i=1}^{m} w_i^2 + m\delta^2.$$

It means that the problem to find minimum of the function f^{δ}_{sqr} is in our case equivalent to the problem to find the minimum of the function f_{sqr}. Hence the problem of finding the most regular perfect matching in a graph $G=(V,E,w,x)$ is equivalent to the problem of finding minimal perfect matching in graph $G=(V,E,w^2)$ and this problem is solvable in polynomial time [11].

5. Conclusions and further research

There are several approaches how to represent uncertainty. In [12] the UWDP in interval arithmetic was introduced. It means that elements of matrix A are not exact values, but we have intervals, to which these values belong. It was proved that this problem is NP-complete and the two-column case is solvable in polynomial time.

The most common approach in systems with uncertainty is the fuzzy arithmetic. Our future plan is to define the UWDP in fuzzy arithmetic. We suppose that the complexity results will not change but it remains an open problem.

The weighted version of the problem was studied in [13]. It is called the weighted uniform workload distribution problem

(WUWDP). Let $S^{\pi} = (s_1^{\pi}, s_2^{\pi}, ..., s_m^{\pi})$ denote the vector of row sums. Let $w_1, w_2, ..., w_m$ be weights of rows. In this version of the problem we require that the values $w_1, s_1^{\pi}, w_2 s_2^{\pi}, ..., w_m s_m^{\pi}$ are as uniform as possible. It is easy to consider that the UWDP is a special case of weighted version, where $w_1 = w_2 = ... = w_m = 1$.

At the end, we can also mention some works which deal with similar problems to the UWDP. For example: the column permutation algorithm of a special case of the UDWP is studied in papers [14] and [15], the uniform k-partition problems are classified and examined in [16 and 17].

Acknowledgment

The research of P. Czimmermann was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic, the Slovak Academy of Sciences under project VEGA 1/0518/15 *"Resilient rescue systems with uncertain accessibility of service"*, and the research of Š. Peško was supported by the Slovak Research and Development Agency under projects APVV-0760-11 *"Designing of fair service systems on transportation networks"* and APVV-14-0658 *"Optimization of urban and regional public personal transport"*.

References

[1] COFFMAN, E. G., YANNAKAKIS, M.: Permuting Elements within Columns of a Matrix in order to Minimize Maximum Row Sum, *Mathematics of Operations Research*, 9(3):384390, 1984.

[2] CERNY, J., VASEK, K., PESKO, S., PALUCH, S., ENGELTHALLER, D.: *Transport Scheduling and its Optimization (in Slovak)*, Research report III-8-9/03, Research Institute of Transport, 1986.

[3] CERNY, J., KLUVANEK, P.: *Principles of Mathematical Theory of Transport (in Slovak)*, VEDA, Bratislava, 1991.

[4] TEGZE, M., VLACH, M.: *On the Matrix Permutation Problem,* Zeitschrift fur Operations- Research, 30(3), 1986.

[5] GAREY, M. R., JOHNSON, D. S.: *Computers and Intractability, A Guide to the Theory of NP-completeness,* Freeman: San Francisco, 1979.

[6] PESKO, S., HAJTMANEK, R.: *Matrix Permutation Problem for Fair Workload Distribution*, Mathematical Methods in Economics, 32nd international conference, Olomouc, 2014.

[7] PESKO, S., KAUKIC, M.: *Stochastic Algorithm for Uniform Workload Distribution Problem*, Proc. of 13th Intern. Conference Informatics, IEEE, Poprad, 2015.

[8] CZIMMERMANN, P., PESKO, S.: The Regular Permutation Scheduling on Graph, *J. of Information, Control and Management Systems,* vol. 1, No. 1, 2003.

[9] CZIMMERMANN, P.: A Note on using Graphs in Regular Scheduling Problems, *Communications - Scientific Letters of the University of Zilina,* vol. 5, No. 4, 2003.

[10] CZIMMERMANN, P.: The Most Regular Perfect Matchings in Graph, *Studies of the University of Zilina, Mathematical Series,* vol. 16, No. 1, 2003.

[11] EVEN, S., KARIV, O.: *An Algorithm for Maximum Matching in General Graphs,* Proc. of 16th Annual Symp. on Foundations of Computer Science IEEE, New York, 1975.

[12] PESKO, S.: The Matrix Permutation Problem in Interval Arithmetic, *J. of Information, Control and Management Systems,* vol. 4, No.1, 2006.

[13] PESKO, S., CERNY, J.: Uniform Splitting in Managerial Decision Making, *Economics and Management,* 9(4), 2006.

[14] PUCCETTI, G., RUSCHENDORF, L.: Computation of Sharp Bounds on the Distribution of a Function of Dependent Risk, *J. of Computational and Applied Mathematics,* 236(7), 2012.

[15] BOUDT, K., VANDUFFEL, S., VERBEKEN, K.: *Block Rearranging Elements within Matrix Columns to Minimize the Variability of Row Sums,* http://ssrn.com/abstract=2634669, 2015.

[16] DELL'OLMO, P., HANSEN, P., PALLOTTINO, S., STORCHI, G.: On Uniform k-partition Problems, *Discrete Applied Mathematics,* 150(1-3), 2005.

[17] PESKO, S.: Minimal Total Area Convex Set Partitioning Problem, *Communications - Scientific Letters of the University of Zilina,* vol. 11, No. 3, 2009.

Lubos Kucera - Igor Gajdac - Martin Mruzek *

SIMULATION OF PARAMETERS INFLUENCING THE ELECTRIC VEHICLE RANGE

Increasing the range of electric vehicles is crucial for electromobility expansion. Using the simulation it is possible to identify the parameters that have a significant impact on the electric vehicle range. Therefore, the analysis of parameters influencing the range was created. Simulation was based on real driving cycles for better accuracy.

Keywords: *Electric vehicle, range, simulation, driving cycle.*

1. Introduction

E-mobility importance in the world has been continuously growing. The range of electric vehicles is considered a major barrier to the acceptance of e-mobility [1]. The battery is the most important component for increasing the range of electric vehicles. New technologies can improve specific energy of the batteries. Also lightweight materials are increasing the electric vehicle range. Another possibility for increasing the range is the use of a range extender [2]. It is a combustion engine which charges the batteries via generator [3 and 4].

Simulations were created in Matlab/Simulink to analyse the parameters affecting the range. Real route as an input to the simulation for better accuracy was used. Routes were measured by using the experimental vehicle Edison.

2. Electric vehicle simulation

Matlab/Simulink was chosen for the simulation. Simulink is an extension of Matlab for modelling and simulation of dynamic systems. The electric vehicle model consists of five main subsystems. Inputs to the simulation are measured data from a real driving cycle. Demanded power is calculated in the dynamic subsystem based on chosen parameters of the vehicle. Speed and efficiency characteristics are defined in the electric motor model by using a lookup table. Efficiency characteristic is defined for the motor and generator mode and takes into account losses in the controller. Discharging and charging of the battery is controlled in the control strategy subsystem. Battery behaviour is determined in the battery subsystem [5, 6, 7, 8 and 9].

The demanded power for vehicle propulsion supplied from the battery is calculated by equation (1).

$$P_{z\_aku} = \left[m\frac{dv}{dt}\lambda + mg(\sin\alpha + f\cos\alpha) + \frac{1}{2}c_x S_x \rho v^2 \right]\frac{v}{\eta} \quad (1)$$

where v is the speed of the vehicle, m is the weight, λ is the rotational inertia factor, α is the road angle, f is the rolling resistance coefficient, c_x is the aerodynamic drag coefficient, S_x is the frontal area, ρ is the air density, η is the efficiency of the drive, where is included the efficiency of the individual elements of the drive such as electric motor and mechanical efficiency [10].

Power provided via recuperation kinetic energy during braking is calculated:

$$P_{do\_aku} = \left[m\frac{dv}{dt}\lambda + mg(\sin\alpha + \cos\alpha) + \frac{1}{2}c_x S_x \rho v^2 \right]\frac{v}{\eta} \quad (2)$$

Total energy flow E_{bat} to the battery is expressed by the formula (3).

$$E_{bat} = \int P_{z\_aku}\,dt - \int P_{do\_aku}\,dt \quad (3)$$

The resulting range of electric vehicle to the battery discharge is expressed:

$$s = \int_{100\%\,SOC}^{0\%SOC} v\,dt \quad (4)$$

* [1]Lubos Kucera, [2]Igor Gajdac, [2]Martin Mruzek
[1]Department of Design and Mechanical Elements, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]University Science Park, University of Zilina, Slovakia
E-mail: lubos.kucera@fstroj.uniza.sk, igor.gajdac@uvp.uniza.sk, martin.mruzek@uvp.uniza.sk

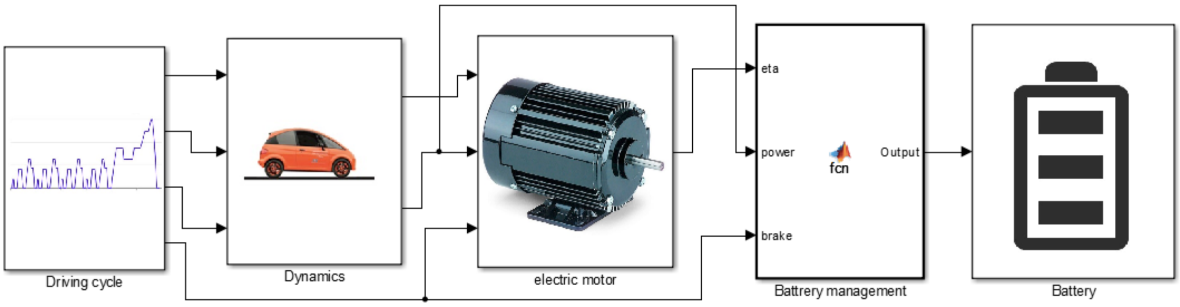


Fig. 1 Electric vehicle model

In Figure 1 is shown the model of the electric vehicle created in Simulink.

3. Real driving cycle

A driving cycle is defined as a speed of a vehicle versus time. Standardised cycles for measuring fuel consumption and harmful emissions are used for the vehicle simulations. However, these cycles are simplified and do not reflect reality [11]. Driving cycles were chosen as an input to the simulation in order to reach the reality. The measurement of driving cycles was performed by the experimental electric vehicle Edison, which allows data logging of dynamic parameters. The Edison was made at the University of Zilina. Curb weight of the vehicle is 1048 kg including batteries. Propulsion is provided by a compact lightweight aluminium, air-cooled asynchronous electric motor AKOE with a nominal output of 16 kW and maximum power of 30 kW, with CURTIS controller and LiFeYPO4 300Ah batteries. The main part of the vehicle is a tubular steel space frame [12].

Two real routes were chosen to the simulations. Route 1 (Fig. 2) consists of driving in a city and its surroundings. The maximum speed on route 1 was 76.5 km.h^{-1} and the average speed was 41.5 km.h^{-1}. Route 2 was (Fig. 3) measured only in the city traffic. The maximum speed on route 2 was 50.2 km.h^{-1} and the average speed was 18.6 km.h^{-1}.

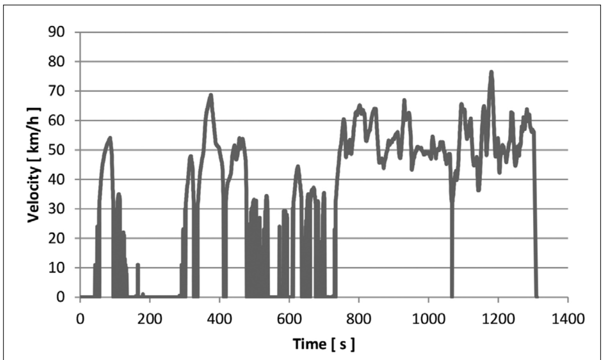


Fig. 2 Course of speed on route 1

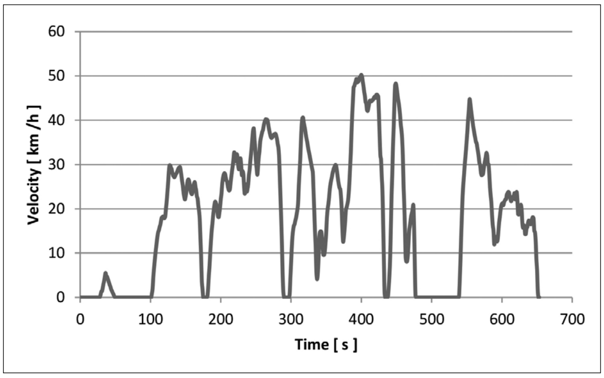


Fig. 3 Course of speed on route 2

4. Analysis of parameters influencing the range

The electric vehicle Edison can also measure energy consumption via its data logging system. Therefore, it is possible to compare the results of simulation with reality. Construction parameters of the experimental vehicle Edison were selected as an input to the simulation. To the curb weight of the EDISON was added the weight of two passengers and cargo. The total vehicle weight was 1250 kg. The 24 kWh battery pack was chosen. It was selected lithium battery with capacity of 24 kWh. The battery was not discharged below 10%. Nominal power of the electric motor was 16 kW.

Validation of the model with reality was performed on routes 1 and 2. Real energy consumption of the electric vehicle 15.41 kWh/100km was measured on route 1. Simulated energy consumption of the electric vehicle was 15.56 kWh/100 km on route 1. In the second case real energy consumption 13.81 kWh/100km was measured on route 2. Simulated energy consumption was 14.35 kWh/100 km on route 2. The difference between reality and simulation was 1% on route 1 and 3.9% on route 2. Impact on the electric vehicle range was simulated on these models with various weights and various aerodynamic drag coefficients. Also the impact of low temperatures was simulated on the range.

Impact of weight

One of the possibilities to increase the range is to reduce the vehicle weight. Figure 4 shows the impact of vehicle weight on the range and Fig. 5 shows the impact of the weight on energy consumption.

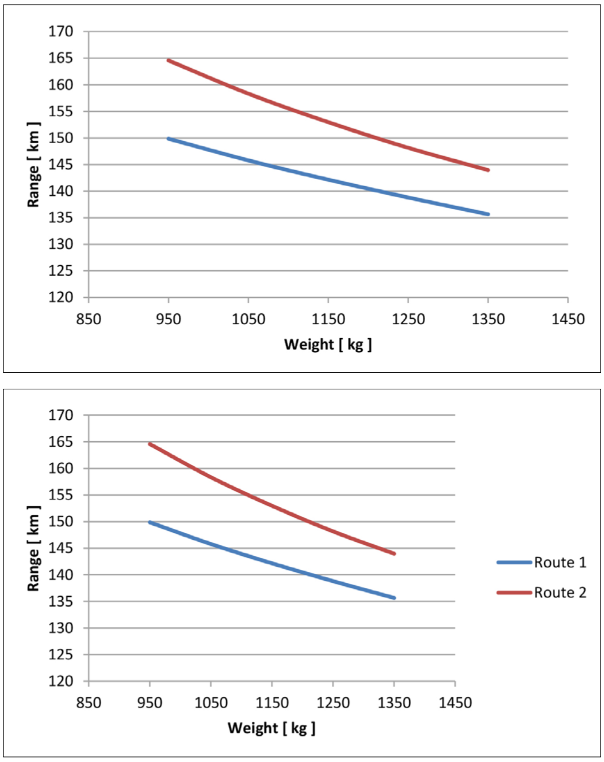


Fig. 4 Impact of weight on the range

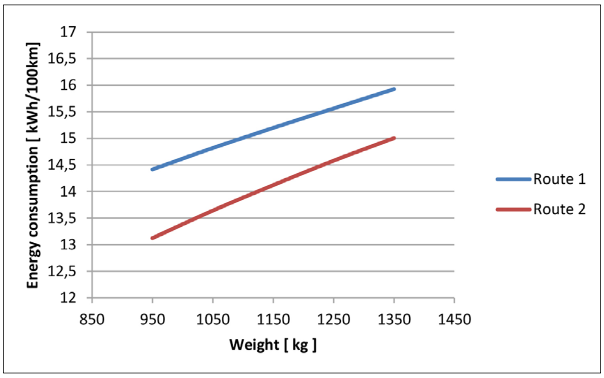


Fig. 5 Impact of weight on energy consumption

Impact of aerodynamic drag coefficient c_x

The aerodynamic drag coefficient value is commonly 0.32 for city vehicles. As you can see in Fig. 6, the aerodynamic drag coefficient had a bigger impact on the range on route 2 because aerodynamic drag rises quadratically with speed. The average speed on route 2 was 18.6 km.h^{-1} and on route 1 was 41.5 km.h^{-1}.

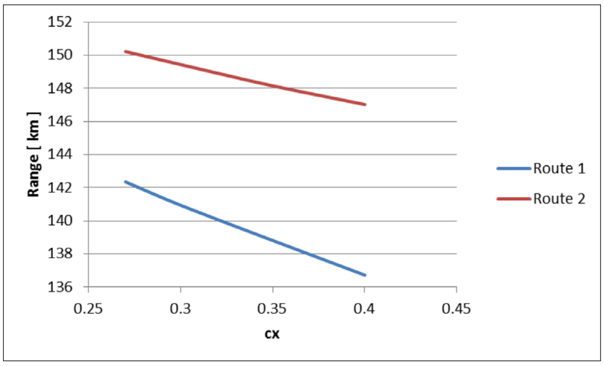


Fig. 6 Impact of c_x on the range, 1250kg vehicle weight

Figure 7 shows the impact of c_x on energy consumption.

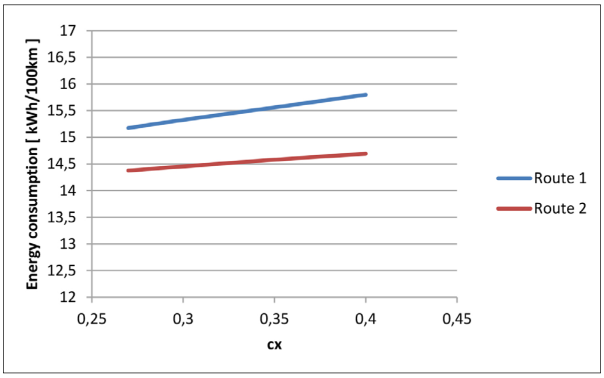


Fig. 7 Impact of c_x on energy consumption, 1250kg vehicle weight

Impact of low temperature

The electric vehicles suffer from significant range loss in sub-zero temperatures. The poor performance of Lithium cells at sub-zero temperatures implies significantly increased internal resistance [13, 14 and 15]. Figure 8 shows behaviour of LiFeYPO4 battery at 0°C and 18°C temperature when vehicle needed power of 10.4 kW for propulsion. Low 0°C temperature caused the battery voltage drop and current increasing. The battery model was modified for low temperatures impact simulation. The battery model was modified based on the testing of EV Edison at 0 °C at the University of Zilina campus. Figure 9 shows how the range which was simulated on the route 1 changed when the temperature was 0 °C. The simulation shows that the range is decreased approximately by 7.5 kilometres in all weight variations. It should be noted that the simulation did not take into account energy which is needed for heating the passenger area. In Figure 10 you can see that energy consumption increased approximately by 0.8 kWh/100km due to low temperature.

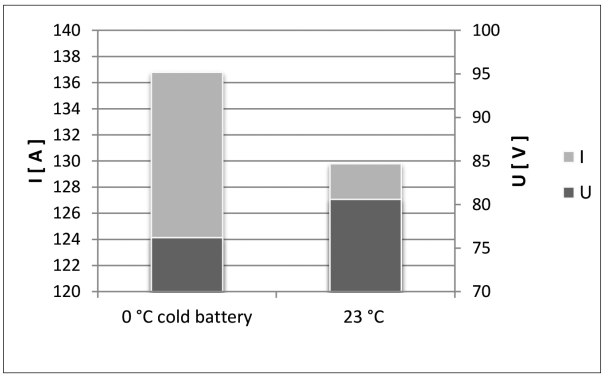


Fig. 8 LiFeYPO4 battery producing 10.4 kW at 0°C and 18°C

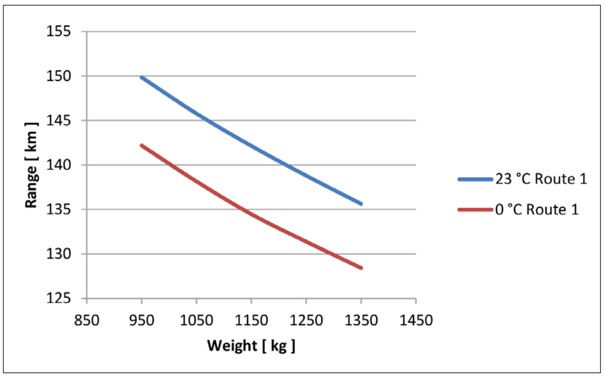


Fig. 9 Impact of low temperatures on the range

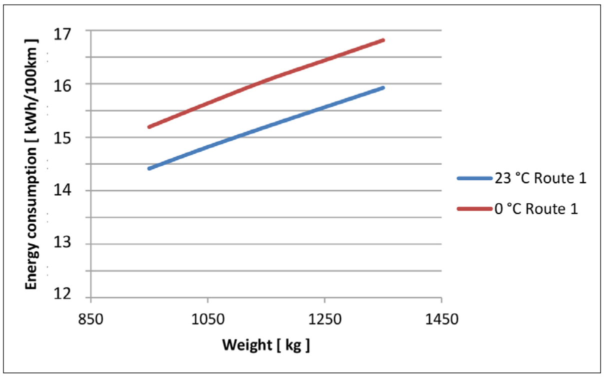


Fig. 10 Impact of low temperatures on energy consumption

5. Conclusion

The simulation results show that for increasing the range of city electric vehicles it is better to focus on decreasing the weight rather than the aerodynamic drag coefficient. Climatic conditions especially low temperatures have significant impact on the range of electric vehicles. When designing electric vehicles it is also beneficial to focus on the systems which can heat the batteries to the appropriate temperature when the battery has low internal resistance. The gained results from simulations, mainly from the simulations of low temperature impact will be used in development of the on-board systems which will inform the driver about the range with better accuracy.

Acknowledgement

This paper is supported by the following project: University Science Park of the University of Zilina (ITMS: 26220220184) supported by the Research&Development Operational Program funded by the European Regional Development Fund and APVV-SUSPP-014-09 – Centrum komponentov dopravnej techniky.

References

[1] FRANKE, T., NEUMAN, I., BUHLER, F., COCORON, P., KREMS, J. F. K.: Experiencing Range in an Electric Vehicle -understanding Psychological Barriers, *Applied Psychology: An International Review*, 61(3), 2012, 368-391.

[2] DEROLLEPOT, R., WEISS, C., KOLLI, Z., FRANKE, T., TRIGUI, R., CHLOND, B., ARMOOGUM, J., STARK, J., KLEMENTSCHITZ, R., BAUMANN, M., PELISSIER, S.: Optimizing Components Size of an Extended Range Electric Vehicle According to the Use Specifications, *Transport Research Arena*, 2014, Paris.

[3] BARTA, D., GALIKOVA, J.: *Wheeled Vehicles - Trolleybuses and Alternative Drives (in Slovak)*. Zilinska univerzita: Zilina, 2013, ISBN 978-80-554-0751-7.

[4] HLAVNA, V., KOVALCIK, A.: *Energy Sources for Automobiles Nonconventional (in Slovak)*, Zilinska univerzita, vol. 1, 2013, ISBN 978-80-554-0788-3.

[5] BRUMERCIK, F., LUKAC, M.: Vehicle Simulation Possibilities, *Logi: Scientific Journal on Transport and Logistics*, vol. 3, 2012, 45-52, ISSN 1804-3216.

[6] GAO, D. W., MI, CH., EMADI, A.: *Modeling and Simulation of Electric and Hybrid Vehicles*, Proc. of the IEEE, 200, 729-745.

[7] NAEHYUCK, CH., DONKYU, B., JEONGMIN, H.: *Power Consumption Characterization, Modeling and Estimation of Electric Vehicles,* 2014 IEEE/ACM Intern. Conference, 2014, 175-182.

[8] MAIA, R., SILVA, M., ARAUJO, R., NUNES, U.: Electric Vehicle Simulator for Energy Consumption Studies in Electric Mobility Systems, *Integrated and Sustainable Transportation System (FISTS),* 2011, ISBN 978-1-4577-0990-6, 227-232.
[9] SCHALTZ, E.: *Electrical Vehicle Design and Modeling,* INTECH, 2011, ISBN 978-953-307-477-1.
[10] EHSANI, M., GAO, Y., EMADI, A.: Electric and Fuel Cell Vehicles, *CRC Press, Florida*, 2010, ISBN 978-1-4200-5398-2.
[11] BARLOW, T. J., LATHAM, S., MCCRAE, I. S., BOULTER, P. G.: A Reference Book of Driving Cycles for Use in the Measurement of Road Vehicle Emissions. 3rd ed. TRL Limited, 2009.
[12] GAJDAC, I., MIKITA, M., KUCERA, L.: EDISON Project of Electric Vehicle at University of Zilina (in Slovak), AI magazine, 2014, 94-96.
[13] JI, Y., WANG, CH. Y.: Heating Strategies for Li-ion Batteries Operated from Subzero Temperatures, Electrochimica Acta, vol. 107, 2013, 664-674.
[14] BULLIS, K.: Electric Vehicles out in the Cold, MIT Technology Review, 2013.
[15] JOHNSON, V., PESARAN, A., SACK, T.: Temperature-Dependent Battery Models for High-Power Lithium-Ion Batteries, 17th Electric Vehicle Symposium Montreal, 2000.

Peter Pechac - Milan Saga - Ardeshir Guran - Leszek Radziszewski *

IMPLEMENTATION OF MEMETIC ALGORITHMS INTO STRUCTURAL OPTIMIZATION

The paper presents an implementation of the memetic algorithm for discrete structural optimization. This algorithm is a combination of genetic algorithm and five local search methods. The proposed memetic algorithm is able to choose the right local search method based on performance which is evaluated in real time during optimization.

Keywords: *Memetic algorithm, structural optimization, pattern search, Nelder-Mead method, nonlinear conjugate gradient method, particle swarm optimization, fully stress design.*

1. Introduction

The term "meme" was coined by evolutionary biologist Richard Dawkins in his book The Selfish Gene where it is described as an equivalent to gene in cultural evolution. Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation [1].

Memetic algorithms represent a group of optimization methods which combine global optimization and local search methods. The global search part is usually based on genetic algorithm or particle swarm optimization. The global optimization part ensures diversity, favors exploration and provides good coverage of the optimization space, while the local search methods favor exploitation and speed up the convergence to local extremes [2]. The choice of local search methods is affected by the problem which is being solved. The no free lunch theorem says that the performance of different methods on different problem classes is not equal [3] and so the choice of the right local optimization method is often left on the user, see - Fig. 1. Problems of engineering mechanics often require computationally expensive evaluation of objective function [4 and 5], so it is crucial to choose the right local search methods. A traditional approach would be that the user selects the local search methods based on his previous experience and observation of their performance on similar problems. This traditional approach requires the user to have this experience and even then it does not guarantee that the best performance is achieved. The memetic algorithm which is proposed in this paper is using multiple local search methods and the algorithm itself is capable of making the decision which method should be used. This decision is based on performance evaluation.

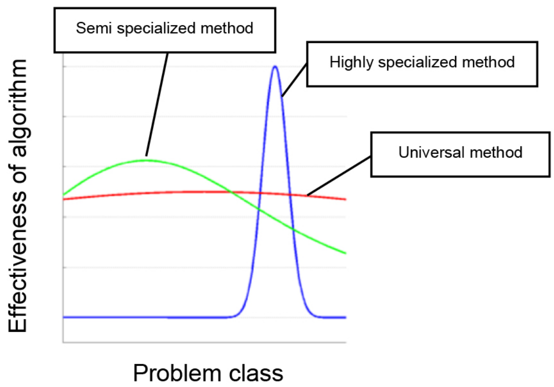


Fig. 1 Effectiveness of algorithm vs. problem class

2. Description of the Memetic Algorithm

The global optimization part of memetic algorithm is based on genetic algorithm. The genetic algorithm is using three genetic operators: selection, crossover and mutation. Rank selection

* [1]Peter Pechac, [1]Milan Saga, [2]Ardeshir Guran, [3]Leszek Radziszewski
[1]Department of Applied Mechanics, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]Institute of Structronics, Ottawa, Canada
[3]Faculty of Mechatronics and Machine Building, Kielce University of Technology, Kielce, Poland
E-mail: saga@uniza.sk

is taking into account both the value of objective function and diversity to prevent premature convergence to local extreme. Uniform crossover ensures that the new offspring inheres equal portion of genes from both parents. The principle of uniform crossover is depicted in Fig. 2. Non-uniform Cauchy mutation performs better than other types of mutation because it has potential to make longer jumps [6]. The Cauchy distribution is depicted in Fig. 3. The high probability of mutation in the beginning helps to improve diversity in population and is approaching zero in the final stages. This probability of mutation is guided by exponential cooling schedule known from simulated annealing which can be seen in Fig. 4. In each generation the best 10% of population is carried to the next generation without changes. This elitist approach saves good solutions from mutation and crossover which could be otherwise lost. However, they are also used as parent chromosomes during selection and crossover. The worst 40% of population is deleted prior to the selection. The new population is assembled from the best 10% which were further improved by local search methods, the offspring generated by crossover and mutation and the rest is generated using "white space search" (WSS) algorithm to complete the full size of population.

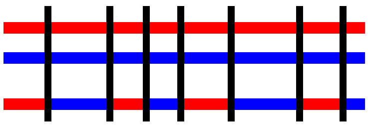


Fig. 2 Uniform crossover

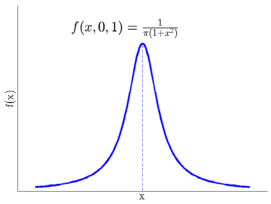


Fig. 3 Cauchy distribution

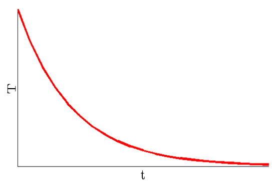


Fig. 4 Exponential cooling schedule

The memetic algorithm has five local search methods at its disposal. These methods include: pattern search method, Nelder-Mead simplex method, Dai-Yuan nonlinear conjugate gradient method with line search, particle swarm optimization and fully stress design method. The sizes of step parameters of the local search methods are guided by linear cooling schedule (see Fig. 5), which is used in simulated annealing. In each generation of genetic algorithm, 10% of the population is improved by local search methods. The memetic algorithm can choose the local search methods based on their performance.

Fig. 5 Linear cooling schedule

In the first iteration all local search methods are tested with equal probability, the performance of methods is then evaluated based on the difference of objective function divided by time spent by the method and averaged for each method. The percentage of candidates improved by local search methods is then divided according to their performance. If the percentage for any method reaches zero it is kept at small non zero value so it has chance to be tested and potentially return to normal use. If the ratio of use of all local search methods reaches zero, the process is restarted and all of them gain the same ratio of use as in the beginning. By always using the best performing method the overall performance is significantly improved. The scheme of the algorithm is in the following Fig. 6.

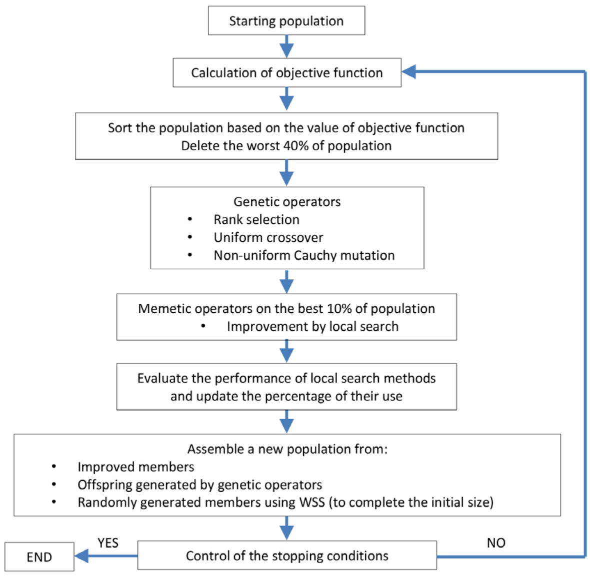


Fig. 6 Scheme of the memetic algorithm

3. Properties of Local Search Methods Used in the Memetic Algorithm

The detailed description of the used local search methods would not fit into the scope of this paper, instead only their main properties will be discussed. The memetic algorithm uses five optimization methods as local search methods: pattern search, Nelder-Mead simplex method, Dai-Yuan nonlinear conjugate gradient method, particle swarm optimization method and fully stress design method. Pattern search method and Nelder-Mead simplex method are deterministic comparative optimization methods, thus they are derivative free and can operate on functions that don't have continuous first derivative (noisy functions). However, they can perform inferior to gradient based methods when used on smooth functions with continuous first derivative. Dai-Yuan nonlinear conjugate gradient method is a fast converging deterministic gradient based method, however as all gradient based methods it can have problems with noisy functions as the numerical calculation of gradient can become problematic. Particle swarm optimization is considered as a stochastic global search method, however its properties can be altered by tuning its parameters (inertia, cognitive and social parameter) to favor exploitation over exploration [7]. Fully stress design method is an engineering approach to structural optimization and despite the fact that its use is limited only to stress constrained problems it is very effective. The fully stress design method can very quickly solve the part of problem which requires minimization of weight with respect to stress constraint, on the other hand, the conventional optimization methods would require much more time to find an equivalent solution. More detailed information on the fully stress design method can be found in [8, 9 and 10]. Each local search method is running only for a few iterations.

4. Numerical Test

The algorithm was tested on a truss structure subjected to stress and displacement constraint. The goal is to minimize price of used material, while satisfying the stress and displacement constraints. Three materials with different price and yield strength were used. The used materials were structural steels S235, S275 and S355. The safety coefficient for stress was k = 3 so the stress limits σ_L for each material were 78.33, 91.66 and 118.33 MPa respectively. The prices p for ton of material were estimated 450, 500 and 600 euro. Density was same for all materials ρ = 7850 kg m^{-3}. The displacement limit was 20 mm. The Young's modulus is the same for all three materials E = 210 000 MPa.

Objective function is calculated as follows

$$f(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m} \rho_i V_i p_i}\,. \tag{1}$$

If the stress constraint is violated, the objective function is penalized as follows

$$f(\boldsymbol{x}) = f(\boldsymbol{x}) + f(\boldsymbol{x})\left(\frac{(\sigma_{\max} - \sigma_L)}{\sigma_L}\right). \tag{2}$$

Similarly, if the displacement constraint is violated, the objective function is penalized as follows

$$f(\boldsymbol{x}) = f(\boldsymbol{x}) + f(\boldsymbol{x})\left(\frac{(u_{\max} - u_L)}{u_L}\right). \tag{3}$$

The structure was divided into 6 optimization groups. The optimization variables were 6 cross-section areas and 10 geometrical parameters. The sizes of cross-section area were divided into 361 discrete values from 500 to 2500 mm^2. For the purpose of keeping the number of optimization variables as low as possible the design variables for cross-sections were combinations of materials and cross-section areas so the total number of available cross-sections was 1083. The geometrical optimization variables were heights of the structure defining its shape. The nodes between the top and bottom of structure were always positioned in the middle of the total height.

The structure was loaded by its own mass and forces of magnitude 15 kN. Boundary conditions and optimization variables are depicted in the following Fig. 7 of the structure using the shape after optimization.

The size of population was set to 1000, maximum number of generations was 500 and the number of iterations of local search in one generation was 2. The resulting axial stresses after optimization are depicted in the following Fig. 8.

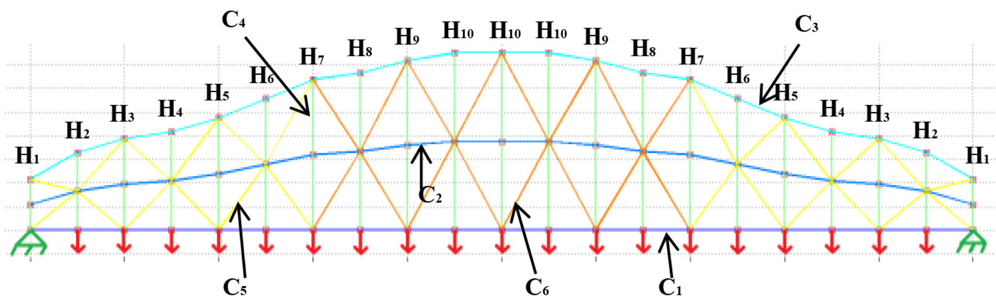


Fig. 7 Boundary conditions (red – loading forces, green - deformation constraints) and optimization variables (c – cross-section, h - height)

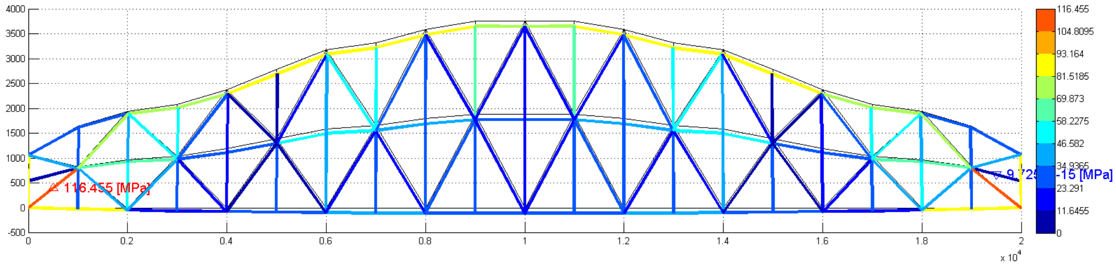


Fig. 8 Axial stress [MPa] after optimization

The history of local search methods use is in the next Fig. 9.

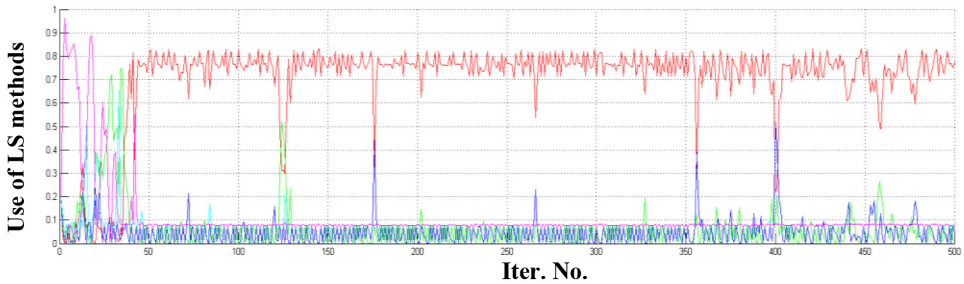


Fig. 9 Use of local search methods vs. iteration step (red - PS, green - NM, blue - NCG, cyan - PSO, magenta - FSD)

The history of objective function can be seen in Fig. 10.

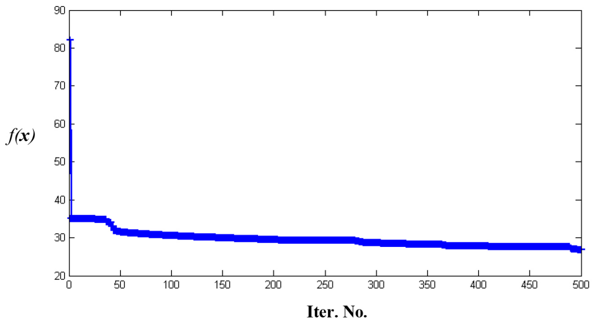


Fig. 10 Objective function vs. iteration step

The value of objective function in optimum found by the algorithm was 26.9 and mass of the structure was 1345.3 kg.

Furthermore, due to stochastic nature of the components which are included in the memetic algorithms, a statistical test was carried out to compare the performance of MA with automated choice of local search method vs. MA with single local search method. We tested it against all five LS methods which have been used in the proposed algorithm.

The size of the population was set to 100, each local search method was allowed to run 2 iteration steps in one generation of GA. The algorithm stopped when it found solution with value of objective function lower than 40. The performance is compared based on the time spent by the algorithm. To provide statistical significance we run each combination 500 times. The results showing percentage of samples vs. computational time can be seen on the following histograms (Figs. 11 - 16).

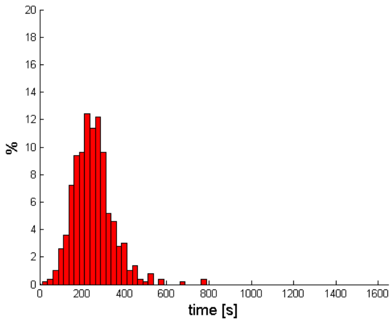


Fig. 11 GA + PS

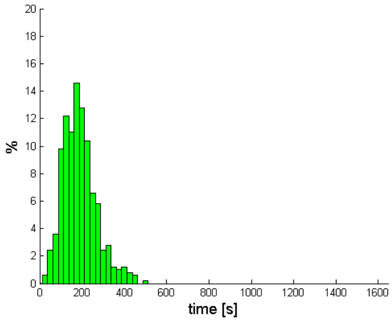


Fig. 12 GA + NM

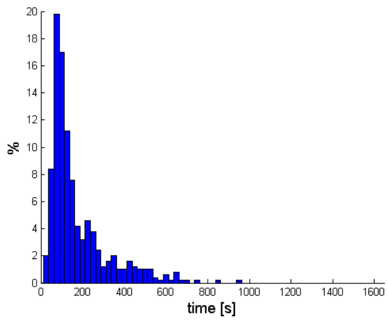


Fig. 13 GA + NCG

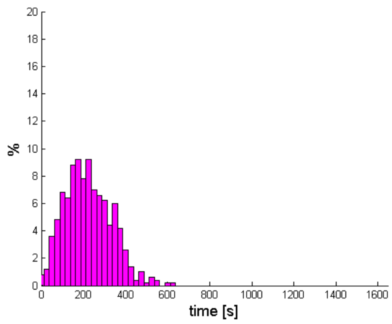


Fig. 15 GA + FSD

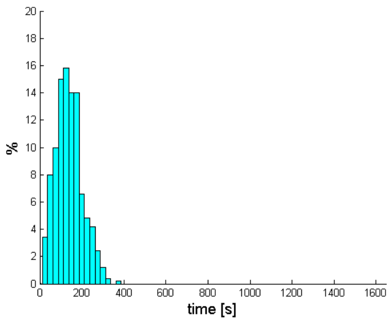


Fig. 14 GA + PSO

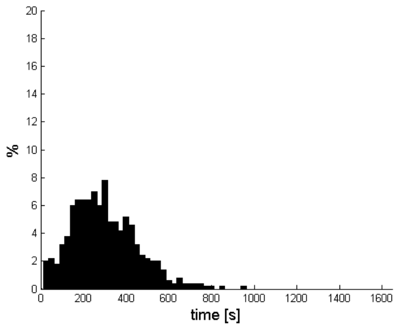


Fig. 16 GA + [PS, NM, NCG, PSO, FSD]

In Fig. 9 you can see that the algorithm first selects FSD method, then Nelder- Mead method and finally the pattern search method. Nonlinear conjugate gradient method and particle swarm optimization method were used only scarcely. Pattern search and Nelder-Mead method are both deterministic and gradient free. Comparison of combinations of GA and single local search methods shows that the Nelder-Mead method is better than simple pattern search, however the automatic choice of local search methods chose pattern search over Nelder-Mead.

5. Conclusion

The best performing combinations were GA+PSO and GA+NCG so they can be considered as the best choices for solution of similar problems. The automatized choice of local search method has not improved performance on this particular test problem as we expected. Further research and testing on multiple test problems is required to definitely confirm whether the proposed idea of automatized choice of local search method improves the performance or not. Its performance could be affected by repeated use of deterministic local search methods on the same solution which did not improve the solution but rather wasted the computational time. The proposed method could be improved by keeping the track of operations carried out on each solution so the algorithm would not use the same deterministic method repeatedly on the same solution and also by using meta-optimization to adjust the parameters of local search methods.

Acknowledgements

This work has been supported by grant project VEGA No. 1/0234/13.

References

[1] DAWKINS, R.: *The Selfish Gene*. New York: Oxford University Press, 1976.

[2] MOSCATO, P., COTTA, C.: *A Gentle Introduction to Memetic Algorithms,* Handbook of Metaheuristics, 105-144, Springer US, 2003.

[3] WOLPERT, D., MACREADY, W.: No Free Lunch Theorems for Optimization, *IEEE Trans. Evol. Comput.*, vol. 1, pp. 67 -82 1997.

[4] A. CZAN, E. TILLOVA, J. SEMCER, J. PILC: Surface and Subsurface Residual Stresses after Machining and their Analysis by x-ray Diffraction. *Communications - Scientific Letters of University of Zilina,* vol. 15, No. 2, 2013, 69-76, ISSN 1335-4205.

[5] ZMINDAK, M., NOVAK, P., PASTOREK, P.: *Finite Element Analysis of Delamination in Steel-reinforced Concrete Reinforced by Fiber Composite (in Slovak),* Experimental and Computational Methods in Engineering, Litomerice, 2015.
[6] YAO, X., LIU, Y., LIN, G.: Evolutionary Programming Made Daster, *IEEE Transaction on Evolutionary Computation*, vol. 3, No. 2, 82-102, 1999.
[7] CERVANTES, A.: *Constriction Factors and Parameters*, 2005, Retrieved from http://tracer.uc3m.es/tws/pso/parameters.html.
[8] SAGA, M., PECHAC, P., VASKO, M.: FSD and its Modifications for Shell Finite Elements. *American J. of Mechanical Engineering,* 1(7), 2013, 427-433.
[9] SAGA, M., DUDINSKY, M., PECHAC, P.: Optimization of thin Shell Structures using FSD Algorithms, 2012, *Communications - Scientific Letters of the University of Zilina,* 14 (3), 32-38.
[10] SAPIETOVA, A., DEKYS, V., VASKO, M.: A Numerical Modeling Rotating Machine Having Unbalance and the Measuring of its Dynamical Properties, *Metalurgija/Metallurgy*, vol. 492010, No. 2, 503-507.

Marek Bruna - Dana Bolibruchova - Petr Prochazka *

NUMERICAL SIMULATION OF MELT FILTRATION PROCESS

The paper aims to comprehensively elaborate the impact of ceramic filter design on filling parameters during the casting of aluminium alloys. Experimental work focuses mainly on the evaluation of velocity, filling time and turbulences in the close vicinity of the filter by numerical simulation. The whole experiment was divided into two stages: in the first stage was evaluated set of three filters with different channels cross--section geometry, and in the second stage was evaluated set composed of four filters with different number of channels and channels diameter.

Keywords: *Filtration, numerical simulation, aluminium alloys.*

1. Introduction

Filling time, flow turbulence and melt velocity are considered as the most important factors during filtration of aluminium alloys, and therefore must be properly analysed in real conditions, as well as with the help of modern simulation systems. Numerical simulation represents a progressive tool that allows observation of processes which are difficult to analyse during real conditions (for example turbulences). For that reason we decided to use professional simulation software ProCAST, which helped us to clarify and analyse the phenomena taking place in the close vicinity of the filter [1].

2. First stage of the experiment

In the first stage of the experiment, three different filters were chosen (with direct channels and with conical channels); geometry of each filter is shown in Fig. 1.

Figure 2 shows the proposed gating system, red colour represents the location of filters. Simple cuboid was chosen as the casting.

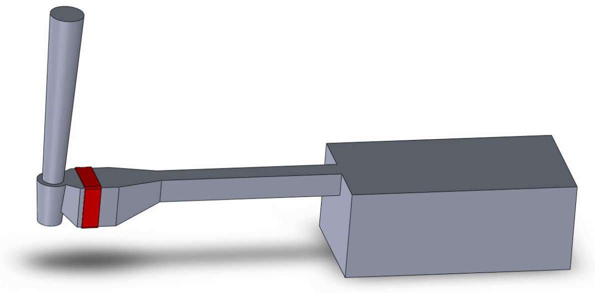


Fig. 2 The gating system shape with the filter location

Preparation of the necessary CAD geometries was followed by FEM meshing in ProCAST software. We used an alternative (more complex) approach instead of the standard function to analyse the melt flow. Individual filters were imported to the

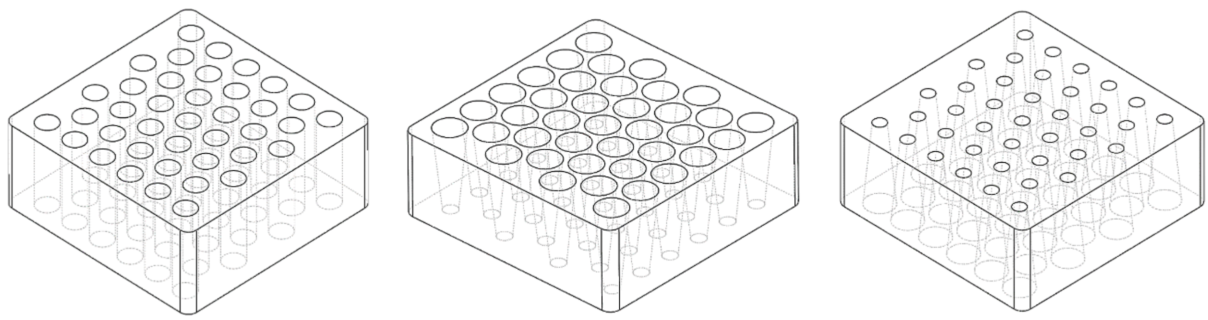


Fig. 1 CAD models of analysed filters

* [1]**Marek Bruna,** [1]**Dana Bolibruchova,** [2]**Petr Prochazka**
[1]Department of Technological Engineering, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]Keramtech, s.r.o., Zacler, Czech Republic
E-mail: rapture.marko@gmail.com

simulation software including the precise channel geometry (based on the CAD files). This fact has a significant impact on the complexity of the entire process, especially on the creation of a 2D and 3D mesh, which was ultimately reflected in the time needed for calculation (approximately 200 computing hours per one variant). However, this method of simulation has great advantages compared to the conventional methods. In the conventional method the software ignores the precise geometry of individual filter channels (it only takes into account its flow capacity), and, therefore, it is impossible to accurately observe the filter effect on the melt flow [2].

Prior to launching the computation itself we had to set the input parameters and boundary conditions. The purpose of experimentation was not a specific case, but mutual comparison of the properties of individual filters in general terms. Therefore, the input data were not based on real-condition values, but were set as a model situation, constant for all three variants, which enabled their mutual comparison. To perform simulation was chosen AlSi7Mg0.3 aluminium alloy poured at 720°C by flow rate 1 kg/s into green sand mould (room temperature 20 °C).

2.1. Results

2.1.1. Filling time

Based on the simulations we achieved the following results:

VARIANT A - 3.3051 s
VARIANT B - 3.8414 s
VARIANT B1 - 3.3256 s

The shortest filling time was achieved in Variant A (filter with direct channels), almost the same time was achieved by Variant B1 (filter with narrower conical channels on the filter inflow side), and an observable increase (by approximately 16 %) occurred in Variant B. Matter of deciding which filling time is optional is complicated because fast filling leads to turbulence-related defects (such as mould erosion, air aspiration, and inclusions). On the other hand, slow filling may cause defects related to premature solidification (such as cold shuts and misruns). In general terms, filling time depends on the cast metal, casting weight, and minimum wall thickness. Achieved results show a way for filter manufacturer and purchasers to find best option while deciding which filter to apply for their specific conditions.

2.1.2. Turbulent energy

In numerical simulations, simplified models are used in addressing turbulent flow due to the complex, and still not fully understood, physical nature of turbulence. Numerical simulation of turbulent flow uses three theoretically different approaches arising from simplified modifications of the flow describing equations.

Currently, three methods are used to solve Navier–Stokes equations that describe laminar and turbulent flow modes. They express the laws of conservation of mass, momentum, heat or other scalar variables. Direct numerical simulation (DNS) solves equations in space and time to get an accurate time course of the observed variable.

The Reynolds-averaged Navier–Stokes (RANS) method gives the variable's mean value in time. A combination of both methods is large eddy simulation (LES), which is used to simulate large eddies precisely (accurately) and solve small eddies by Reynolds equations. 99% of the models used in technical practice are those based on the RANS method. This method is also used by ProCAST simulation software.

It should be also emphasised that, when evaluating the flow of metal through the filter, simulation software does not consider the gradual reduction of the filter flow capacity due to channels blockage by inclusions. Priming (time needed to fill the filter channels with melt) is followed by a unifom flow of melt until the entire gating system and casting are filled [3].

Variant A

Figure 3 shows the turbulence intensity at cross-section of the gating system - time 0.6 sec. The colour scale expresses the melt turbulent energy level, where purple means zero kinetic turbulent energy and with graduated shades it passes to red, which represents 350 $cm^2.s^{-2}$. Larger turbulence occurs on the outflow side of the filter in the lower part than in the upper part. However, we will be able to draw specific conclusions in terms of turbulence intensity only upon comparison with the other variants.

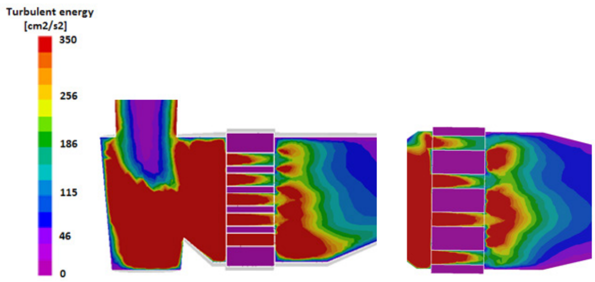


Fig. 3 Turbulence intensity - side view and top view (variant A, time 0.6 sec)

Variant B

The filling in Variant B is of a different nature compared to variant A. Turbulence on the outflow side of the filter is significantly smaller than in variant A. Figure 4 shows the turbulence intensity – cross-section of the gating system at the time 0.6 sec. Results will be described in conclusion.

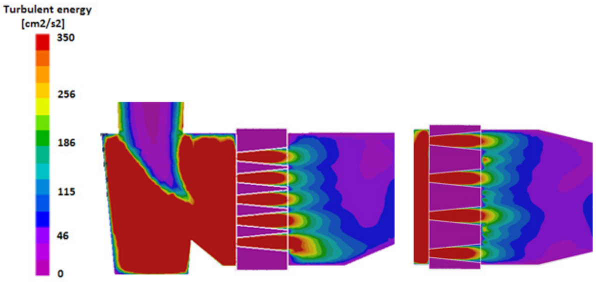


Fig. 4 Turbulence intensity – side view and top view (variant B, time 0.6 sec)

Variant B1

Figure 5 shows the turbulence intensity - cross-section of the gating system at the time 0.6 sec. Variant B1 shows the largest turbulence intensity on the outflow side compared to the other variants (the largest share of red).

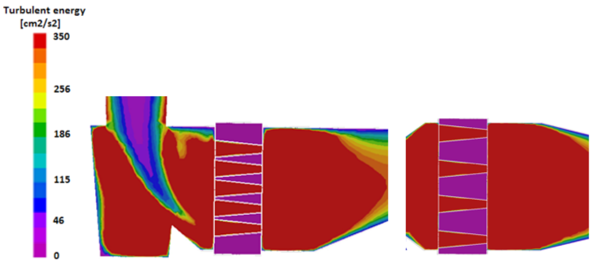


Fig. 5 Turbulence intensity - side view and top view (variant B1, time 0.6 sec)

2.2. Evaluation of the first stage of experiments

In terms of the mould cavity filling time we can conclude that variant A is an alternative with the shortest ***filling time*** (3.3051 sec), almost the same time was achieved by variant B1 (3.3256 sec), and an observable increase by approximately 16 % (3.8414 sec) occurred in variant B. Which variant is optimal is hard to say, it is dependent on the specific case. Main goal was to show possibilities of simulation software regarding filtration.

In terms of ***flow optimisation*** and the associated turbulences, variant B can be considered as the optimal alternative based on the results criterions. Variant B1 features the highest ***velocity*** on the outflow side of the filter, uneven flow and large turbulences. These turbulences are probably caused by nozzle-like shape of channels, which supports increase of velocity on the outflow side of the filter. These phenomena can be evaluated as negative effect leading to higher melt reoxidation and possibility of porosity nucleation during solidification [3 and 4].

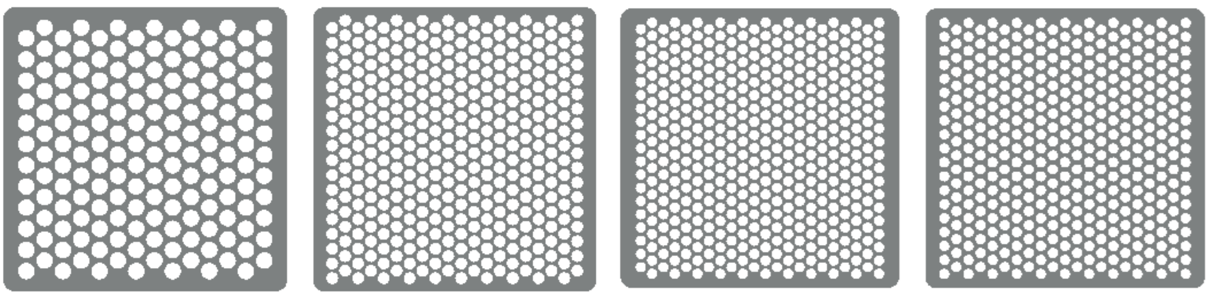
3. Second stage of the experiment

During the second stage of the experiment were used 4 different ceramic filters. Models of the filters used and their identification are shown in Fig. 6. The filter parameters are shown in Table 1. The objective of the second stage was to evaluate turbulence activity and velocity of the melt in the filter close vicinity (while flow area is changing due to amount of channels and their diameter).

Parameters of the filters Table 1

	Filter 0810	*Filter 0812*	*Filter 0735*	*Filter 0825*
Channel diameter [mm]	Ø 3.2	Ø 2.3	Ø 2	Ø 2
Flow area [mm^2]	1351	1495	1348	1223

In order to be able to implement the above-mentioned methodology of filter properties investigation, it was first and foremost necessary to design and calculate a gating system. Design of the gating system required determination of the cross sections of individual parts of the gating system and determination of the controlling cross section. In the first stage we calculated the flow cross section (S_f) of the ceramic filters used based on the sum of



a) Filter 0810 *b) Filter 0812* *c) Filter 0735* *d) Filter 0825*

Fig. 6 Experimentally tested filters

areas of individual channels. Subsequently, we determined the controlling cross section S_r. In the experiments we used filters for aluminium alloy casting, therefore, we determined the resulting ratio $S_f : S_r$ in the interval 4 to 8. Scheme of measuring system is shown in Fig. 7. Proposed measuring system has not been used only for simulation purposes and was constructed also for real conditions experiments, but those will not be evaluated in this article.

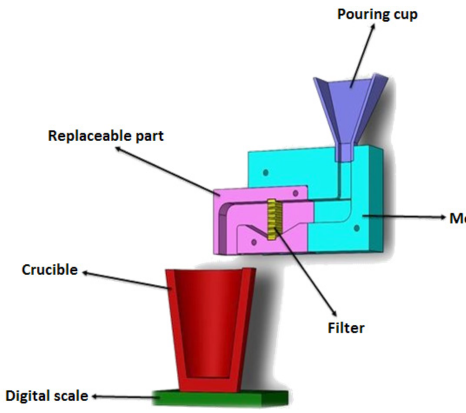


Fig. 7 Scheme of the measuring system

3.1. Melt velocity

Filter impact on the melt velocity in the section plane leading through the middle of the gating system (Fig. 8) was investigated. Velocity development was examined at the time **0.65**

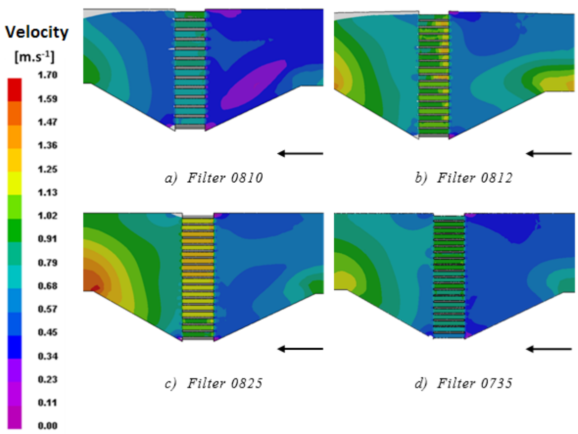


Fig. 8 Melt velocity - time 0.65 sec.

sec, when the gating system was almost filled. The colour scale expresses the intensity level of the melt velocity similar to the previous case. Simulations show that the greatest deceleration of the melt occurs by introducing a 0810 filter, which consists of channels with a largest diameter among the investigated filters. This phenomenon is attributable to the fact that narrow filter channels may act as nozzles on the melt. Comparison of the 0735 and 0825 filter simulations, which have the same size of channels (Ø 2 mm), clearly shows that melt velocity is affected also by the size of the flow area. Flow area affects the size of metallostatic pressure in the way that the larger the flow area, the more easily depressurisation occurs. At the same time the melt velocity on the outflow side of the filter is also influenced to a great extent by the primary shock wave on the filter [4].

3.2. Turbulent energy

The colour scale expresses the melt turbulent energy level, where purple corresponds to zero kinetic turbulent energy and with graduated shades it passes to red, which represents 200 $cm^2.s^{-2}$ (Fig. 9).

When evaluating turbulence, or the melt kinetic turbulent energy in the separation plane at the time 0.65 sec, optimal results were achieved by a 0812 filter, which exhibits on the outflow side of the filter the lowest level of turbulent flow in comparison to other filters (Fig. 9b). Although it has been proved that flowing through narrower filter channels accelerates the melt movement, the channel hydraulic diameter is so small that the Reynolds number reaches lower values than those for a 0810 filter having the largest channel diameter.

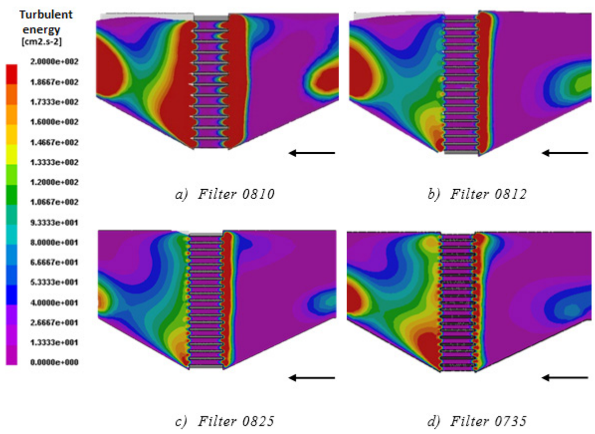


Fig. 9 Melt turbulence - time 0.65 sec.

4. Conclusion

The objective of this paper was to utilise numerical simulation for analysis of important parameters when casting aluminium alloys through ceramic filters. Partial objectives following the individual stages of experiments have confirmed that simulation is a useful tool even for this field of foundry industry. Nowadays, it is obvious that simulations are used not only for theoretical research, but they can be fully implemented even into practical tasks with tangible results, and can assist in choosing a suitable alternative from the evaluated set.

References

[1] CAMBELL, J.: *Casting Practice - the 10 Rules of Castings,* Oxford: Butterworth-Heinemann, 2010., 224 p. ISBN 07506-4791–4.

[2] PASTIRCAK, R., SLADEK, A.: Foundry Processes and Computer Simulation. *Materials Engineering,* No. 10, pp. 119-122. 2003, ISSN 1335-0803.

[3] BRUNA, M.: Reoxidation Processes Analysis during Filtration of Aluminium Alloys. *Technolog,* vol. 5, No. 1, pp. 31-36. 2013, ISSN 1337-8996.

[4] PASTIRCAK, R.: Effect of Low Pressure Application during Solidification on Microstructure of AlSi Alloys, *Manufacturing Technology,* vol. 14, No. 3. 2014, pp. 397-402, ISSN 1213-2489.

Radoslav Konar - Marek Patek - Michal Sventek *

NUMERICAL SIMULATION OF RESIDUAL STRESSES AND DISTORTIONS OF T-JOINT WELDING FOR BRIDGE CONSTRUCTION APPLICATION

Prediction of the distortions and residual stresses of the welded steel structures is still very important to reduction of the costs, mainly during the optimization of welding process. If there is necessity of the higher welding current to ensure weld penetration, distortions might be influenced by appropriate welding sequence. During optimization process, numerical analysis of welding process, based on finite element method, can be used to prediction of distortions at different welding sequences without necessity of preparing high amount of experimental samples. Numerical simulation of T-joint welding process is presented in this article, together with verification of the analysis by thermocouple measurements and contactless measurement of distortions.

Keywords: *Distortions, finite element method, residual stress, SYSWELD, welding simulation.*

1. Introduction

Welding distortions in a structure or component can result in the degradation of its dimensional tolerances followed by costly rectifications and possible delays in the production line. Additionally, welding residual stresses can influence the fatigue performance and buckling strength of the product. Therefore, understanding and controlling the formation of weld induced distortions and residual stresses are of the utmost importance in the manufacturing industry [1].

During the last three decades, with the evolution of computing capabilities, the finite element (FE) method has proven itself as an alternative and acceptable tool for prediction of welding residual stresses and distortions. The FE simulations for the welding process can be divided into three categories, though they all have a common motive which is assessment of welding residual stresses and distortions. The three categories are as follows:

1. The simulations in which the influence of welding process specifications (heat input, plate thickness, weld sequence etc.) are investigated.
2. The simulations in which a novel numerical strategy is used to get the results with acceptable accuracy but with less computational time.
3. The simulations in which FE input parameters (mesh size, element type, heat source model, material properties and material models etc.) are manipulated to study their influence on residual stresses and distortions [1].

Application of the first category allows optimization of the manufacturing process before assembly of the final construction. However, better FE analysis is performed if experimental samples are prepared to verification. Samples can have lower dimensions that are still significant advantage of the numerical analysis. In the article, application of the first FE category to prediction of angular deformations and residual stresses during T-joint welding is presented. Results are verified by thermocouple measurements and contactless deformation measurement.

In principle, a finite element simulation of the welding process consists of two main parts: thermal analysis and mechanical stress analysis. In thermal analysis, the temperature field is determined as a function of time for each integration point. This temperature time-history is used as an input into the thermal stress analysis. Herein, the thermal solution can be sequentially or fully coupled with the mechanical solution of the structure. Because the rate of heat generation due to mechanical dissipation energy can be neglected in the heat transfer analysis, a sequentially coupled thermal-stress analysis is commonly applied for the simulation of a welding process in which a thermal analysis is followed by a stress analysis [2 - 6].

* [1]**Radoslav Konar,** [1]**Marek Patek,** [2]**Michal Sventek**
[1]Department of Technological Engineering, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]MONT IRP, s.r.o., Zilina, Slovakia
E-mail: radoslav.konar@fstroj.uniza.sk

2. Experimental measurements and results

Experimental works consisted of the measurements of thermal cycles by the thermocouples and contactless distortions measurement by the TRITOP system.

Each measurement was performed on the T-joint made of S355J2+N steel plates with *12 mm* thickness welded by the fillet weld from both sides of the web. Welds were prepared by automatic MAG welding process by FastMIG Basic KM500 device at welding parameters according to Table 1. Welding wire INEFIL (G 42 4M G3Si1 in terms of EN ISO 14341-A) with *Ø1.2* mm diameter was used as a filler material. During the process, shielded gas consisting of 18 % CO_2 + 82 % Ar was used.

Welding parameters Table 1

Weld pass	Current I_z [A]	Voltage U_z [V]	Welding speed v_z [mm.s^{-1}]	Arc power Q [kJ.mm^{-1}]
1	421.7	41.4	5.86	2.38
2	394.8	39.3	5.88	2.11

2.1. Measurements of thermal cycles

Samples with length of *750 mm* were used during thermocouples measurements. Base plate width was *200 mm* and web was *100 mm* high. Thermal cycles were measured by 9 thermocouples type K, placed in the base plate from the bottom side in three planes (3 thermocouples in each plane). Thermocouples were in different depth and distance of the center of the sample to better description of temperature distribution, mainly in heat affected zone (HAZ) of the welds. Position of the thermocouples TC03, TC04 and TC 05 used for the verification of the thermal behavior during experimental measurement and FE analysis is shown in Fig. 1. Temperatures were during welding measured with the frequency of *10 Hz*.

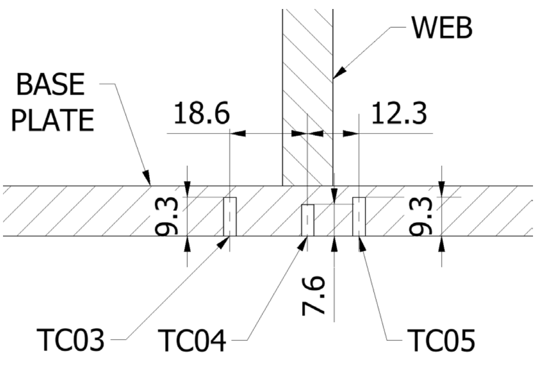


Fig. 1 Position of the thermocouples TC03, TC04 and TC05

Characteristic curve of the thermal cycles measured by thermocouples T03 and T05 placed in the same depth (*9.3 mm*) but in different distance of the base plate vertical axis (T03 – *18.6 mm* in 1st weld HAZ, T05 – *12.3 mm* in 2nd weld HAZ) is shown in Fig. 2. Position of the thermocouple TC04 was very close to the weld penetration (see Fig. 8b), which led to destruction of this thermocouple during the welding process. Thermal cycle thus could not be used to verification of the results.

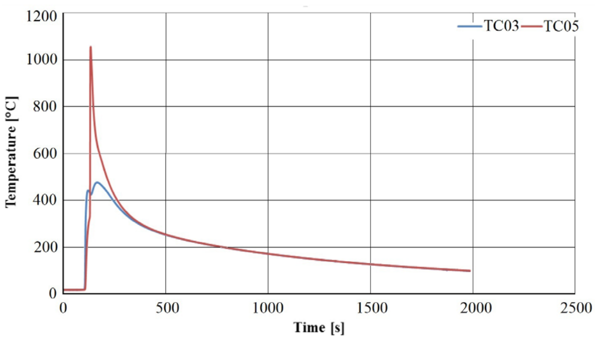


Fig. 2 Characteristic curve of temperature distribution in HAZ of the weld joints

2.2. Distortions measurements

Distortions of the T-joint were determined by the displacement of the points placed on the sample by optical measuring system TRITOP. The TRITOP system is an industrial optical measuring system that acquires 3D coordinates of selected points of a measuring object without contact. In order to carry out photogrammetric measurements, the object needs to be recorded from various directions with a high-end digital camera. It is not necessary to maintain exact camera positions. Some coded measuring markers are applied to the fixture or directly to the object and are used for the automatic evaluation process. Scale bars are positioned next to the object. Their dimensions are transferred to the measurement to provide scale and to observe the measuring accuracy. Surface points to be measured are marked with removable markers or adapters. Photographs prepared by digital camera are after creating processed in the appropriate software.

Experimental samples were during the distortion measurement welded by two welding sequences (Fig. 3). First sequence was consecutive welding with one welding tractor. One fillet weld was prepared at first, and after the tractor was transferred to other side of web, second weld was made. Second sample was welded by two welding tractors at once. The second welding source started to work with technological delay of *25 s* after the first one.

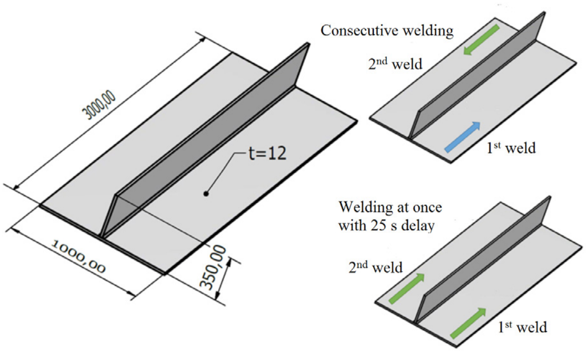


Fig. 3 Dimensions of the sample and welding sequences during distortions measurements

Distortions of the T-joint were processed in the TRITOP software and each point's coordinates were transferred to graphic output of sample distortions. Example of the distortions for second welding sequence (welding at once) is shown in Fig. 4.

Coordinates of the points served for the calculation of the angular distortion by the law of cosines and difference of the computed angles between the web and base plate on the both sides of the T-joint. Angular distortion was computed in the 5 cross-sections, where the 1st one was situated in the front of the sample (beginning of the 1st weld) and the others were placed in intervals *150 mm* long. Calculated angular distortions are shown in Table 2. Positive value represents decreasing (closing) of the angle between the base plate and web, and negative value increasing (opening) of the angle.

3. Numerical simulation of welding

FE analysis of welding was prepared to prediction of temperature distribution during welding, distortions and residual stresses after welding.

3.1. Input parameters

Due to relatively large dimensions of the experimental samples, simplification of the structure by 2D geometrical model was applied, which represents cross-section of the T-joint. Simplification not only allowed shorter computational time, but also finer meshing in the weld area in relatively low number of the elements. It increases precision of the welding analysis.

Two geometrical models were prepared during analysis to ensure possibility of comparison of the experimental and numerical results (different dimensions of the base plate and web were used in experiments). Dimensions of the model to thermal analysis corresponded with sample from thermocouples

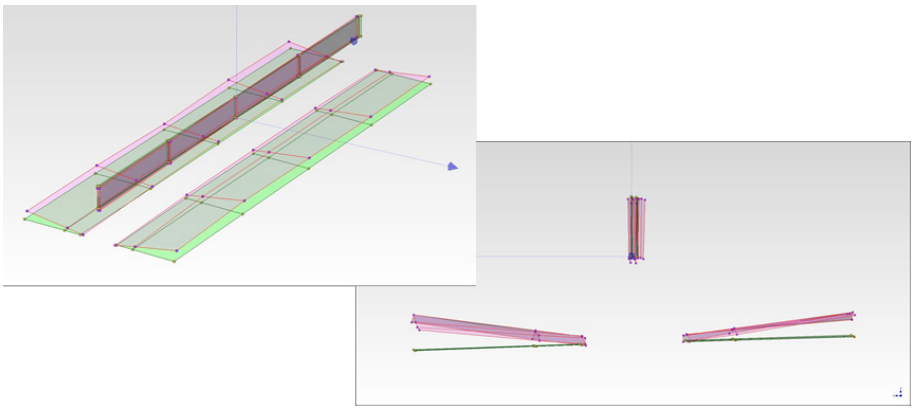


Fig. 4 Distortions of the sample after welding with technological delay 25 s (green – before welding; red – after welding)

Angular deformation obtained by optical measurement — Table 2

Cross-section	Consecutive welding		Welding at once	
	Right side (1st weld)	**Left side (2nd weld)**	**Right side (1st weld)**	**Left side (2nd weld)**
1	0.7408°	-0.8958°	1.0180°	-0.7137°
2	1.3464°	-1.1177°	0.6101°	-1.0746°
3	1.4745°	-1.2073°	0.4941°	-0.7797°
4	1.5448°	-1.3362°	0.7777°	-0.6003°
5	0.9160°	-0.9167°	0.7848°	-0.4983°
Average	1.2045°	-1.0947°	0.7369°	-0.7333°

measurement and mechanical analysis model had the same dimensions as distortion measurement sample.

Geometrical model was created in AutoCAD program and consequently meshed in Visual-Mesh software. Thermal source movement was defined by trajectory path and so-called reference that serve to definition of the local coordinate system. Both curves (lines) were formed by 1D elements. Along with weld parts and curves of the thermal source movement, a plane (curve) of heat exchange has to be defined. In the prepared mesh, the curve formed as a contour of the 2D model served for this purpose. Final model of the T-joint was formed by *2784* 2D elements and *802* 1D elements (Fig. 5).

Mechanical properties of the FE model were defined by material database of SYSWELD software for S355J2G3 steel. Since material used in experimental measurements had the same strength grade as material in software database, it was not necessary to modify the database.

Welding process is in numerical simulations defined by heat source model. Arc welding processes (MMA, MIG/MAG, TIG, etc.) are mostly defined by double-ellipsoidal heat source (so-called Goldak's source) [1, 2 and 4]. Geometry and heat source parameters of the double-ellipsoid that was defined according to macrograph of the experimental sample and real heat input are shown in Fig. 6 and Table 3.

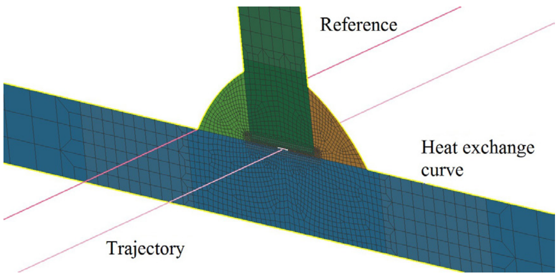


Fig. 5 FE mesh

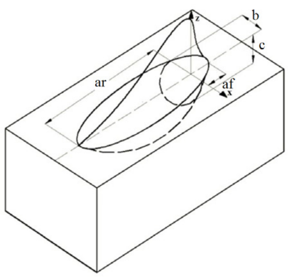


Fig. 6 Geometry of double-ellipsoidal heat source model [1,2,4]

Welding speed of the heat sources corresponded with speed during experimental measurements (shown in Table 1).

During welding, a considerable amount of heat is lost by radiation and convection, while radiating loses are dominating at higher temperatures near the weld zone and convection loses are significant at low temperatures away from the weld line. In this analysis, radiation is modeled by standard Stefan-Boltzman relation, in which the main parameter is radiation emissivity ε [7]. Constant value of the emissivity was maintained with the value of ε = 0.8 that corresponds value with oxidized surfaces of materials [8]. Heat loss by convection depends on convective heat transfer coefficient, or on heat flux q [7]. In performed analysis, constant heat flux was defined with the value of *8* $W.m^{-2}$. Applied heat flux value was controlled according to cooling curve measured by thermocouples.

Mechanical analysis requires except of the material properties, heat source definition also definition of the clamping conditions. In FE analysis, clamping was defined by removing of the freedom degrees for the two points according to Fig. 7.

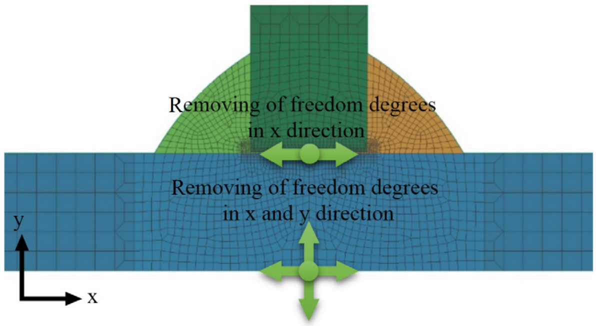


Fig. 7 Clamping conditions of FE analysis

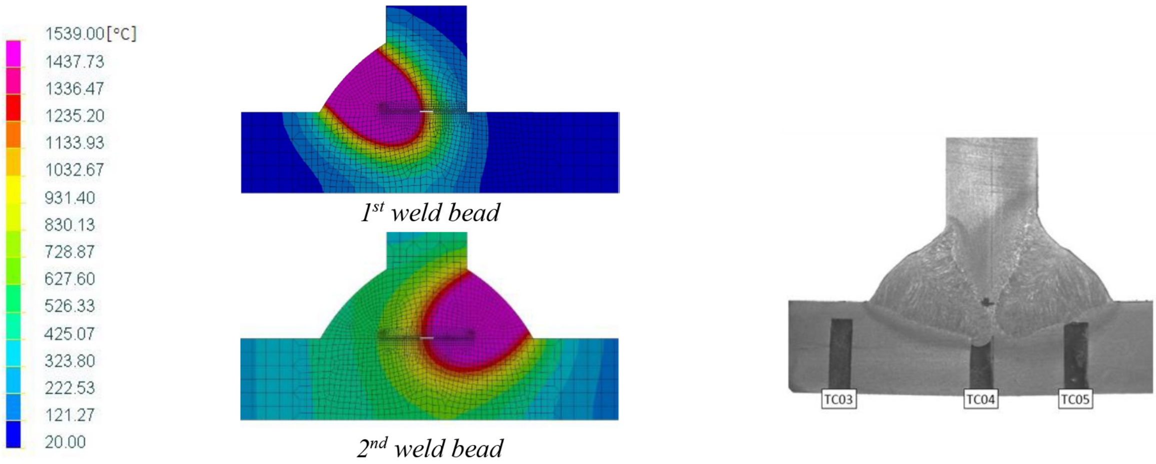
3.2. Thermal analysis results

Results of the thermal analysis are thermal fields and thermal cycles during welding. Comparison of the thermal fields and weld joint macrograph from thermocouples measurement is shown in Fig. 8.

Temperature distribution from the FE analysis of welding corresponds with the liquid metal penetration during welding from the macrograph. Slight differences might be observed in the upper part of the welds caused by the shape of the heat source model that does not allow widening of the thermal field in the

Heat source parameters — Table 3

Bead	Q_f $[W.mm^{-3}]$	Q_r $[W.mm^{-3}]$	a_f $[mm]$	a_r $[mm]$	b $[mm]$	c $[mm]$	x_0 $[mm]$	a_y $[°]$	y_0 $[mm]$	z_0 $[mm]$
1.	18.62	14	2	4	5.5	9.4	8	-49	0	-2.2
2.	12.635	9.5	2	4	7.5	10	-8	56	0	-2.8



a) thermal fields *b) weld macrograph*

Fig. 8 Thermal fields of the first and second weld bead in comparison with macrograph of weld joint after consecutive welding

Parameters of thermal cycles after experimental measurements and FE analysis — Table 4

Thermocouple (FE node)	Exp. / Sim.	T_{max} [°C]	$t_{8/5}$ [s]	r_{300} [°C.s^{-1}]	t_{100} [s]
TC03	Exp.	476	-	0.52	1798
	Sim.	462	-	0.85	1804
TC05	Exp.	1055	64.8	0.54	1842
	Sim.	1297	40.5	0.86	1837

outer part of the weld. Thermal cycles were obtained in the places of corresponding with the thermocouples position, from which the characteristic parameters of the cycle (maximal temperature - T_{max}, cooling time from 800 °C to 500°C - $t_{8/5}$, cooling rate in 300 °C - r_{300} and cooling time from T_{max} to 100 °C - t_{100}) were defined and compared with those obtained by experimental measurements (Table 4).

3.3. Mechanical analysis results

Values of the angular distortions were computed from the displacements of the nodes placed on the edges of the planes and with application of the law of cosines, similarly to experimental measurements. Results of the angular distortions are shown in Table 5.

4. Discussion

Temperature cycles obtained by numerical simulation achieved high agreement with the experimentally determined curves. The highest agreement was at the cooling time from the maximum temperature to *100 °C*, where the differences were lower than *1 %*. High match was also in the maximum temperature of the cycle, but with increasing temperature the agreement decreased. It can be caused by even slight differences in actual and computational thermal conductivity of the material, but also by imperfect coupling of the thermocouple to the material or the position of the thermocouple in the hole (ceramic cover of thermocouple in the hole acts as a place with different thermal properties as welded material). Higher differences were found in the parameters $t_{8/5}$, and r_{300} that are significantly influenced by real heat losses from the material. Only constant values of thermal emissivity and heat flux was applied, but some authors recommend temperature dependent parameters [9]. Simplification of the heat transfer process might result to obtained

Angular distortions obtained by numerical analysis — Table 5

Consecutive welding		Welding at once	
Right side (1st weld)	Left side (2nd weld)	Right side (1st weld)	Left side (2nd weld)
1.285°	-1.072°	0.548°	-0.792°

differences of cooling parameters. High agreement of maximum temperature of the cycle and cooling times along with the shape of temperature fields corresponding with the macrostructure are still sufficient conditions for the prediction possibility of residual stresses and distortions.

In the T-joints welding an angular distortion has an important role that is in addition to the welding parameters also related to the welding sequence. Welding parameters have to be in some cases higher to ensure weld penetration. In this case, it is possible to reduce deformation of the plates by application of different welding sequences, to the choice of which it is also possible to use numerical simulation. In the experimental measurements and FE simulations were used two welding sequences with the same welding parameters. Lower distortions were experimentally measured and also numerically predicted in the welding at once with a delay of *25 s*. Results of the simulation were in very high agreement with the experimental results for both welding sequences.

5. Conclusions

Results obtained by the numerical simulation of the T-joint with subsequent verification by experimental measurements might be concluded as follows:

1) Numerical simulation of welding offers a possibility of high agreement of the thermal cycles and parameters T_{max} and t_{100} with experiments without a necessity of material database modification according to real chemical composition and without temperature dependent thermal emissivity and heat flux if steel with corresponding strength grade is used.
2) Angular distortion obtained by FE analysis achieves high agreement with experimental results in the application of 2D FE model.
3) Experimental and simulated results showed lower angular distortion of the T-joint in the welding at once with the *25 s* delay compared to consecutive welding by one welding source around the sample.

Numerical analysis with the parameters used in this article can be further applied to distortion prediction of the steel structures with larger dimension and at different welding sequences.

Acknowledgement

This work has been supported by Scientific Grant Agency of Ministry of Education of the Slovak Republic, grants VEGA 1/0610/12 and KEGA 034ŽU-4/2015. Authors acknowledge the grant agency for support.

References

[1] BHATTI, A. et al.: Influence of Thermo-mechanical Material Properties of Different Steel Grades on Welding Residual Stresses and Angular Distortion. *Materials and Design*, 65, 2015, pp. 878-889.
[2] PERIC, M. et al.: Numerical Analysis and Experimental Investigation of Welding Residual Stresses and Distortions in a T-joint Fillet Weld. *Materials and Design*, 53, 2014, pp. 1052-1063.
[3] DOPJERA, D., MICIAN, M.: The Detection of Artificially Made Defects in Welded Joint with Ultrasonic Defectoscopy Phased Array. *Manufacturing Technology*, vol. 14, No. 1, 2014, pp. 12-17.
[4] NOVAK, P., MESKO, J., ZMINDAK, M.: Finite Element Implementation of Multi-pass Fillet Weld with Phase Changes Phased Array. *Manufacturing Technology*, vol. 13, No. 1, 2013, pp. 79-85.
[5] FABIAN, P., JANKEJECH, P., KYSELOVA, M.: Simulation of Roundness, Hardness and Microstructure of Bearing Rings with thin Cross Sections by using SYSWELD, *Communications - Scientific Letters of the University of Zilina*, vol. 16., No. 3A, 2014, pp. 124-129.
[6] KIK, T. et al.: Numerical Simulations of Heat Treatment Processes. *Applied Mechanics and Materials,* vol. 809-810, 2015, pp. 799-804, doi:10.4028/www.scientific.net /AMM.809-810.799.
[7] *SYSWELD Engineering Guide to Training and Toolbox*. France: ESI Group, 2006.
[8] POIRIER, D.R, GEIGER, G.H.: *Transport Phenomena in Materials Processing*. The Minerals Metals and Materials Society, 1994.
[9] MICHALERIS, P., DEBICCARI, A.: Prediction of welding distortions, *Welding J.*, vol. 76, 1997, pp. 172-181.

Alexander Rengevic – Darina Kumicakova *

NEW POSSIBILITIES OF ROBOT ARM MOTION SIMULATION

The importance of computer simulations of the industrial robot arm motions at the technological scene becomes more important especially in connection with industrial applications of the new types of robot arm kinematic structures or with solving the human robot safety collaboration in industry. The article is aimed at the simulation of the basic (primitive) tasks executed by industrial robot with utilization of software packages of the "open-source" platform ROS. ROS is the software of modular architecture that extends the capabilities of the different robot types programming and controlling.

Keywords: *Computer simulation, industrial robot, ROS, program.*

1. Introduction

Computer simulation belongs to group of important tools for aid of industrial robotic development. Process of the simulation enables engineers to solve many tasks related to designing of the industrial robots construction and configuration, and also to test them at the robotized workplace virtual model by simulation. These tasks type can be classified as "basic tasks" which are usually solved with support of the complex program packages known as CAD/CAE and CAPP systems [1 and 2]. Their tools allow user to create the detailed models of the robot individual constructional parts, to define kinematic joints between them, to create the robot arm assembly and simulate its motions, velocities, accelerations and acting forces and torques, to test the design of the robot. For solving the complex tasks related to the industrial robot programming and implementation of the various input/output devices, it is necessary to use the programs belonging to the group of computer aids of manufacturing engineering (CAPE). Such programs consist of the special functions applicable only to the certain robot type and to solving the specific tasks. These programs enable to simulate: process of the robot programming, changes of the robot control system settings, controlling of input/output devices, grasping of parts and manipulation with them and also to test the robot working cycle at the virtual model of a robotized workplace. The Department of Automation and Production Systems has the practical experiences with solving the tasks related to the simulation of robotized welding and deburring, and also with robots cooperation. The Department has been using a special program of computer aided robot control and programming [3].

Till this time, the industrial robotics was focused most of all on automation of welding and painting processes, and manipulation with objects. But at the present the new trends in industrial robot applications are related to the request to involve them to the collaboration with human [4] or to place them at a mobile platform to be able to move inside a factory [5]. The programming of such industrial robots is much more complex and it requires an implementation of various specific devices or control algorithms which are not standard in current robotics. The solution of these problems is the object of interest of different development groups in the world. The effort to engage these specialists into the unified platform resulted in the project ROS Industrial that brings new possibilities to solve such complex tasks in industrial robotics [6] implementing a unified open-source platform The Robotic Operating System (ROS). This platform is freely accessible for researchers from around the world and enables them to cooperate and exchange information with each other and determine the future direction of the ROS.

In the framework of currently "running" European research projects [7] the several research workplaces (IPA Fraunhofer, Danish Technological Institute, TU Delft, LMS-University of Patras, etc.) participate in the research of tasks connected with intelligent and safety interactive Human-Robot (H-R) cooperation for assembly plants of the future. Research in individual solved projects is most of all aimed at hardware and software aids of symbiotic assembly solution and application of new technologies in area of the shared workspace operations sensing and controlling. The open-source system ROS is used for tasks relating to the H-R cooperation. This system consists

* **Alexander Rengevic – Darina Kumicakova**
Department of Automation and Production Systems, Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: alexander.rengevic@fstroj.uniza.sk

of tools as 3D sensing, automatic generation of trajectories or solution of inverse kinematic tasks [8].

The research of possibilities to utilize the platform ROS for creation of an intelligent safety system for the laboratory robot Fanuc LR MATE 200iC is now starting at the Department of Automation and Production Systems. Such a safety system should allow the robot to obtain online information from a surrounding workplace, recognize the objects in a monitored space and adapt the robot arm planned motions to current state of the risk of unwanted collisions with an object in the workplace. The article presents the basic information about the structure, tools of open-source platform ROS and selected steps related to the solving of the robot workplace simulation in ROS.

2. ROS and ROS Industrial

The ROS platform represents developmental environment that is used for creation of reliable programs for mobile and service robots control with utilisation of various types of hardware, drivers, libraries and other devices [9]. The ROS platform consists of the group of packages that include files starting the required tasks. The goal is to create the control program composed of several independent programs that can work together. Such starting program is called "nod". Single nods communicate with each other in environment called "Master" that is a very important part of the ROS platform. Dependencies between single nods in the framework of Master environment is possible to view with help of tool "rqt graph" - see example shown in Fig. 1. The ellipse represents a single nod. Arrows represent relations between single nods which send or receive messages related to the topic. The nod/turtlesim represents a simulation environment and it receives information from nod/teleop_turtle that provides motion commands. Both nods send information to the nod/rosout that collects published data from all available nods [10].

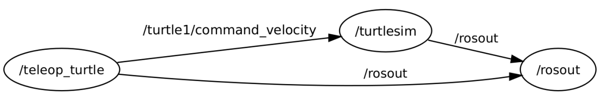


Fig.1 Visualisation of dependence between individual nodes under the ROS

When creating programs in ROS the build system catkin is used. The build system represents a tool that is able to translate the source code to files, libraries, binary files and scripts. When creating new projects in ROS it is necessary first to define the workspace. With regard to the directory the whole files essential for the right function of the created application must be placed. Structural arrangement of the working directory, packages, and codes which are used in the build system catkin is in Fig. 2.

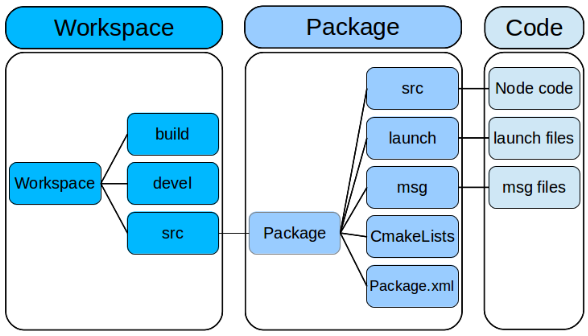


Fig. 2 Structure of directories and files within the build system catkin

ROS expands current possibilities of industrial robotics with new tools for spatial perception, recognition of objects, faces, gesture and motion tracking, stereovision and more. It is based on the support of the standard used hardware within robotics like industrial robots from known manufacturers (ABB, Adept, Yaskawa Motoman, Universal Robots or Fanuc), security elements and different types of end effectors and grippers. Besides all this, it has the ability to work with hardware which is not directly specific for usage in industrial robotics, for example, different types of image sensors and space depth sensors [9].

3. Simulation of robot working activity

The process of simulation is preceded by some basic steps. It is about choosing the right simulation tools, configuration of simulation tools and system linking. This section will describe each simulation capabilities and the environment in which the simulation will take place.

3.1. Simulation facilities

For the purposes of verifying and testing possibilities offered by the ROS platform – a notebook ASUS X550C computing device was used. The parameters of this device are stated in Table 1. The used operating system is preferentially intended for the ROS platform with codename Hydro. In GitHub repository which represents a worldwide stack of data for ROS community are available packages which are able to create programs for industrial robots Fanuc. For the use of an industrial robot Fanuc LRMate 200iC within platform ROS several initialization steps were executed. These initialization steps were able to recompile files created with ROS to formats compatible with the robot control system. This allowed the mutual communication between the robot and the ROS.

Table 1

ASUS X550C	
Processor	Intel Core i5-3317U
System memory	4016 MB DDR3
Graphic accelerator	Nvidia GeForce 720M 2GB DDR3
Operating system	Ubuntu 12.04.5 LTE Precise Pengoline

3.2. Communication between the robot control system Fanuc R30iA Mate and ROS

For the compilation of these files the program Fanuc Roboguide V7 was used. This program contains an essential function for compilation of ROS files to format supported by the robot control system. After compilation this folders were uploaded to a robot control unit of the real robot. After this procedure configuration and initialization of the files was provided. For the exchange of information between the control system of the robot and ROS the ROS configuration had to be setup. This included a set of network address, communication port and start up of the supporting communication programs. For the complete recovery of communication it will be necessary to add the robot control system with the options R623 and R648. These options will be able to exchange data between both systems.

3.3. Communication between Kinect device and ROS

The next device which will be necessary to create an intelligent safety system is a sensor used for spatial vision. For confirmation of the capabilities of spatial vision implementation a Kinect device was chosen. The functionality and communication ability of this device was tested with the use of computer with operating system Microsoft Windows 7 and corresponding drivers. After confirmation of the hardware functionality the device was connected to the ASUS X550C notebook with corresponding packages. Subsequently, the application for verifying the basic setups for the Kinect device was launched.

As can be seen in Fig. 3 on the right, the quality of image from rgb camera is sufficient enough. In the figure on the left is an image from a depth camera which shows the distance between the person and object in the space with the assistance of the color scale. Based on these two figures it can be concluded that the sensor Primesense is working correctly. The implementation of the picture from the depth camera in simulation environment is described in the experimental part.

3.3. Simulation environment

For creation of the intelligent safety system the program MoveIt! was used. This program is able to define the kinematics of the robot, motion and trajectory planning, collision detection, spatial perception and more in the environment of ROS platform. For the purposes of simulation the MoveIt! a graphical interface Rviz was used.

Rviz is able to easily detect and adjust actual settings of the simulating environment, simulated model and also monitor and display output data from the sensors that the robot is equipped with.

The first step in the creation of simulation environment was the preparation of the mathematical model of the industrial robot. For the purpose of simulation of the model in fanuc_drivers a package from GitHub repository was used. This eliminated the need to create a model of a robot in CAD system and convert it with to URDF (Universal Robotic Description Format) supported by ROS.

The next step was to create a configuration package for the chosen robot model. In this step it was necessary to define a way how to solve the inverse kinematics with the help of one of the

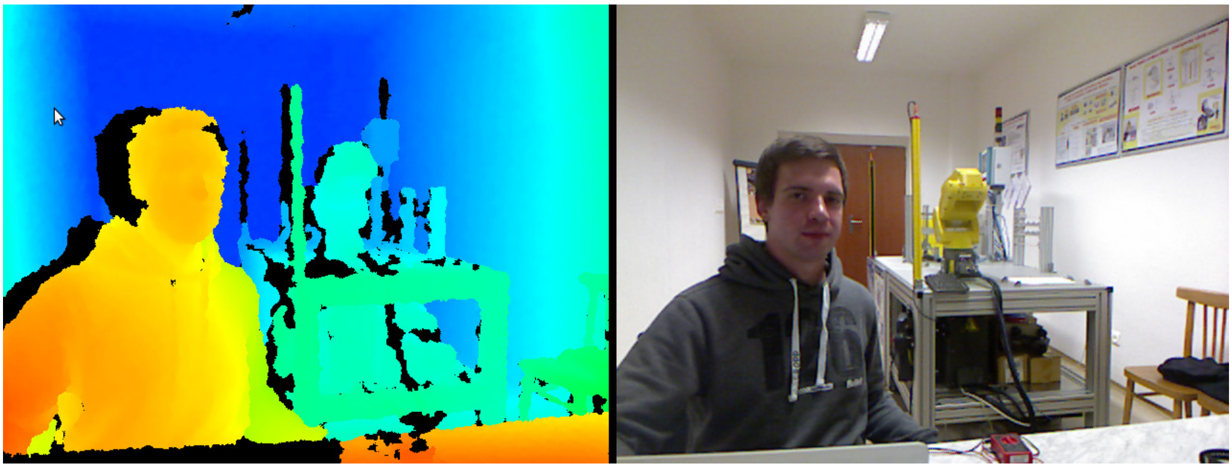


Fig. 3 Verifying the functionality of Kinect device

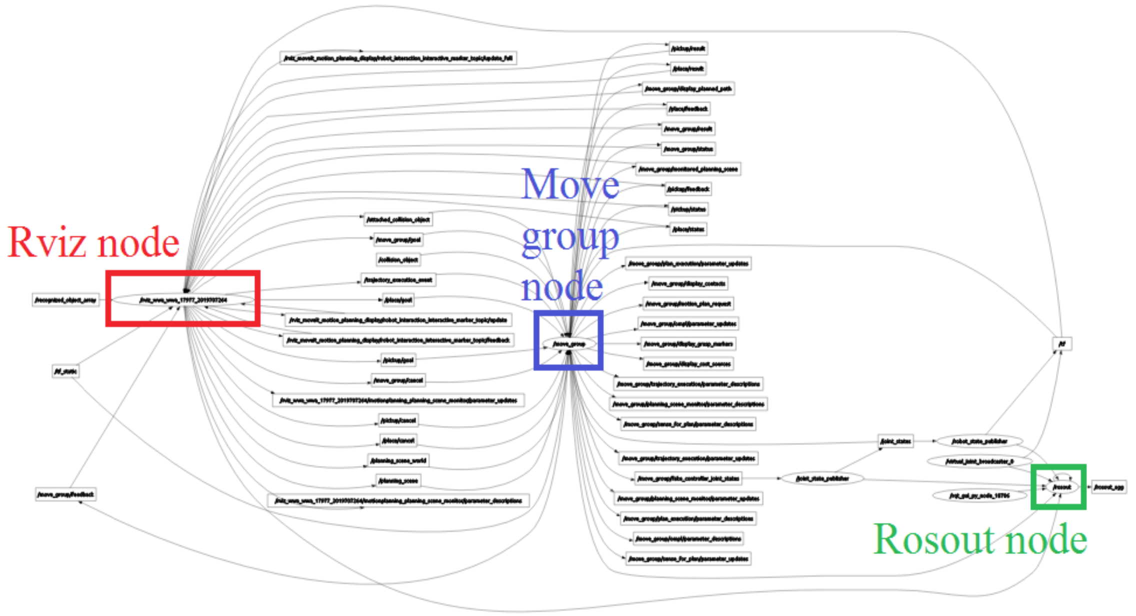


Fig. 4 Dependence between nodes after simulating environment launch

existing packages. The chosen library KDL (The Kinematics and Dynamics Library) is standardly used in ROS for industrial robots with six controlled axis. Except this, the collision states of joints, starting positions were setup and the range of the joints was verified. In the following step configuration files necessary for work in MoveIt! were generated. After launching the generated file in the Rviz environment we went in the definition of a move planner. This planner helps to specify the results of the solved kinematics. As a motion planner within ROS the OMPL (Open Motion Planning Library) was used. This library offers several ways of solutions depending on the chosen algorithm. For the purposes of experiment we chose a PRM (Probabilistic Roadmap Method) planner.

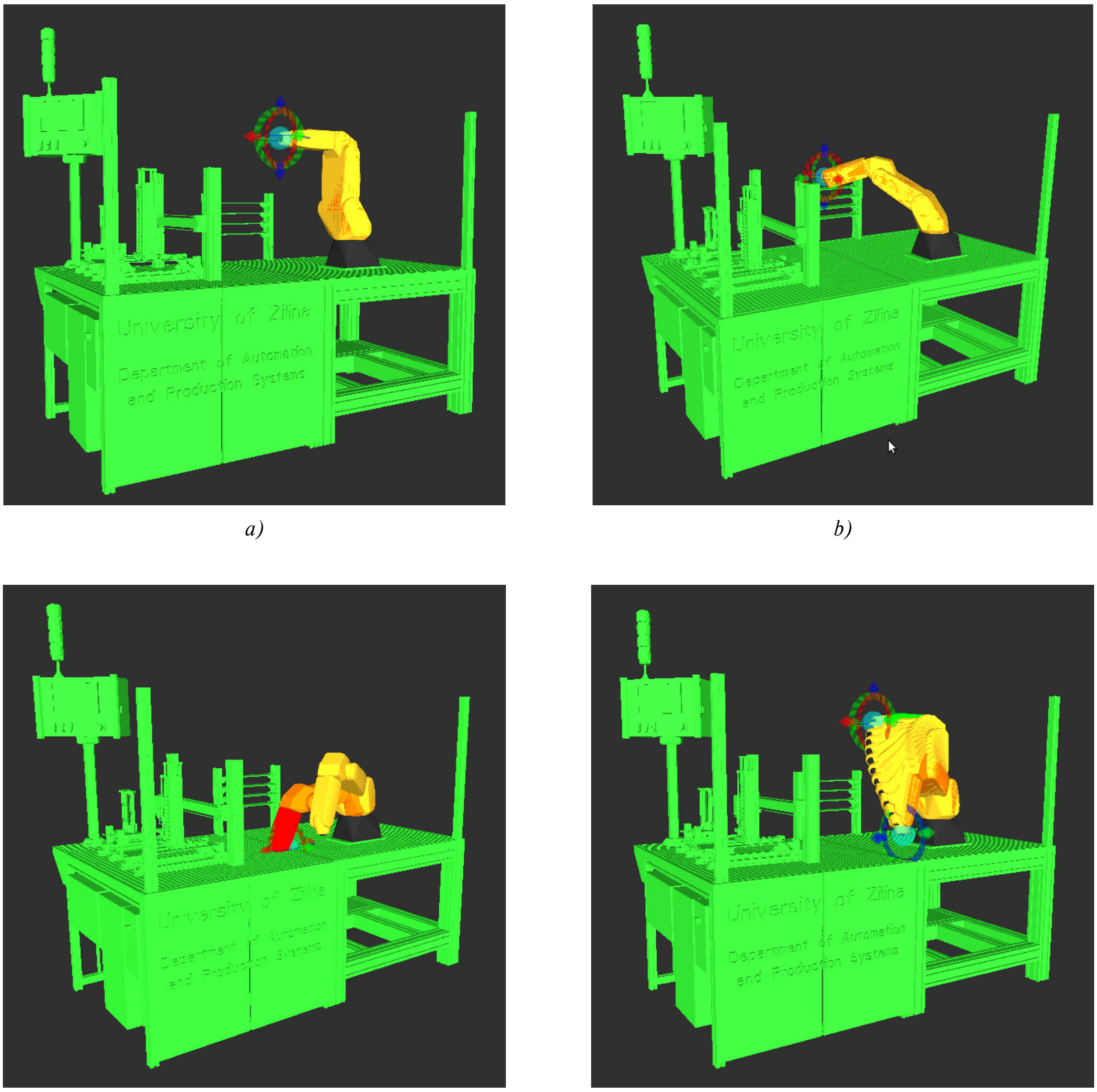
For the launch of the simulation environment it was necessary to use several nodes in which mutual dependences are represented in Fig. 4. As can be seen, the dependence net is quite complicated. The main nodes can be described as rviz node which represents the basics of simulation environments (marked in red color), move_group node which is responsible for the execution of the moving commands (marked in blue color) and node rosout (marked in green color) whose function is described in Chapter 2.

4. Simulation experiments

As first, the work of the industrial robot was simulated. The goal of the robot was to achieve required points within the operational space. The robot arm was, at the beginning of the simulation, in the starting position described in Fig. 5a. This position of the robotic arm was saved in the position database. From this database it is possible to load the position which was saved before. With the help of an orientation element located in the middle of the robot flange (in the tool center point, TCP) the whole positions were set. These positions are essential for the creating of the required robot working activity. One of the achieved positions is described in Fig. 5b.

The simulation served for the verification of functionality of the selected robot model and selected library for the purposes of solving inverse kinematics. The robot collision with the objects in the working space or self-collision was simulated. Recorded unwanted collision states were indicated in red color representing appropriate part of the robot arm. (Fig. 5c) After performing the execution for planned motion of the robotic arm the system evaluates the planned motion as faulty and does not allow the execution. The possibility of describing the trajectory of the robotic arm (Fig. 5d) and the motion in loop was available. Mathematical model of the robot which was selected for the simulation is not enough processed and does not contain the end effector by which the objects in the workspace are caught. Obtaining the basic information about behavior of the simulation is sufficient enough.

The display of output data obtained from the Kinect device was tested in the environment Rviz. Figure 6 represents the real workplace displayed as a point cloud. The point cloud represents a group of a bigger number of points where each point has a defined position in space. Each point has assigned information about color obtained from the rgb camera with which the Kinect device is equipped. In defining the settings in Rviz environment it was also possible to choose the resulting picture from different kinds of displays. Figure 6a displays a standard view of the



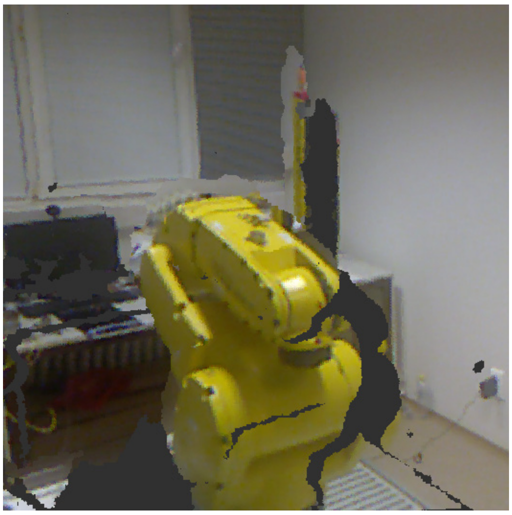
a) *b)* *c)* *d)*

Fig. 5 Examples of the robot simulation, a) initial state, b) achievement of defined point, c) collision between the robot arm and table, d) display of the robot trajectory

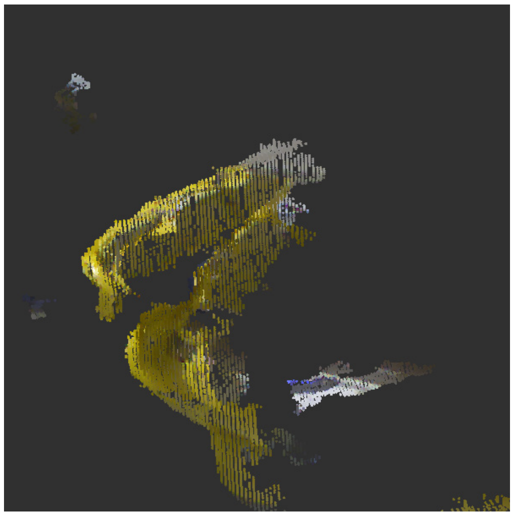
scanned space by using the Kinect device in the Rviz. After turning the view in program (Figure 6b) it is obvious that points are distributed within the simulated space. From the figures it is obvious that using a single Kinect device is not sufficient for the purposes of detecting and recognizing objects in the monitored area. That is why for the solution of creating an intelligent safety system we will consider a multi sensor system.

5. Conclusion

The article presents individual steps of testing the ROS platform utilization for the purpose to propose an intelligent safety system for the robot Fanuc LR Mate 200iC control. We set and verified communication between the robot Fanuc control system and ROS and also between Kinect and ROS. The functionality of the robot selected model was verified in the ROS simulation environment through the simple simulation experiments. Achieved results showed that ROS platform can be



a)



b)

Fig.6 View of workspace of robot in Rviz, a) direct view, b) at an angle

used for research of possibilities of controlling the robot working activities in laboratory conditions of the department. On the other hand, the obtained experiences and knowledge led us to specification of other tasks which will be implemented in the coming period:

- To create the new mathematical model of industrial robot with an integrated model of robot´s end effector that allows us to perform a better simulation.
- To setup, activate and verify the communication between the robot FANUC LR Mate 200iC control system and the ROS applications.
- To elaborate methods of recognizing the particular objects in the monitored space and implement created function into the ROS simulation environment. This allows us to create the new and advanced applications for the robot arm motions controlling.

Acknowledgement

This paper was created within the framework of a preparation phase of solution of the national project VEGA 1/0632/16 "Research of possibilities of intelligent safety systems solution for hybrid assembly systems based on interactive human-robot cooperation with aid of application of new technologies in the field of sensing of activities on the shared workspace".

References

[1] CUBONOVA, N.: Postprocessing of CL Data in CAD/CAM System Edgecam using the Constructor of Postprocessors, *Manufacturing technology: J. for Science, Research, and Production,* vol. 13, No. 2, 2013, 158-164, ISSN 1213-2489.

[2] GROZAV, S. D., CECLAN, V. A., KURIC, I.: *Determining Deformation Force in Orbital Deformation and Comparing it with other Conventional Methods,* International conference on Quality and innovation in engineering and management, 227-230, ISBN 978-973-662-978.

[3] KUMICAKOVA, D., RENGEVIC, A.: Automation of Manufacturing Technologies with Utilization of Industrial Robots, Applied Computer Science (ACS), vol. 11, No. 3, 2015, 6-18, ISSN 1895-3735.

[4] RENGEVIC, A., KUMICAKOVA, D.: Safety Cooperation between Human and Robot, TRANSCOM 2015: 11th European conference of young researchers and scientists, 257-262, ISBN 978-80-554-1047-0.

[5] BRUMSON, B.: New Applications for Mobile Robots, Robotic Industries Association, http://www.robotics.org/content-detail.cfm/Industrial-Robotics-Industry-Insights/New-Applications-for-Mobile-Robots/content_id/3362, 2015.

[6] http://rosindustrial.org/

[7] www.eu-robotics.net,

[8] http://www.smartroboticsys.eu/?lang=en

[9] http://www.ros.org/

[10] O'KANE J. M.: *A Gentle Introduction to ROS,* University of South Carolina, Columbia, ISBN 978-14-92143-23-9, 2014.

Martin Krajcovic - Andrej Stefanik - Luboslav Dulina *

LOGISTICS PROCESSES AND SYSTEMS DESIGN USING COMPUTER SIMULATION

In the course of logistics systems designing, great emphasis has been recently put on speed, quality and flexibility of the proposed solutions. The logistics system draft requires then not only an analysis of the logistics systems requirements, a draft of several solution variants and their static verification but also dynamic verification of performance of the designed system in real conditions in a form of computer simulation. The following article deals with the issues of the use of computer simulation for logistics processes and systems designing.

Keywords: *Logistics, simulation, design.*

1. Logistics Processes and Systems Designing

The aim of the logistics processes and systems designing is to draft such a logistics system in the pre-implementation phase which could meet all the necessary technical (optimum use of resources, high level of provided services, system flexibility, high transparency, etc.) as well as economic (return of investments, logistics costs, productivity, etc.) parameters without additional adjustments and interventions into the real system in the phase of its implementation and operation [2 and 3].

The current pressure on fast innovations in the factory means higher demands on logistics systems designing from the point of view of the laboriousness reduction, time and costs consumption for implementation of the entire designing process and increase of quality, complexity and information ability of outputs generated from the process [4].

Due to the above mentioned reasons it is possible to summarize the basic designing process requirements:

- Quick draft of new solutions.
- Observance of the system approaches when designing.
- Designing of logistics systems as a part of the digital factory concept.
- Interactive design of a new logistics system.
- A possibility of a running inspection and evaluation of the proposed solution variants.
- Application of the optimization approaches in the individual stages of the logistics system draft.
- Suitable visualization and presentation of the designing outputs.
- A possibility of dynamic verification of the proposed solution [2].

General procedures of logistics system designing include the following stages (Fig. 1):

1. Preparation, processing and analyzing of the input data.
2. Material flow analysis and transport performances calculation.
3. Analysis of existing means.
4. Processing of the logistics system variants – a static draft of the system.
5. Computer simulation – dynamic dimensioning of the material flow components [2].

1.1. Preparation, Processing and Analyzing of the Input Data

The basic information source for the designing is mainly a database from the technical preparation of the production. The data for logistics system designing has to include information about:

- Products (P – Product), which will be manipulated, transported and stored in the logistics system (products types, bills of materials, material consumption standards, construction parameters, physical characteristics of the individual items, etc.).
- Quantities (Q – Quantity), which will be manipulated, transported and stored in the logistics system (overall quantities of transport, time progress of the requirements –

* **Martin Krajcovic, Andrej Stefanik, Luboslav Dulina**
Department of Industrial Engineering, Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: martin.krajcovic@fstroj.uniza.sk

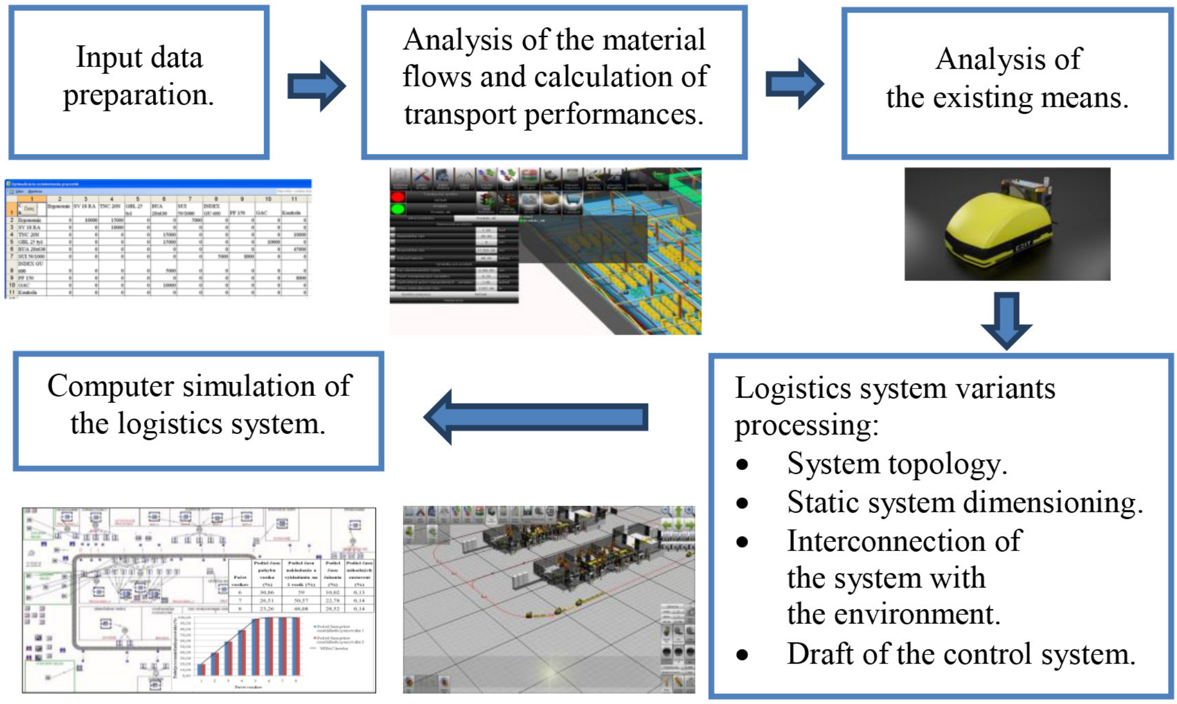


Fig. 1 General procedure of the logistics system designing [2]

seasonality, irregularity, transport, storage and manipulation units, etc.).

- Routings (R – Routing) of the individual items in the production and circulation (production procedures, assembly procedures, transportation relations, etc.).
- Supporting services (S – Supporting Services) necessary for the logistics activities implementation (maintenance and modernization of the logistics devices, quality management, packaging, etc.).
- Time (T – Time) periods of the individual activities in the logistics processes (loading and unloading time, transport time, preparation times, stock operations time, etc.) [4].

1.2. Material Flow Analysis and Transport Performances Calculation

The aim of this stage is a detailed analysis of the input data collected in the previous step and the analysis can provide the information about the current state of the logistics system, requirements for the parameters design and dimensioning of the future logistics system [4 and 5].

The base of the analysis is transported material items and their movements in the logistics or production system [2, 3 and 6].

From the range of products analysis of the transported parts point of view great emphasis is put on the use of tools which enable to analyze all items and to sort them into smaller subgroups (families) for which the common elements or subsystems of logistics system can be designed. The most often used tools are:

- Classification of material items on the basis of construction similarity (shape, dimensions, material, etc.).
- P – Q analysis: classification of the material items on the basis of transported volumes.
- P – P analysis: classification of the material items on the basis of their influence on the overall value flow.
- P – G analysis: classification of the material items on the basis of weight [7].

From the material flows point of view we examine quantitative (intensity of a material flow) and qualitative (movement direction, movement routes) aspect of a material flow. Methods and tools used for the analysis and optimization of the material flow can be divided into the following groups:

- Recording methods: progress lists, a where from – where to table.
- Graphical methods: line scheme, spaghetti diagram, Sankey diagram.
- Optimization methods (analytic, heuristic, metaheuristic): linear and dynamic programming, triangle method, CRAFT, CORELAP, ALDEP, and the like [8].

A synthetic indicator resulting from the analysis of transported parts and their movement is a calculation of transport performances and transport costs:

$$p = \sum_{i=1}^{m} \sum_{j=i+1}^{m} f_{ij} * d_{ij}, \qquad (1)$$

$$z = \sum_{i=1}^{m} \sum_{j=i+1}^{m} f_{ij} * c_{ij} * d_{ij}, \qquad (2)$$

where:

p - an overall transport performance (pcs . m / time unit),
z - overall transport costs (€/ time unit),
f_{ij} - intensity of material flows between objects i and j (pcs / time unit),
c_{ij} - transport costs for transport of one material unit per one distance unit between nobjects i and j (€/ (m . pcs)),
d_{ij} - distance between objects i and j (m) [8].

1.3. Analysis of Existing Means

A part of the analytic stage of logistics system designing is an analysis of available resources. The resources necessary for the logistics process include:

- Transport, manipulation and storage units – they form passive elements of a logistics system, enable effective manipulation, transportation and storage of material items.
- Means of transport and manipulation devices – active parts of the logistics transport system; devices which provide relocation of material from the place of supply to the place of consumption.
- Storage facilities – devices which provide temporary storage of material before it is used.
- Stuff – workers responsible for management and operation of logistic activities.
- Areas – production areas or factory areas, which are connected with the performance of logistics activities [9].

The analysis of existing means provides an assessment of structure, quantity, technical parameters and efficiency of use of the existing means in logistics. In the draft stage the required structure of means is always compared to the existing one and a solution of disproportions between both data files (procurement of missing means, elimination of redundant means) is also a part of the draft. The need for means and the real level of their usage is one of the indicators of productivity and economy assessment of a logistics system [10].

1.4. Processing of the Logistics System Variants – a Static Draft of the System

The proposed stage consists in a draft of several variants of the logistics system, which meet the final criteria of the project. Each variant has to deal with:

1. Logistics system topology: the system topology describes the spatial solution of a logistics system, i.e. the arrangements of the individual transport devices, a draft of transport routes, a definition of loading and unloading sites, location of storage facilities, etc. The system topology itself is affected by results of an analysis of the transported range of products and material flows and by the selected suitable methods of transport, manipulation and storage.
2. Static dimensioning of material flow elements: it represents a stage of capacity dimensioning of a logistics system, a selection of specific types of transport, manipulation and storage means and definition of their parameters (bearing capacity, dimensions, speed, loading and unloading time) and a calculation of the overall need of the individual types of devices used in the logistics system is implemented for the proposed topology of a logistics system.
3. Interconnection of the system with the environment: the designed logistics system has to be effectively interconnected to the previous and subsequent factory processes. Due to this fact great emphasis is in the project of a logistics system draft put on the designing of input and output places of the logistics system. The aim is to create such a way of passing and receiving of material items between the system and its environment so that laboriousness and times of waiting for material are reduced.
4. A control system draft: whereas the previous three steps deal with the draft of a part of the logistics system as for the physical material flow, the last step consists of a draft of information flows in the logistics system which includes collection, processing, storing (logistic record keeping) and use of information in the processes of planning and controlling of logistics processing. The control system draft is a key part of the logistics system draft because the organization factors have the greatest influence on the logistics processes progress; the technical parameters are of significantly lower importance [2 and 6].

1.5. Computer Simulation – Dynamic Dimensioning of the Material Flow Components

On the basis of the proposed concept of the logistics system it is important to implement the dynamic evaluation of system performance in the final stage of the logistics system project and complete the proposed logistics project in detail. It is possible to implement the analysis from the physical rules and sensors signals point of view (e.g. simulation of the dependency of transport time and changes of speed caused by events on the transport route – an obstacle on the route, loading/unloading of material, reduced speed of a logistics element, etc.) as well as the analysis based on the change of system condition in a form of a discrete or connected event-oriented simulation [10 and 11].

The verification enables not only to define an optimum number of logistics means but also to check the performance in emergency situation which can be a malfunction of a part of a production line, a malfunction in the preparation supermarket, blocking of the logistics means by an obstacle or verification for how long it is possible to put one or more trucks out of operation because of planned od unplanned maintenance. Compared to the static calculation it provides more relevant information and it is also possible, by means of visualization, to familiarize with the functions of a given system before the implementation itself. It saves not only time but also costs linked with the problem solution in the situation of real production use. Besides standard problem types we encounter the problems related to specific conditions of use in a form of crossroads or cooperation of several types of logistics means. The result is a debugged system, which meets all the necessary requirements with predefined possible critical areas [2 and 9].

In case of need it is possible to check the entire material flow, to find the basic information like running production time, average unfinished production or to define the size of input, output or intermediate operation storage. Considering the use of theoretical classification of likelihood and simulated malfunctions, idle times and changing of production type or consecutiveness of the individual production processes it is possible to gain the real results, which correspond with the performance of the real production system [2 and 9].

2. Examples of Use of the Computer Simulation in the Logistics System Designing

A simulation application when solving the problems of logistics system draft will be demonstrated on an example from industrial practice. It concerns a draft and dimensioning of logistics system supply of production workplaces by necessary material inputs.

2.1 Problem Description

The production process is implemented in one production hall and the logistics process is ensured through an automated guided vehicle (AGV) system. Material is transported to the individual workplaces in defined sequences from the entrance storehouse. Material is located in box pallets which are transported to the place of need by an AGV towing unit in specialized C-frames (manipulation platforms), see Fig. 2.

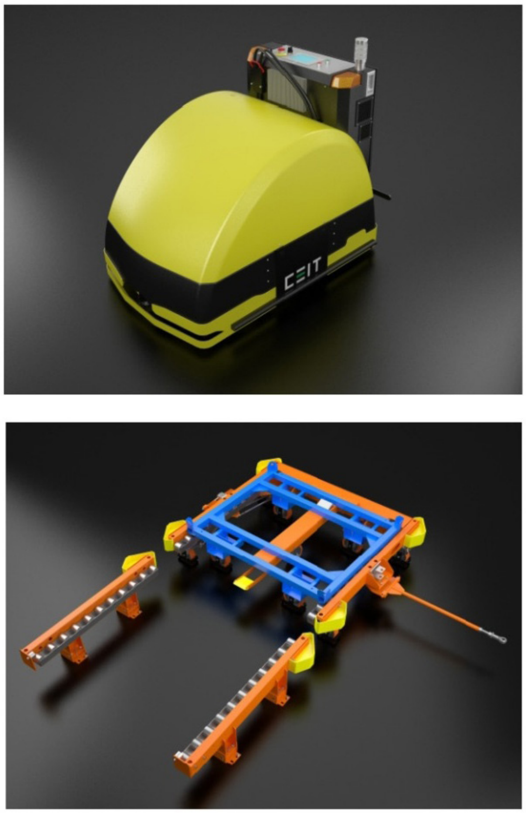


Fig. 2 Transport means used in the logistics system

Material is delivered from the entrance storehouse. There are currently nine catching positions in three time sequences (12 min, 24 min, 48 min) designed within the storehouse. The sequences are designed on the basis of production cycles of individual workplaces. The arrangement of the catching positions is not optimal at the moment and, therefore, it is necessary, based on the monitored indicators, to define a new arrangement of the sequences in the storehouse.

2.2. Static Draft of the Logistics System

There are currently several software solutions for support of the static draft of production and logistics systems like visTABLE from Plavis, Tecnomatix FactoryCAD/FLOW from Siemens or Malaga V3 from Zip Industrieplanung. A software tool CEIT Table was used in the described problem solution due to availability of advanced functions for designing of logistics and transport systems.

The static draft in the environment of CEIT Table software requires the following input data:

- Transported items.
- Volume of transport for the individual items.
- Used manipulation units.

Fig. 3 Example of the transport system draft using the CEIT Table software

- Number of pieces in the manipulation units.
- Available time capacity of a transport device.
- Speed of a transport device.
- Times of loading and unloading.

The CEIT Table system includes graphic environment which enables to design a layout of the logistics system in 3D environment and subsequently on the basis of a 3D model of the layout and entered data to automatically specify the overall laboriousness of transport and the necessary number of transport devices (Fig. 3).

2.3. Dynamic Simulation of the Designed System

Dynamic simulation is used for detailed verification of drafts when designing production, assembly and logistics processes. An advantage of the dynamic simulation is modeling and implementation of dynamics into the designed processes and a possibility to examine the influence of mutual interaction of the individual elements of the designed systems. From the logistics system draft and optimization point of view it is primarily applied in the following areas:

- Testing of changes in the transport system and their impact on production/assembly system.
- Testing of various goods supply systems (direct supplying on request, transport circuits and the like) and their assessment.
- Assessment of changes in transport devices types and their capacities on system.
- Testing of changes in priorities of transport tasks and their impact on the system's ability to provide the necessary tasks.
- Monitoring of the influence of crossroads and transport junctions solutions on the utilization of transport devices.
- Testing of impact of the transport devices failure on the material supply.
- Assessment of impact of loading/unloading processes at input/output on the amount of transport devices waiting for loading / unloading.

Storing is also a part of logistics processes, which represents a point of contact of transport and production/ assembly tasks. The simulation is primarily aimed at collection of statistical data and their visualization so it can quickly and simply evaluate various influences on the size related level of storehouses like:

- Influence on planning and different priority rules of tasks on the intermediate operation storehouses.

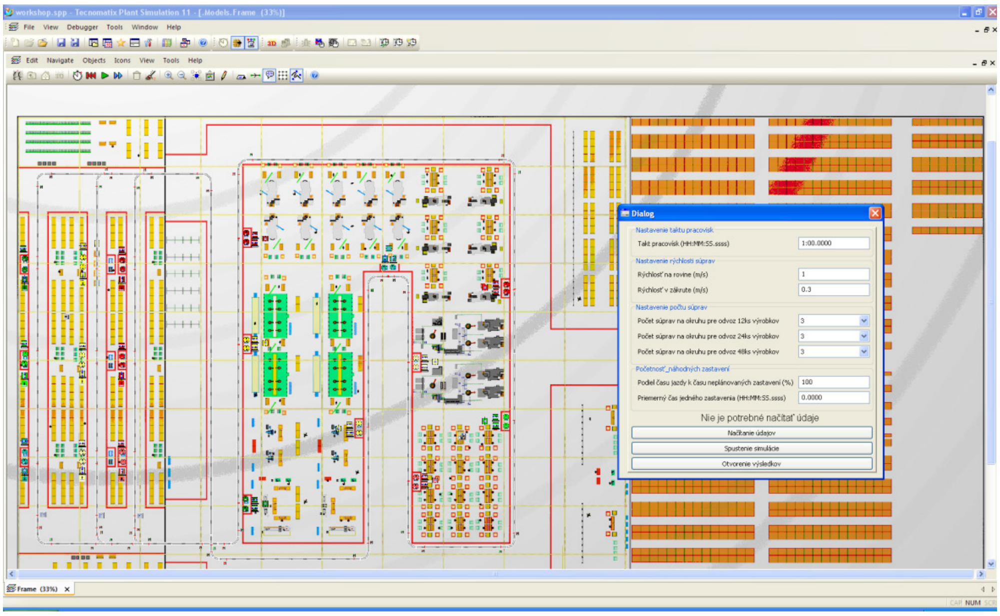


Fig. 4 Simulation model of the logistics system created in the environment of the Plant Simulation software

- Influence of shift pattern in the workplaces in different parts of production/assembly/logistics system on the maximum level of storehouses.
- Influence of random phenomena – breakdowns and type changing – on the overall storehouse capacity, which a factory is forced to create to provide the maximum storehouse utilization.

In the solution of the presented problem a software tool Plant Simulation from SIEMENS was used for the purposes of dynamic simulation. The simulation model of the proposed logistics system is pictured in Fig. 4.

The following input parameters of the model, which can be modified during the implementaion of individual simulation runs, were defined within the testing of the individual variants of workplaces supplying by material items:

- Products processing cycle.
- Vehicle combination speed on flat surface.
- Vehicle combination speed in on bends.
- Number of vehicle combinations on the individual circuits.
- Drive time/accidental stopover time ratio (stopovers due to obstacle on the road).
- Average time of an accidental stopover.

The aim of the simulation is to design and dimension the logistics system in terms of capacity so that it can fulfill the basic function of non-stop supplying of workplaces by means of necessary inputs and at the same time to ensure the optimum utilization of the transport system.

The simulation software Plant Simulation provides both graphic and data outputs when evaluating the individual simulation experiments. A graph of material supply performance in the individual workplaces of a line is pictured in Fig. 5. This parameter was monitored while making a decision about the optimum number of trucks used in the designed transport system.

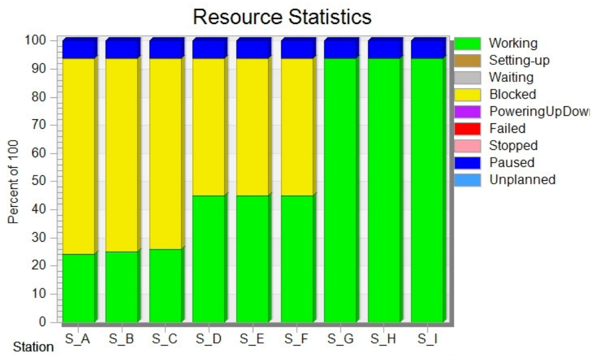


Fig. 5 Graph of material supply performance in the workplaces

Results of the individual simulation experiments are depicted in Table 1 and Fig. 6. The results show that the target criterion was met only in case of simulation experiments number 10 and 22. The remaining results are not satisfying due to unsuitable average utilization of the production system workplaces. The unsatisfactory simulation results were in most of simulation experiments caused by blockage in the first workplace.

Simulation experiments results Table 1

Experiment no.	Tuggers number			Utilisaton of workplace									Satisfy?	Note
	sequence = 12 min	sequence = 24 min	sequence = 48 min	A	B	C	D	E	F	G	H	I		
1	1	1	1	53.84	54.7	55.5	100	100	100	100	100	100	NO	
2	1	1	2	28.75	28.88	29.68	53.03	55	55	100	100	100	NO	Blocked by first workplace
3	1	1	3	13.75	13.88	14.68	21.25	21.25	21.25	100	100	100	NO	Blocked by first workplace
4	1	2	1	48.75	48.97	49.77	100	100	100	100	100	100	NO	
5	1	2	2	28.75	28.88	29.68	100	100	100	100	100	100	NO	Blocked by first workplace
6	1	2	3	13.75	13.75	13.75	36.25	36.25	36.25	100	100	100	NO	Blocked by first workplace
7	1	3	1	28.75	28.92	29.73	100	100	100	100	100	100	NO	
8	1	3	2	22.58	23.44	23.75	85	85	85	100	100	100	NO	Blocked by first workplace
9	1	3	3	14.01	14.01	14.01	52.8	51.41	52.35	100	100	100	NO	Blocked by first workplace
10	2	1	1	100	100	100	100	100	100	100	100	100	YES	
11	2	1	2	38.23	39.09	39.89	37.89	40	39.74	100	100	100	NO	Blocked by first workplace
12	2	1	3	23.75	23.88	24.68	21.25	21.25	21.25	100	100	100	NO	Blocked by first workplace
13	2	2	1	78.35	77.55	77.55	100	100	100	100	100	100	NO	
14	2	2	2	28.75	29.12	29.93	52.89	55	54.74	100	100	100	NO	Blocked by first workplace
15	2	2	3	21.25	21.25	21.25	36.25	36.25	36.25	100	100	100	NO	Blocked by first workplace
16	2	3	1	53.41	52.73	52.73	100	100	100	100	100	100	NO	
17	2	3	2	33.75	33.75	33.75	77.89	79.71	79.71	100	100	100	NO	Blocked by first workplace
18	2	3	3	23.75	23.75	23.75	52.89	55	54.74	100	100	100	NO	Blocked by first workplace
19	3	1	1	100	100	100	85.12	85.25	85.12	100	100	100	NO	
20	3	1	2	50.52	51.38	52.18	31.25	31.25	31.25	100	100	100	NO	Blocked by first workplace
21	3	1	3	30.52	31.38	32.18	21.25	21.25	21.25	100	100	100	NO	Blocked by first workplace
22	3	2	1	100	100	100	100	100	100	100	100	100	YES	
23	3	2	2	43.23	44.09	44.89	52.89	54.71	54.71	100	100	100	NO	Blocked by first workplace
24	3	2	3	28.75	28.75	28.75	36.25	36.25	36.25	100	100	100	NO	Blocked by first workplace
25	3	3	1	73.91	72.38	72.38	100	100	100	100	100	100	NO	
26	3	3	2	41.25	41.62	42.43	77.89	80	79.74	100	100	100	NO	Blocked by first workplace
27	3	3	3	30.52	31.38	32.18	51.25	51.25	51.25	100	100	100	NO	Blocked by first workplace

where A – I are significations of workplaces

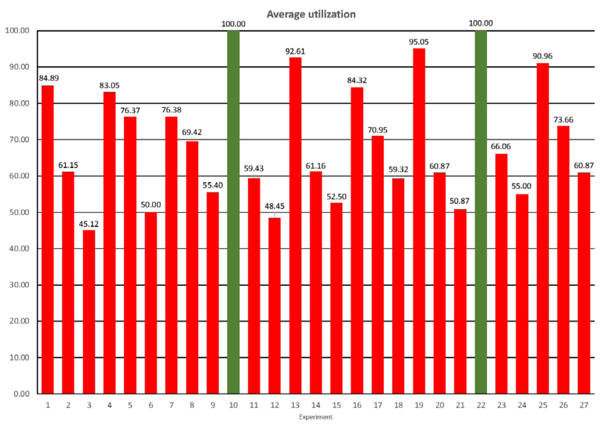


Fig. 6 Average utilization of workplaces achieved in the individual simulation experiments

3. Conclusion

Computer simulation currently represents an important designing and optimization tool of logistics processes and systems. Its main advantage lies in a possibility, in case of a correct construction of a simulation model and after its verification and validation, to simulate with a very high accuracy the behavior of a logistics system in real conditions. The individual simulation experiments are used for modeling of various situations and scenarios, which can occur in reality and enable to assess their impact on achievement of target indicators of the designed system. Computer simulation is, therefore, a very important tool when implementing "what-if" analyses, which are a part and parcel of the conceptual and detailed designing of production, assembly and logistics systems. This is the reason why computer simulation currently forms an integral part of digital factory and PLM (Product Lifecycle Management) concept of software solutions.

References

[1] HNAT, J.: *Virtual Factory Framework.* Industrial Engineering Moves the World - InvEnt 2012, University of Zilina, 2012, 56-59, ISBN 978-80-554-0542-1.

[2] KRAJCOVIC, M. et al.: Intelligent Manufacturing Systems in Concept of Digital Factory. *Communications - Scientific Letters of the University of Zilina*, vol. 15, No. 2, 2013, Zilina, 77-87, ISSN 1335-4205.

[3] MICIETA, B, BINASOVA, V., HALUSKA, M.: The Approaches of Advanced Industrial Engineering in Next Generation Manufacturing Systems. *Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3A, 2014, Zilina, pp. 101-105, ISSN 1335-4205

[4] LI, J., MEERKOV, S. M.: *Production Systems Engineering.* New York: Springer, 2009, 666 p., ISBN 0387755780
[5] ALTUNTAS, S., SELIM, H.: Facility Layout using Weighted Association Rule-based Data Mining Algorithms: Evaluation with Simulation. *Expert Systems with Applications*, 39 (1), 2012, pp. 3-13.
[6] PLINTA, D., WIĘCEK, D.: *Production Systems Design*. Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej, Bielsko-Biala, 2012, p. 122, ISBN 978-83-63713-06-5
[7] JERBI, A. et al.: *Functional vs Cellular Layout: Using Simulation as a Comparison Tool.* 3rd intern. conference on advances in mechanical engineering and mechanics, 2006.
[8] GREASLEY, A.: Using Simulation for Facility Design: A Case Study. *Simulation Modelling Practice and Theory*, 16 (6), 2008, pp. 670-677.
[9] EBBESMEYER, P. et al.: *Designing Flexible Production Systems with Virtual Reality.* Proc. of the ASME 2001 Design Engineering Technical Conference on Computers and Information, Pittsburgh, 2001.
[10] HOVANEC, M, SINAY, J., PACAYOVA, H.: *Application of Proactive Ergonomics Utilizing Digital Plant Methods Based on Augmented Reality as a Tool Improving Prevention for Employees.* Intern. Symposium on Occupational Safety and Hygiene: 02. 2014: Guimares, pp. 182-185, Guimares: SPOSHO, 2014, ISBN 978-989-98203-2-6.
[11] GASOVA, M.: Modern Methods in Ergonomics Research and European Legislation. *Advanced Industrial Engineering*. Monograph. Publisher: Wydawnictwo Fundacji Centrum Nowych Technologii : Bielsko Biala, 2015, 1st ed., pp. 147-164, ISNB 978-83-927531-7-9.

Andrej Czan - Michal Sajgalik - Jozef Holubjak - Marianna Piesova - Tatiana Czanova *

SIMULATION OF RESIDUAL STRESSES IN SURFACE AND SUBSURFACE LAYER AFTER MACHINING

This article deals with non-destructive detection method of residual stress as analytical tool of expected distribution of residual stress in surface and sub-surface layers. Based on measurements, which can consists of cloud of points, lines or surface map of residual stress, we can simulate potential distribution of residual stress in surface and subsurface layers via mathematical-statistical procedures. Right identification of residual stress and its distribution can improve the prediction of failures and similar damage incidences due to workload over lifetime of components and also, it can be used as evaluation parameter of suitability of applied manufacturing technological operations.

Keywords: *X-ray diffractometry, residual stress, non-destructive detection method, simulation of distribution.*

1. Introduction

Residual stresses are an integral part of manufactured workpieces, whether they are introduced deliberately, as a part of the design, as a by-product of a process carried out during the manufacturing process, or are present as the product of the component's service history. Residual stresses have cumulative effect to the existing stresses in the workpieces as a result of service loads [1 and 2]. In practice, residual stress is stress which affects the entire volume of part or the majority of itself, i.e. macroscopic character. This includes the stress in infinitely thin or large area. It is important that violation of compactness of part causes the change in macro-geometry. Methods of machining, casting, forming, etc. can cause these changes. For full classification, it should be noted that residual stresses are called sometimes as technological stresses, because they arise from the action of technological processes during the producing of parts. Direction of residual stress (tension or compression) depends on the kind of deformation. Permanent residual stresses have the largest share on the functionality of part, and they cannot be detected by conventional methods [3 and 4].

Obviously, to realize the benefits of understanding the residual stresses in parts and structures, tools are needed to measure them. Several techniques are available, with varying degrees of sophistication. Some of them are rather limited in their application, but one stands out as having widespread applications and being readily available [5 and 6].

X-ray diffraction is a specific method that can measure residual stress quantitatively in crystalline and semi-crystalline materials, which include virtually all metals and their alloys, and most ceramic materials. It is a non-destructive detection technology in many applications, and is widely accepted by the engineering community, being the subject of SAE and ASTM publications, which provide reliable sources of information on methods to ensure repeatability and reliability in the results of measurements. Because individual measurements are non-destructive, they can be replicated to demonstrate their statistical reliability [7, 8 and 9].

This article is focused on determining stress characteristics and their simulated prediction of distribution in surface and subsurface layers.

2. Residual stress

Stress conditions are the ones of demonstrations of used machining technologies. After manufacturing process, they remain in parts and constructions and they operate continuously even without load. Their impact significantly affects the functionality of machined surfaces. [2, 10 and 11].

* **Andrej Czan, Michal Sajgalik, Jozef Holubjak, Marianna Piesova, Tatiana Czanova**
Department of Machining and Manufacturing Technology, Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: andrej.czan@fstroj.uniza.sk

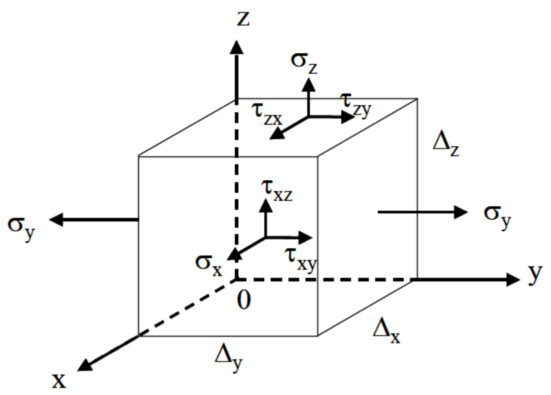


Fig. 1 Scheme of triaxial stress acting on an elemental unit cube [10]

Normal stress is defined as the stress acting normal to the surface of a plane; the plane on which these stresses are acting is usually denoted by subscripts. For example, consider the general case as shown in Fig. 1, where stresses acting normal to the faces of an elemental cube are identified by the subscripts that also identify the direction in which the stress acts, e.g., σ_x is the normal stress acting in the *x* direction. Since σ_x is a normal stress it must act on the plane perpendicular to the *x* direction. The convention used is that positive values of normal stress denote tensile stress, and negative values denote a compressive stress [10, 12, and 13].

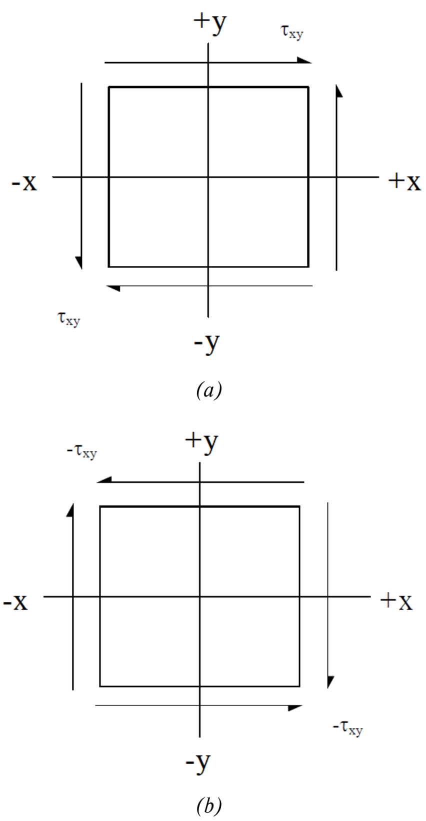


Fig. 2 Sign convention for shear stress - (a) Positive, (b) negative [10]

A shear stress acts perpendicular to the plane on which the normal stress is acting. Two subscripts are used to define the shear stress, the first denotes the plane on which the shear-stress is acting and the second denotes the direction in which the shear stress is acting. Since a plane is most easily defined by its normal, the first subscript refers to this. For example, τ_{zx} is the shear stress on the plane perpendicular to the *z*-axis in the direction of the *x*-axis. The sign convention for shear stress is shown in Fig. 2, which follows Timoshenko's notation. That is, a shear stress is positive if it points in the positive direction on the positive face of a unit cube. It is negative if it points in the negative direction of a positive face. All of the shear stresses in (a) are positive shear stresses regardless of the type of normal stresses that are present, likewise all the shear stresses in (b) are negative shear stresses [10 and 14].

3. Principles of measurement of residual stress by X-ray diffraction

The residual stress determined using X-ray diffraction is the arithmetic average stress in a volume of material defined by the irradiated area, which may vary from square centimeters to square millimeters, and the depth of penetration of the X-ray beam. The linear absorption coefficient of the material for the radiation used governs the depth of penetration, which can vary considerably. However, in iron, nickel, and aluminium-based alloys, 50% of the radiation is diffracted from a layer approximately 0.005 mm deep for the radiations generally used for stress measurement. This shallow depth of penetration allows determination of macro and microscopic residual stresses as functions of depth, with depth resolution varying from 10 to 100 times larger than potentiality of other methods. Although in principle virtually any inter-planar spacing may be used to measure strain in the crystal lattice, the availability of the wavelengths produced by commercial X-ray tubes limits the choice to a few possible planes. The choice of a diffraction peak selected for residual stress measurement impacts significantly on the precision of the method. The higher the diffraction angle, the greater the precision. Practical techniques generally require diffraction angles, *2θ*, greater than 120° (Fig. 3) [8 and 15].

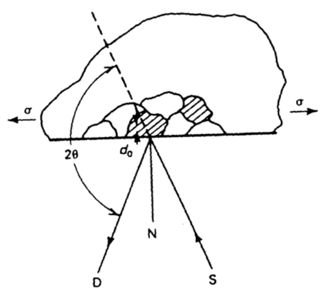


Fig. 3 Principles of X-ray diffraction stress measurement [8]

Plane-stress elastic model X-ray diffraction stress measurement is confined to the surface of the sample. Electro-polishing is used to expose new surfaces for subsurface measurement. In the exposed surface layer, a condition of plane stress is assumed to exist. That is, a stress distribution described by principal stresses σ_1 and σ_2 exists in the plane of the surface, and no stress is assumed perpendicular to the surface, $\sigma_3 = 0$.However, a strain component perpendicular to the surface ε_3 exists as a result of the Poisson's ratio contractions caused by the two principal stresses (Fig. 4) [3].

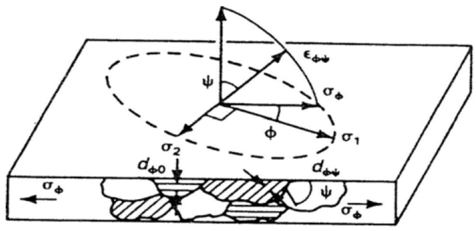


Fig. 4 Plane-stress elastic model [8]

4. Experiment conditions

Experiment was performed on constructional non-doped steel C56E2, which is used for production of bearing rings and other rolling parts of bearings. This steel is also suitable for production of shafts, gear wheels, pivots etc. It is characterized by worse weldability, and it is used mostly for production of parts with higher wear resistance.

Sample for the experiment was machined by turning technology on CNC turning center. We monitored the residual stress with constant cutting speed (v_c=200m.min^{-1} for roughing and v_c=250m.min^{-1} for finishing) and depth (a_p=3mm for roughing and a_p=0.5mm for finishing) and increasing values of feed (f=0.17÷0.55mm for roughing and f=0.05÷0.1mm for finishing).

The measuring of residual normal and shear stress was performed with Proto XRD diffractometer (Fig. 5) and electro-chemical polisher for removing the minimal layers of surface up to 200μm.

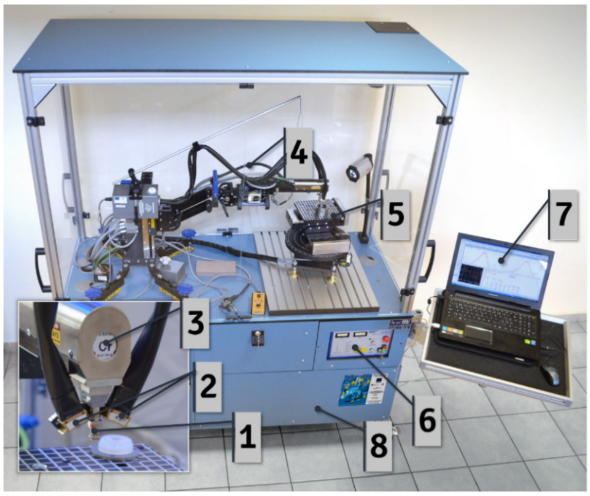


Fig 5 X-ray diffraction residual stress measurement system (stationary: 1 - collimator; 2 – two detectors for diffraction cone; 3 - X-ray tube; 4 - Cobralink® flexible arm; 5 - positionable and rotary table; 6 – control unit; 7 – PC with software; 8 - laboratory stand

From the theory of elasticity the relationship between residual stress (σ) and strain (ε) on the sample surface under plane stress is given by the Bragg equation, $\lambda=2d \sin \theta$, relating incident X-ray wavelength (λ), lattice inter-planar spacing (d) and diffraction angle (θ) (Fig. 6). The direction of maximum residual stress, which can be tensile or compressive, is assumed to occur in the cutting direction during most machining operations.

The residual stress field at a point, assuming a condition of plane stress, can be described by the minimum and maximum normal principal residual stresses, the maximum shear stress, and the orientation of the maximum stress relative to some reference direction. The minimum stress is always perpendicular to the maximum. The maximum and minimum normal residual stresses, shown as σ_1 and σ_2 in Fig. 4, and their orientation relative to a

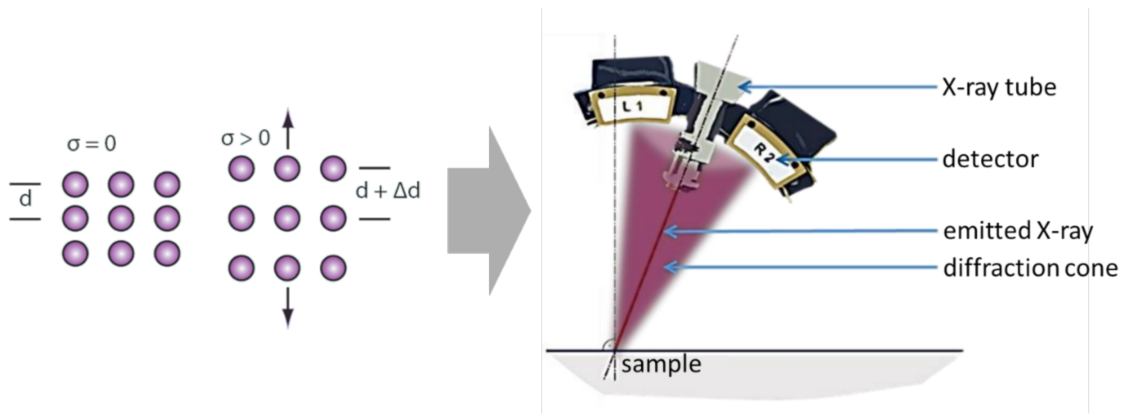


Fig. 6 Principle of measuring of residual stress by X-ray diffractometry based on Brag's Law

reference direction can be calculated along with the maximum shear stress using Mohr's circle for stress, if the stress σ_ϕ is determined for three different values of ϕ.

5. Experimental results

Input data for simulation were measured as cloud of points or square map (10×10mm) consisting of points regularly disposed within it. Measurements were made in the surface layer (depth of 0mm) and regular subsurface layers up to 200μm. In the following figures we can see the selected courses of normal residual stress through the measured depth in one place (at one point).

Selected Figs. 7 and 8 present the courses of residual stress through the measured depth. Almost all the cases, with the given cutting conditions, have a similar course. In the surface layer and higher subsurface layers, the residual stress had compressive character. With the increasing depth, these residual stresses decreased, or changed character to the tensile stress.

From measured data, we computed via STATISTICA software the simulation of distribution of normal residual stress in three directions: direction of cutting speed, direction of cutting feed (or speed of cutting feed) and direction of depth. Subsequently, 3D graphs of these distributions were generated.

3D simulation of distribution of normal residual stress after roughing operations shows changes of normal residual stress in surface layer in direction of cutting speed and cutting feed (Fig. 9a). No step changes can be seen: normal residual stresses are in the range between 175MPa and 245MPa with compressive character. Roughing operations with cutting depth 3mm are relatively suitable for next finishing operations without inheritance of visible signs of possible rise of errors. Simulation into the depth (Fig. 9b) of part material shows residual stress with tensile character in deeper layers and a very large range of values (from 292MPa of compressive residual stress up to 74MPa of tensile residual stress). The risk of errors increased, but in

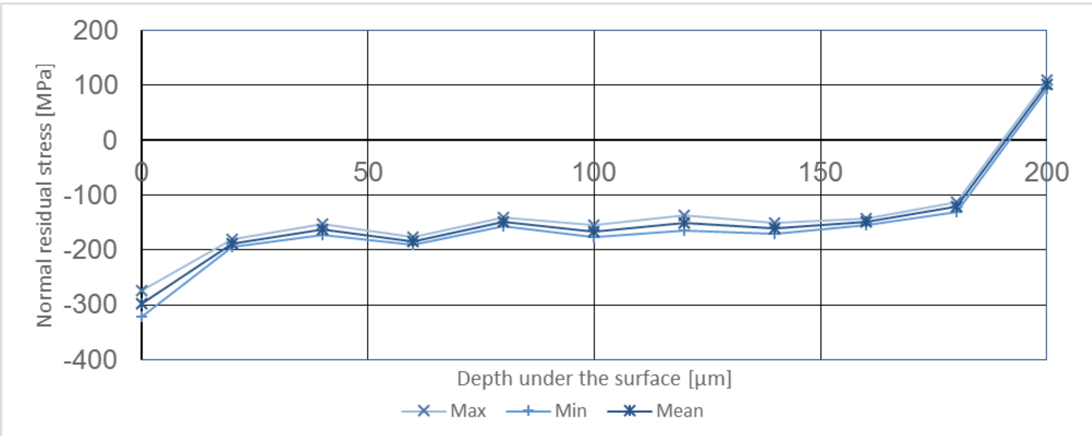


Fig. 7 Measured values of normal residual stress into the depth after roughing (cutting conditions: v_c = 200m.min^{-1}, f = 0.17mm, a_p = 3mm)

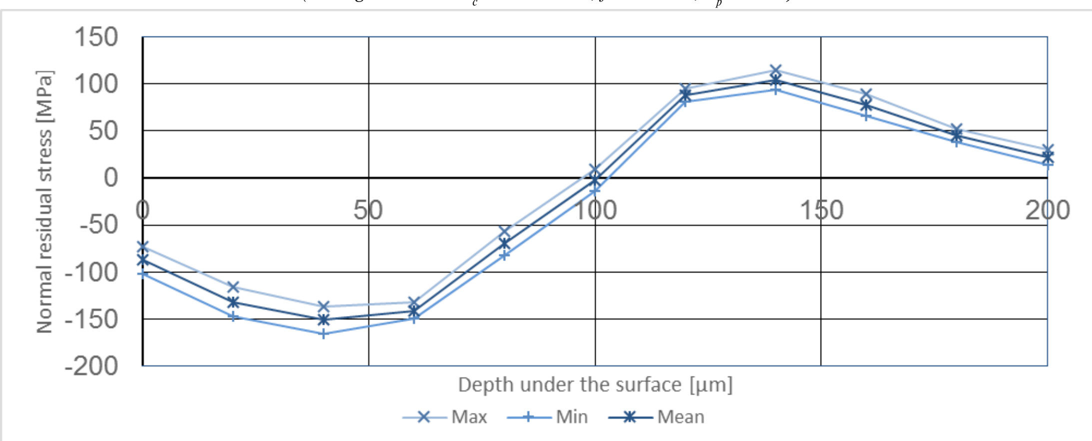
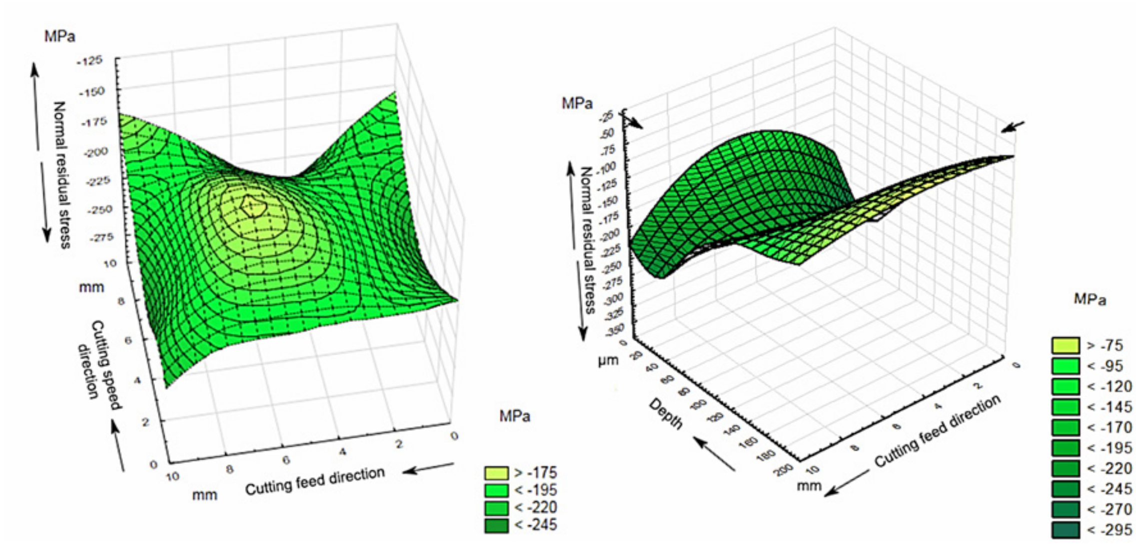
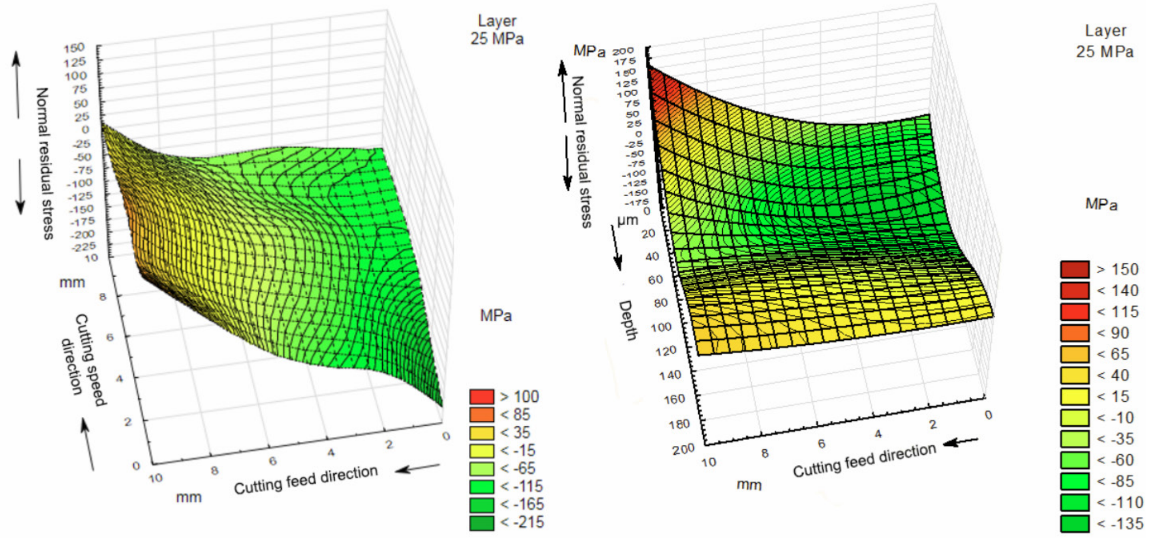


Fig. 8 Measured values of normal residual stress into the depth after finishing (cutting conditions: v_c = 250m.min^{-1}, f = 0.05mm, a_p = 0.5mm)



a) b)

Fig. 9 Simulation of distribution of normal residual stress in various directions after roughing operations (cutting conditions: v_c = 200m.min^{-1}, f = 0.35mm, a_p = 3mm)



a) b)

Fig. 10 Simulation of distribution of normal residual stress in various directions after roughing operations (cutting conditions: v_c = 250m.min^{-1}, f = 0.075mm, a_p = 0.5mm)

many causes it can be improved by finishing operations or other treatments of surface.

Simulation of distribution of normal residual stress, for finishing operations, shows rise of tensile residual stress with direction of cutting feed (Fig. 10a). It can be caused by gradual increase in the volume of extruded material by the cutting tool due to relatively low cutting depth (a_p=0.5mm). Simulation into the depth (Fig. 10b) shows increased appearance of residual stress with tensile character. The range of normal residual stress on surface layer and deepest measured layer (200μm) is again very large. It is not suitable for final state of part. This problem can be solved via improvements of surface and subsurface layers by applying other finishing methods such as shot peening.

6. Discussion and conclusion

Measurement of residual stress is a topical problem by which we can predict the functional properties of parts. Some destructive methods allow to measure residual stress in subsurface layers, but due to its principles, they are not so detailed and accurate. Principles of X-ray diffraction can measure the residual stress at microstructure levels and so push the boundaries of measuring in this area. Automation of this device allows to measure not only individual points or small lines, but we can obtain cloud of points or maps of residual stress distribution. Through the mathematical-statistical procedures, we can minimalize the number of required measured points and so create sufficiently precise simulations of residual stress distribution. And finally, from this data we can see a detailed course of residual stress depending on cutting conditions into the depth of machined material. In presented experiments, we can see that the course of residual stress can be changed dramatically in the range of several micrometres. By applying these simulations, we can design suitable cutting conditions and so to minimalize the rise of potential errors, which can cause formation of cracks and subsequent failures.

Acknowledgments

This article was funded by the University of Zilina project OPVav-2009/2.2/04-SORO number (26220220101) – “Intelligent system for nondestructive technologies on evaluation for the functional properties of components of X-ray diffraction.”

References

[1] GRIFFITHS, C. A.: *Closing the Loop on Product Integrity on Bearings with Engineered Source Approval, Split Ballbearing Features,* Company Newsletter, 1989.

[2] GANEV N., KRAUS I.: *X-ray Diffraction Measurement of Residual Stresses*, Material Structure, vol. 9, No. 2, 2002.

[3] TIITTO, K. et al.: Testing Shot Peening Stresses in the Field*, The Shot Peener,* vol. 4, 1991, ISSN 1069-2010.

[4] CEP, R., PETRU, J., ZLAMAL, T., VALICEK, J., HARNICAROVA, M., JANASEK, A., CEPOVA, L.: *Influence of Feed Speed on Machined Surface Quality.* Proc. of 22nd Intern. Conference on Metallurgy and Materials, pp. 1033-1038, 2013, ISBN 978-808729441-3

[5] PALA, Z., KOLARIK, K., GANEV, N., CAPEK, J.: *Study of Residual Stress Surface Distribution on Laser Welded Steel Sheets,* 51st Annual of the Intern. Scientific Conference on Experimental Stress Analysis, EAN 2013

[6] SAPIETOVA, A., SAGA, M., NOVAK, P.: Multi-software Platform for Solving of Multibody Systems Synthesis, *Communications - Scientific Letters of the University of Zilina*, vol. 14, No. 3, 43-48, 2012, ISSN 1335-4205.

[7] PREVEY, P. S.: *X-ray Diffraction Residual Stress Techniques*, Metals Handbook 10, Metals Park: American Society for Metals*,* pp. 380-392, 1986.

[8] FITZPATRICK, M. E. et al.: Determination of Residual Stresses by X-ray Diffraction. *Measurement Good Practice,* No. 52, 2005, National Physical Laboratory, Teddington.

[9] STANCEKOVA, D., SEMCER, J., HOLUBJAK, J., DRBUL, M.: Machinability of Nano-structured Biomaterials for Dental Implants. *Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3A, 2014, pp. 96-100, 2014, ISSN 1335-4205.

[10] BREZINA, M., KUPCA, L.: *Stress Corrosion Cracking of Austenitic Stainless Steel in a Nuclear Power Plant Environment,* Environment inducted cracking of materials, Department of Structural Analysis : VUJE Inc, Trnava, pp. 431-436, 2008.

[11] CZAN, A., TILLOVA, E., SEMCER, J., PILC, J.: *Surface and Subsurface Residual Stresses after Machining and Their Analysis by X-ray Diffraction*. Communications – Scientific Letters of the University of Zilina, vol. 15, No. 2, 2013, pp. 69-76, ISSN 1335-4205

[12] NAPRSTKOVA, N., HOLESOVSKY, F.: *Admeasurement of Grinding Wheel Loss at FPTM*. 24th Intern. colloquium Advanced manufacturing and repair technologies, pp. 159-164, 2007, ISBN 978-80-7194-962-6.

[13] VASILKO, K., PILC, J.: New Technological Knowledge of the Rotary Turning Tool, *J. Manufacturing Technology*, vol. 13, No. 4, December 2013, pp. 571-575, ISSN: 1213-2489.

[14] CEP, R., JANASEK, A., PETRU, J., CEPOVA, L., CZAN, A., VALICEK, J.: Hard Machinable Machining of Cobalt-based Superalloy, *Manufacturing Technology* XIII/13, UJEP : Usti n. Labem, pp. 226-231, 2013, ISSN 1213-2489.

[15] CZAN, A., SAJGALIK, M., MARTIKAN, A., MRAZIK, J.*: Observation of Dynamic Processes in Cutting Zone when Machining Nickel Alloys. Communications - Scientific Letters of the University of Zilina*, vol. 16, No. 3A, pp. 161-168, 2014, ISSN 1335-4205.

Alzbeta Sapietova - Vladimir Dekys *

USE OF MSC.ADAMS SOFTWARE PRODUCT IN MODELLING VIBRATION SOURCES

This paper presents the methodology of rotating device modelling in order to facilitate correct evaluation of selected, usually malfunction (degraded) states of rotating machines of mechanical systems, especially of gearboxes. The analysis was performed using mathematical models utilising MSC.ADAMS software product to solve the dynamics of a system of bodies. The correctness of the proposed method is verified by qualitative comparison of simulation results with the presented symptoms of malfunction states and also by comparison with real measurement results from technical practice.

Keywords: *Dynamics, rotating machine, gear mesh frequency, frequency spectrum, MSC.ADAMS.*

1. Introduction

Correct operation of machinery inevitably requires alignment of machine geometry in hundredths of millimetres (depending on the shaft speed). Shaft misalignment results in increased reactions, which the bearings must absorb, it induces mechanical looseness, damages to seals, allows fluids and impurities to enter the bearings, which considerably decreases the machine operating life and, consequently, the operating life of the whole machinery. Carrying out purposeful maintenance of machines requires a correct interpretation of malfunction causes. Very frequent claims are that malfunction is caused by the bearing, seals or others. However, these are the consequences and not the initial causes of malfunction. In general, there is a lack of correlation between causes and consequences of malfunctions [1].

One of non-destructive methods allowing identifying and consequently eliminating the problem is vibration diagnostics of rotating machinery. It uses vibrations, which are generated by machines in operation, as a source of information about the way of its working. Vibration diagnostics is also an important tool of modern productive and proactive methods of machinery maintenance. Using this tool, maintenance of machines is planned according to the actual state, and, therefore, many useless preventive revisions are eliminated, which leads to important savings of spare parts and time for machinery repair. It gives information about regularly monitored machines, by which we can prolong the planned periods of shutdowns or prepare in advance for the repair of specific monitored machine node etc. [2, 3, 4 and 5].

However, the above is diagnostics of malfunction caused by experimental method right in the operation. In the recent decades many software products solving dynamic problems of rotating machines were developed, therefore, computer simulation and building of virtual prototypes (VP) is nowadays an essential part of every technical solution [6]. Often, there is a question how to create a VP so that the results of its mathematical solution correspond to physical values or values acquired by experimental measurement.

The paper presents the creation of a virtual prototype of gearing, and this prototype is used to generate and analyse dynamic effects measurable in the place of rotary links between the shaft and the frame. These effects are registered as vibration velocity. We then use the generated courses of vibration velocity to obtain velocity spectrum, and this spectrum will be compared with the spectra from diagnostic measurements. In this way we can verify the created virtual model, adjust its appropriate variables to select a suitable solver in MSC.ADAMS software.

2. Manifestations of the gearbox normal spectrum

The gearbox normal spectrum (gearbox without damage) is characterised by the presence of revolution frequencies of the two shafts f_1, f_2 (with gear wheels t_1, t_2) and tooth frequency (including

* **Alzbeta Sapietova, Vladimir Dekys**
Department of Applied Mechanics, Faculty of Mechanical Engineering, University of Zilina, Slovakia,
E-mail: alzbeta.sapietova@fstroj.uniza.sk

its higher harmonics). Using the distance of the sideband Δf we can determine the source responsible for generating the tooth frequency (Fig. 1). However, sidebands are usually not pronounced. Normal speed spectrum is usually characterised by the following peaks (Fig. 1) [1 and 7]:

- revolution frequency of individual shafts,
- generating the tooth frequency $f_{GMF} = f_1 \cdot t_1 = f_2 \cdot t_2$ and its multiples $2xf_{GMF}$, $3xf_{GMF}$ with sidebands. Sidebands may not be significantly pronounced and their distance from the tooth frequency (and its harmonics) very often coincides with the revolution frequency of one of the shafts,
- however, tooth frequencies and their multiples are not significant, often their amplitude decreases with increasing multiple of revolution frequency, and eigenfrequency of one of the wheels does not appear in the spectrum.

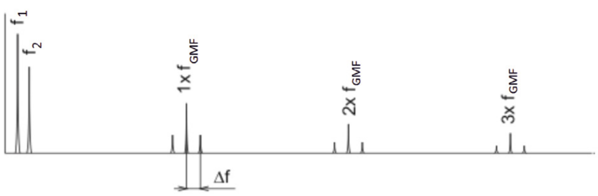


Fig. 1 Normal spectrum of the gearbox, Δf is the sideband distance.

It should be noted that the term "normal spectrum" is hereby incorporated in an intuitive sense [8 and 9]. Normal spectrum corresponds to a state where the object performs required functions, even if it features ongoing degradation processes that usually generate other peaks in the spectrum. The term "normal spectrum" is then equated to a descriptive expression "time-stable nature of spectrum (often determined using periodically repeated measurements), with peak levels not exceeding certain limits".

3. Modelling a spur gearing

We used the MSC.Adams/View software environment to create a VP (Fig. 2) for modelling and error analysis of the gear transmission mechanism. The VP consists of solid bodies, the shaft is modelled by a geometric element "cylinder" and the spur gearing is imported from the 3D parametric modelling software, Pro Engineer.

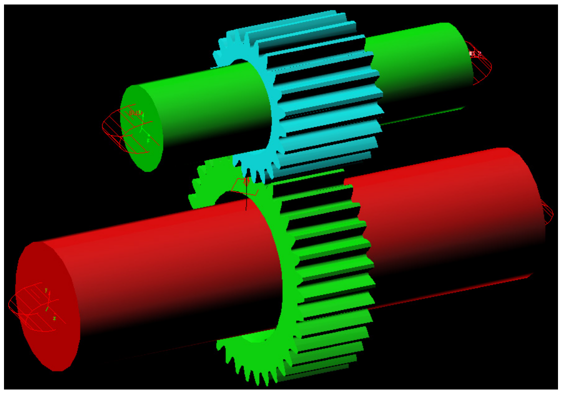


Fig. 2 Spur gearing virtual prototype in MSC. ADAMS

4. Shaft mounting on a bushing elastic fastener

The gearing VP was attached to the frame in the MSC. ADAMS environment using a bushing force bond, which is force based connections of members. The advantage of this link is the fact that it can simulate the system elasticity (flexibility). It allows connection of two bodies. Bushing has 6 force degrees of freedom. Three forces and three moments act in the axes of the link coordinate system. The software calculates components of forces and moments. In general, the bushing properties are defined by six stiffness coefficients, six damping coefficients and optional prestress condition. Instead of the prestress value the initial conditions can be defined also by initial displacement [10].

It follows from the above that between two bodies linked by bushing, there are applied three forces: F_x, F_y and F_z, and three moments: T_x, T_y and T_z, acting in the axes of a rectangular Cartesian coordinate system. These are a linear function of translational and rotational movement between the two coordinate systems of these bodies.

The general mathematical formulation of the bushing force element is [10]:

$$\begin{bmatrix} F_X \\ F_y \\ F_Z \\ T_X \\ T_Y \\ T_Z \end{bmatrix} = - \begin{bmatrix} k_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & k_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & k_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & k_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & k_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & k_{66} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ \alpha \\ \beta \\ \gamma \end{bmatrix} - \begin{bmatrix} c_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix} \cdot \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\alpha} \\ \dot{\beta} \\ \dot{\gamma} \end{bmatrix} + \begin{bmatrix} F_{X0} \\ F_{y0} \\ F_{Z0} \\ T_{X0} \\ T_{Y0} \\ T_{Z0} \end{bmatrix} \quad (1)$$

where:
F_x, F_y and F_z are the moving forces in the coordinate system of the J marker,
x, y and *z* are the bushing deformation vector components in the coordinate system of the J marker,
$\dot{x}, \dot{y}, a\ \dot{z}$, is the deformation vector derivation by time, i.e. velocity of displacements,
F_{x0}, F_{y0} and F_{z0}, (T_{x0}, T_{y0} and T_{z0}), are the values of constant prestress components of forces (moments) in the coordinate system of the J marker,
T_x, T_y and T_z are the moments of forces in the coordinate system of the J marker,
α, β, and *γ* are the small rotation angles of the I marker against the J marker,
$\dot{\alpha}, \dot{\beta}, a\ \dot{\gamma}$ are the values of angular velocities of the I marker coordinate system expressed in the coordinate system of the J marker,
k_{11} - k_{66} are the values of stiffness,
c_{11} - c_{66} are the values of damping.

Bushing applies to the balance of the forces and moments of the I marker to the J marker as follows:

$$\mathbf{F}_j = -\mathbf{F}_i, \tag{2}$$

$$\mathbf{T}_j = -\mathbf{T}_i - \delta \times \mathbf{F}_i, \tag{3}$$

where **δ** is the instantaneous deformation vector between the markers from J to I. While the force applied to the J marker is equal and oppositely oriented to the force applied on the I marker, the torque in the J marker is usually not equal to the torque in the I marker, because the moment arm is a function of the bushing element deformation, which can be different in connected bodies [11].

Based on the actual (real) mounting, the shaft is mounted in two bushings, where on one side the axial parameter values are zeroed (axially free), thereby achieving the effect of a possible axial displacement in the mounting.

5. Force link in the gearing

The type of link used in the gearing is persisting contact. This force link is represented by a model in the form of a nonlinear spring-damper system, in which the stiffness takes into account the elasticity of contacting surfaces and damping reflects energy dissipation, where the bodies are not separated after colliding [12]. The calculation of contact forces is carried out separately for each point, and the resulting effect is their sum. Reaction forces generated by contact during simulation are an important aspect for assessing the behaviour of the gear transmission. Theoretically, the forces in the contact of the teeth in engagement should have a common bearer lying on a joint tangent of root circles (Fig. 3). However, the lines of action of these reactions in the VP work cycle are often not identical (Fig. 4). It should be noted that that the reactions are generated by the contact function, and, therefore, correct adjustment of the contact is important. The goal is to achieve the effect when the gear wheel surfaces roll away perfectly upon each other, which increases the VP's ability to react to smaller irregularities in the surface geometry.

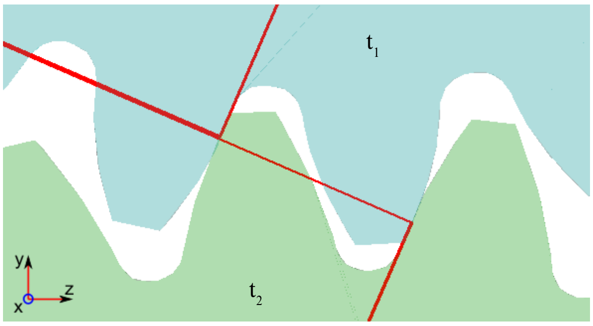


Fig. 3 The lines of action of reactions in the gearing - the ideal state

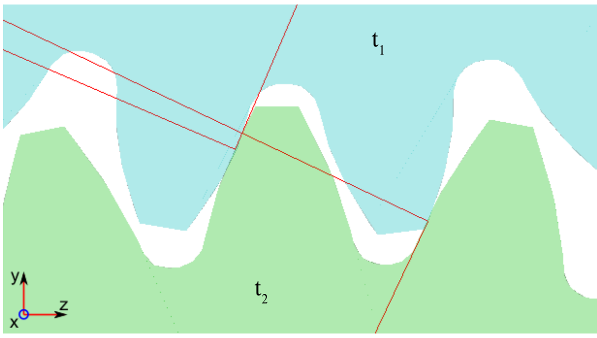


Fig. 4 The lines of action of reactions in the gearing - the lines of action do not overlap

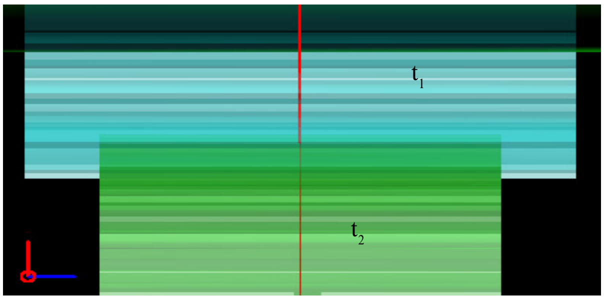


Fig. 5 Side view of the VP, reactions in the monitoring of their perpendicularity to the axis of rotation, reaction effects (red straight line), reactions lie on one bearer

When observing the contact forces between the tooth system (gearing), it is further necessary to ensure their compliance with perpendicularity to the axis of rotation. Otherwise, misalignment of the shafts is present in the frequency spectrum. Then we can talk about a possible misalignment caused by incorrect setting of bushings. Figure 5 shows an illustrative example of an approximately ideal state. If an angle is formed between reactions, then it is appropriate to debug the VP using one of the sensitivity or optimisation analyses (DS/DOE/DO) of MSC. ADAMS software [13 and 14].

6. Input parameters

For the VP of spur gearing we determined the following parameters and dependencies:

- The number of teeth is predetermined by this device's geometry and configuration, this will be the key parameter for detecting tooth frequencies and transmission ratio: t_1 = 23, t_2 = 36.
- Input revolution frequency f_{in} = 25Hz, derived from the value of input revolutions 1500 rpm. This will be the key parameter for detecting tooth frequencies.
- Output revolution frequency f_{out} is derived based on the input revolution frequency and the transmission ratio $f_{out} = f_{in} \cdot t_1/t_2$= 15.9Hz.
- Input angular velocity ω_1 is a parameter derived from the input revolutions. Transients were removed using the function STEP (step(time, 0, 0, 1, 157)), thus defining a smooth start and subsequent maintaining of the angular velocity at a constant value.
  $\omega_1 = 2\pi f = 2 \cdot 3.14 \cdot 25$ = 157rad/sec = 9000deg/sec,
  angular velocity of the shaft of the output gear wheel is then ω_2 = 5750deg/sec.
- The tooth frequency f_{GMF} corresponding to the pair of wheels is dependent on the revolution frequencies of corresponding shafts and wheels mounted to them.
- The pair of wheels will generate the first tooth frequency $1xf_{GMF} = 23 \cdot f_{in} = 36 \cdot f_{out}$ = 575Hz and other harmonic multiples $2xf_{GMF}$, $3xf_{GMF}$, $4xf_{GMF}$.
- Simulation parameters: simulation time – 5s, simulation time step – 0.0005s, solver type Newmark – it works better than the standard default Gauss solver when using force links and small time steps,
- Parameters of contact between the wheels remained almost the same, except for stiffness 1.10^7 [N/mm] and penetration depth 1.10^{-3} [mm] in order to achieve a minimum overlap of geometry of teeth in engagement and thus increased surface sensitivity to small changes (damage).

7. Sensitivity analysis of mounting parameters in bushing links

In this section we present sensitivity analysis to determine optimal parameters of stiffness K and damping C in bushing links.

First, we used the Design Study (DS) method based on a change in the stiffness parameter K in bushing links to monitor the size of constant angular velocity of the output shaft ω_2. It should be understood that DS solving was adjusted to finding an effective value of ω_2 (RMS). RMS values were determined only from the revolution-stable time section in post-processing.

Damping was set to a constant value $C=1x10^3$[Ns/mm]. It follows from the solution in Table 1 that stiffness K has a minor effect on the course of the output angular velocity ω_2 at constant damping C. For this reason we carried out advanced analysis using the Design of Experiment (DOE) method of searching for effective combinations of parameters C and K [15].

DS sensitivity analysis for stiffness parameter K for target ω_2=5750deg/s — Table 1

trial	achieved ω_2 [deg/s]	stiffness K [N/mm]
1.	611.1	1.10^3
2.	4231.3	1.10^4
3.	5767.1	1.10^5
4.	5757.9	1.10^6
5.	5756.8	1.10^7
6.	5771.7	1.10^8
7.	5769.1	1.10^9

Table 2 shows several selected, more significant combinations achieved within the DOE method solving, i.e. mapping the effective combinations of parameters K and C (the values in the table are RMS from the stabilised part of the courses after 2.5s, i.e. after the transient aftermath). Of the total 49 combinations, trial No. 17 has the RMS value closest to the target value of ω_2, both in the value as well as in the quality of the course.

DOE method for input variables K and C, and for target value of ω_2 — Table 2

trial	achieved ω_2 [deg/s]	stiffness K [N/mm]	damping C [Ns/mm]
4.	5756.7	1.10	1.107
5.	5757.9	1.10	1.106
17.	5751.1	1.103	1.105
18.	5756.6	1.10^3	1.10^6
25.	5754.6	1.10^4	1.10^6
32.	5753.9	1.10^5	1.10^6
39.	5755.2	1.10^6	1.10^6
46	5754.2	1.10^7	1.10^6
49	5756.8	1.10^7	1.10^9

Monitoring the velocity on the system output included also monitoring the frequency spectra of the tooth frequency. Since this evaluation requires to be carried out by the solver itself, we can present in this case that high damping causes suppression of amplitudes in the frequency spectrum.

8. Monitored parameters

The state of the evaluated transmission system can be determined by observing the properties that are specific for this system configuration. These specific properties are time and frequency spectra of vibration velocities obtained from the time courses of vibration on the shaft bearing housings. Figure 6 shows the course of velocities on the output bearing observed in radial directions Y (vertical) and Z (horizontal). Figure 7 shows a detail of the course of these velocities from the stable part of the spectrum.

Figure 8 shows the FFT (fast Fourier transform) frequency spectrum corresponding to the above presented course of velocities in the radial direction Z. The course shows a dominant first tooth frequency. The spectrum also includes its harmonic multiples with insignificant sidebands.

9. Practical example of normal spectrum

Figure 9 shows a normal spectrum of vibration velocity on the output shaft of a 3-speed gearbox of a cement mill. The spectrum

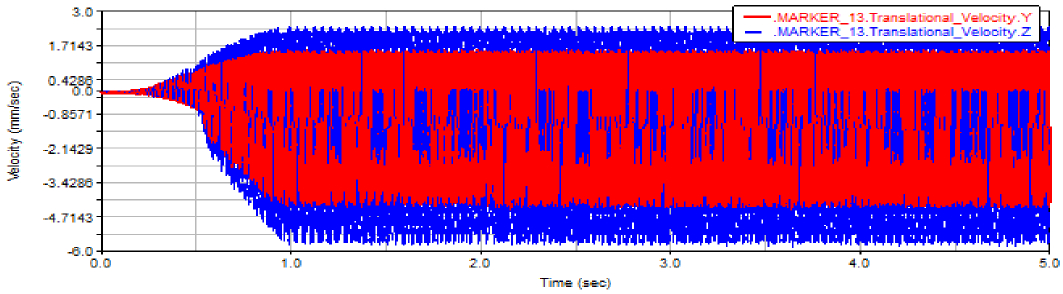


Fig. 6 Velocities in the radial direction on the output bearing

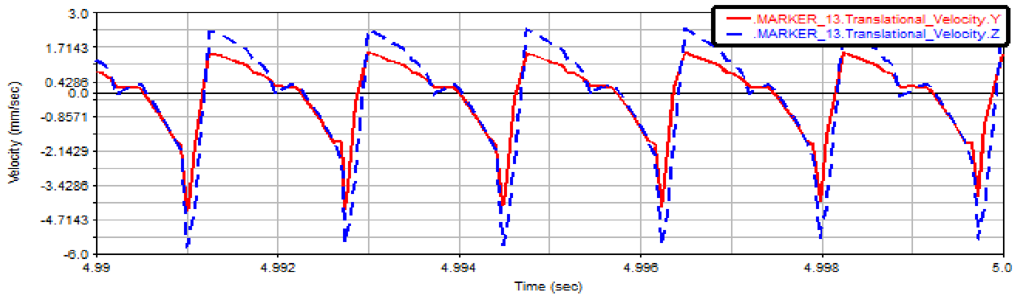


Fig. 7 Detail of velocity in the radial direction on the output bearing

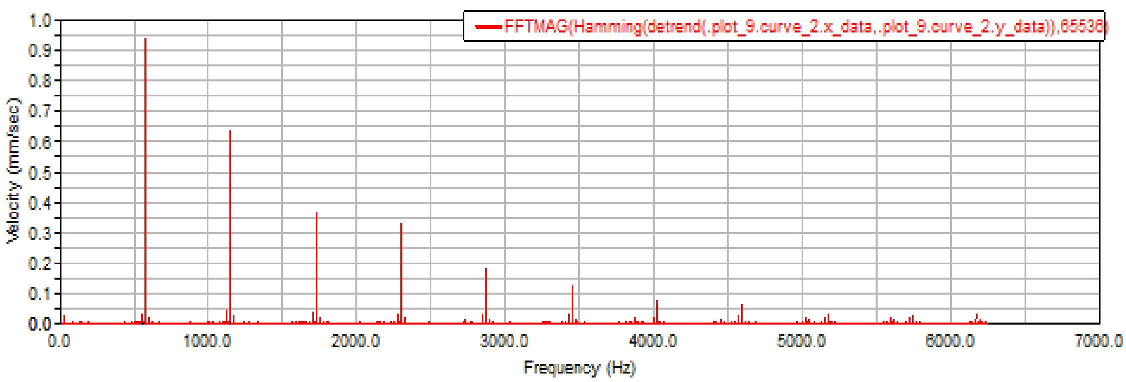


Fig. 8 FFT velocities in the radial direction on the output bearing

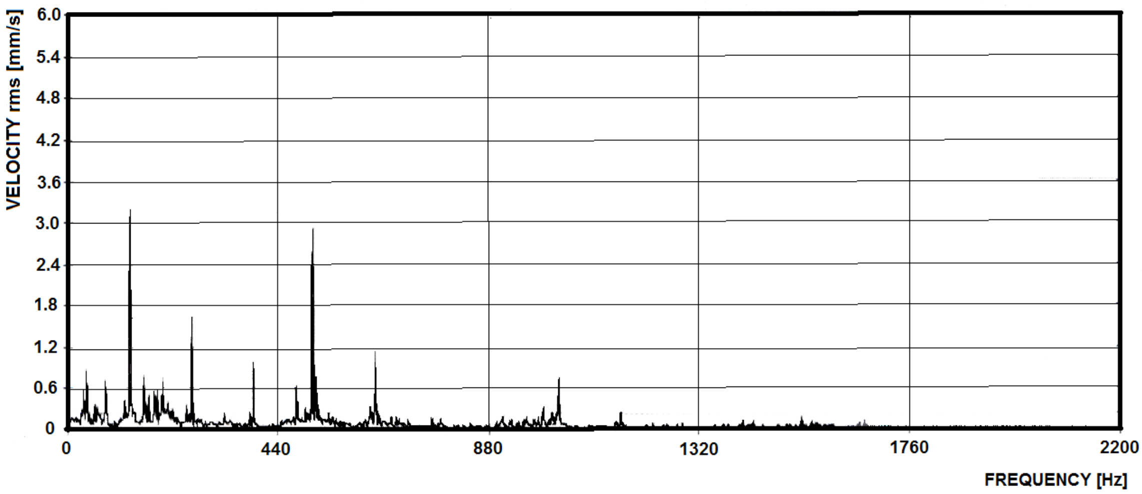


Fig. 9 Normal FFT spectrum of a raw mill gearbox, characterised by decrease in higher multiples of tooth frequencies

includes lower, but significant multiples of f_{GMF3} (40.4Hz, 120Hz, 160.4Hz, output gear), two dominant peaks at multiples of f_{GMF2} (130.4Hz, 259.7Hz), and significant multiples of the tooth frequency f_{GMF1} (513.3Hz and 1026.6Hz, input gear). This measurement clearly shows a decrease of amplitudes of harmonic tooth frequencies for individual gear speeds.

10. Conclusion

The presented methodology of modelling a system of gear wheels seems satisfactory, since the expected results were achieved in both time spectrum (Fig. 6) and frequency spectrum (Fig. 7).

When modelling a system of gear wheels it is necessary to achieve that tooth surfaces perfectly roll away over each other, which increases the virtual prototype sensitivity to minor surface imperfections. This means we need to employ sensitivity analysis to correctly set both the lines of action of reactions of contact in the gearing and also optimal parameters of stiffness and damping in the links. The shaft is mounted in two bushing-type links – one of them has zero axial values of the parameters, thereby achieving the effect of a possible axial displacement in the shaft mounting. Parameters of contact between the gearing remained almost constant, except for stiffness $1x10^7$ [N/mm] and penetration depth $1x10^{-3}$ [mm] (achieved minimal overlap of geometries) [16 - 17].

In the created VP, the courses of kinematic parameters and FFT spectra were qualitatively identical to the predicted and practically measured ones. This created VP of a gearing may further serve to modelling failure states in the gearing, such as misalignment, damage to teeth, etc. This will allow analysing and mutually comparing, for example, the manifestations of undamaged and damaged gearboxes.

Acknowledgement

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0736-12 and the Slovak Grant Agency VEGA 1/0983/15.

References

[1] BILOSOVA, A., BILOS, J.: *Vibration Diagnostics (in Czech),* Ostrava, 2012.
[2] KREIDL, M., SMID, R.: *Technical Diagnostics (in Slovak),* BEN: Praha 2006, ISBN 80-73000-158-6.
[3] KREIDL, M., et al.: *Diagnostic Systems (in Czech),* CVUT: Praha 2001, ISBN 80-01-02349-4.
[4] SHEFFER, D., GIRHARD, T.: *Practical Machinery Vibration Analysis and Predictive [*
[5] BAKOWSKI, A., RADZISZEWSKI, L., ZMINDAK, M.: Analysis of Uncertainty in Pressure Measurement with a Piezoelectric Transducer for a Combustion Engine. *Applied Mechanics and Materials*, vol. 827, 2015, pp. 77-82.
[6] VAVRO, J., VAVRO, J., jr., VAVROVA, A., KOVACIKOVA, P.: Dynamic Analysis of Winding Mechanisms for Manufacturing of Raw Tyres. *Hutnicke listy*, No. 64, pp.189-192.

[7] DEKYS, V., NOVAK, P.: Report: Condition Monitoring of Bearings in Test Laboratory PSL. University of Zilina: SjF: Department of Applied Mechanics, No. 1, June 2005.

[8] BERRY, J., E.: *Intensive Vibration Diagnostics.* SKF Condition Monitoring, Zaltbommel, June 1985, Intensive condition monitoring course: Technical Associates: Charlotte, 1985.

[9] ADAMS, M., L.: *Rotating Machinery Vibration from Analysis to Troubleshooting,* New York: Marcel Dekker, 2001, ISBN 0-8247-0258-1.

[10] KIVINIEMI, T., HOLOPAINEN, T.: *Modelling of Flexible Members for Simulation of Vehicle Dynamics* [online], 1999, [cit. 2015-04-07], http://www.vtt.fi/inf/julkaisut/muut/ 1990s/valb 424.pdf.

[11] KOPAS, P., VASKO, M., HANDRIK, M.: Computational Modeling of the Microplasticization State in the Nodular Cast Iron. *Applied Mechanics and Materials,* 2014, vol. 474, pp. 285-290, ISSN 16609336.

[12] HYBEN, B., SAPIETA, M.: *Multi-software Solution of Optimization for Selected Parameters Machinery in Terms of their Mechanical Properties*. Computational mechanics 2012, Pilsen: University of West Bohemia, 2012. ISBN 978-80-261-0157-4 - CD-ROM, [2] s.

[13] SAGA, M., KOPAS P., VASKO, M.: Some Computational Aspects of Vehicle Shell Frames Optimization Subjected to Fatigue Life. *Communications - Scientific Letters of the University of Zilina,* vol. 12, No. 4, 2010, pp. 73-79.

[14] SAPIETOVA, A., SAPIETA, M., HYBEN, B.: Sensitivity Analysis Application for Multibody System Synthesis. *Applied Mechanics and Materials,* vol. 420, 2013, pp. 68-73, Online available since 2013/Sep/27 at www.scientific.net © (2013) Trans Tech Publications, Switzerland doi:10.4028/www.scientific.net/AMM.420.68 ISSN: 1660-9336.

[15] ZMINDAK, M., RADZISZEWSKI, L., PELAGIC, Z., FALAT, M.: Fem/bem Techniques for Modelling of Local Fields in Contact Mechanics. *Communications - Scientific Letters of the University of Zilina*, vol. 17, No. 3, 2015, pp. 37-46.

[16] SAGA, M., VASKO, M., KOPAS, P., JAKUBOVICOVA, L.: Numerical Algorithm for Beam Residual Stress Identification. *Communications - scientific letters of the University of Zilina*, vol. 16, No. 3A, 2014, pp. 13-18, ISSN 1335-4205.

[17] KOPAS, P., SAGA, M.: In-phase Multiaxial Fatigue Experimental Analysis of Welded Cylindrical 6063-T66 Aluminium Alloy Specimens. *Manufacturing Technology*, vol.13, No. 1, 2013, pp. 59-64, ISSN 1213-2489.

Slavomir Hrcek - Robert Kohar - Stefan Medvecky - Branislav Krchnavy - Marcel Caco *

DETERMINATION OF THE OPTIMUM PRELOAD OF ROLLER BEARINGS IN TERMS OF THEIR LIFE USING FEM

The objective of this paper is to describe the procedure of correct determination of the optimum preload of tapered roller bearings in terms of their fatigue life, using the procedure laid down in ISO 281, and by analysis of contact stress state (analytically and by the FEM) with application of the Lundberg-Palmgren theory. The procedures will be applied to a specific example of the fit of the output shaft of a gearbox in a pair of tapered roller bearings designed specifically for this application.

Keywords: *Preload, tapered roller bearings, radial load distribution, fatigue life, Lundberg - Palmgren, FEM.*

1. Introduction

Roller bearings are one of the basic structural components of machines, used to fit the rotating parts in frames or housings. They capture and transmit reaction forces from external loading. Typical applications of rolling bearings are fits of shafts in gearboxes and transmissions. However, there are also special fits where the shaft is fitted in tapered roller or ball angular contact bearings, and proper preload must be determined to ensure their proper function [1 and 2]. The proper preload inferred between a pair of tapered roller or ball bearings with angular contact has a great impact on the service life of both the bearings forming the fit [3]. Improperly determined preload can significantly reduce their life or even damage the whole transmission. Figure 1 shows a typical shaft fit in two tapered roller bearings in the "O" configuration [4 and 5].

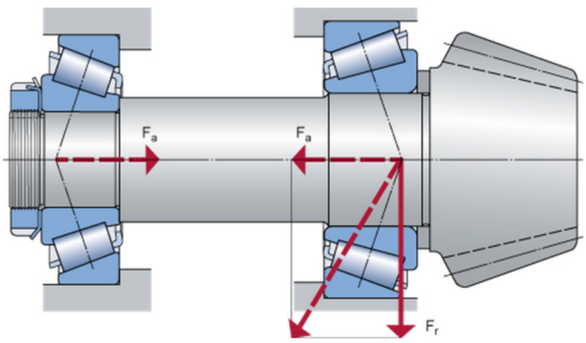


Fig. 1 Typical shaft fit in a pair of tapered roller bearings

Proposal of proper preload and its impact on bearing life will be shown using the example of the output shaft fit deposit in an 8-speed gearbox in a pair of tapered roller bearings in the "O" configuration. The gearbox output shaft is fitted, due to dimensions, in atypical single-row tapered roller bearings 32008 with the basic parameters as follows: $C_r = 79.784\ kN$, $e = 0.318, Y = 1.881$, the inner ring diameter d = 40 mm, the outer ring diameter D = 80 mm, the bearing width (broadness) B = 21.8 mm, nominal contact angle α = 12°, with the number of rolling elements Z = 17 and effective length L_{we} = 14.3 mm. The rolling element profile geometry consists of an unmodified logarithmic profile. Figure 2 shows a scheme of the 8-speed gearbox, the output shaft is fit in bearings marked A5 a B5.

The aim was to design the proper fit of the output shaft and determine optimal bearing preload for the gearbox working cycle specified in Table 1.

2. Calculation of equivalent dynamic load on the gearbox output shaft bearings

The load acting on rolling bearings constituting the fit of the output shaft is based on the forces acting on the gearwheel at the end of the shaft, as shown in Fig. 3 [6]. At the end of the shaft, gearwheel is used with straight teeth in order to show more clearly the difference in the preload influence on the life calculated through equivalent dynamic load. The same is true for gears with a small angle of teeth inclination β.

* **Slavomir Hrcek, Robert Kohar, Stefan Medvecky, Branislav Krchnavy, Marcel Caco**
Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: slavomir.hrcek@fstroj.uniza.sk

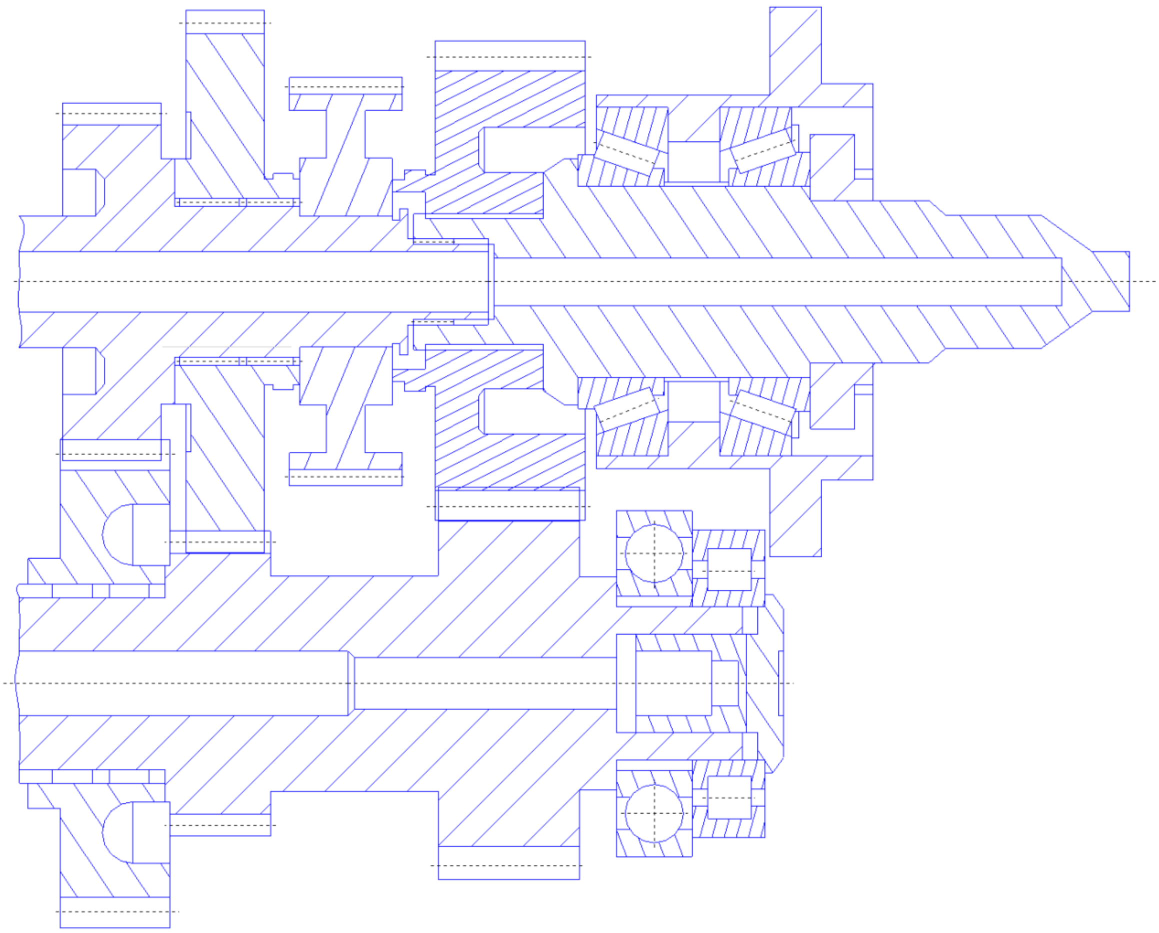


Fig. 2 Scheme of the 8-speed gearbox

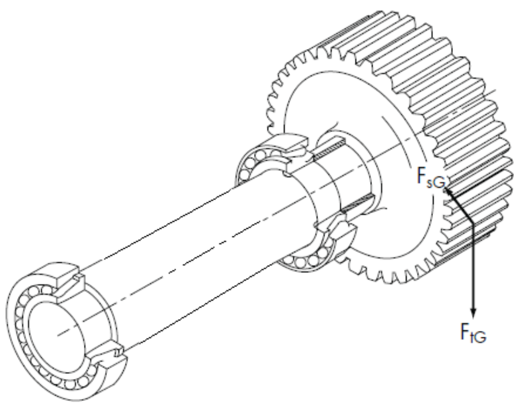


Fig. 3 Output shaft fit

Force F_{tG} is the tangential component of forces from the gears, force F_{sG} is the radial component of forces from the gears and this component is transmitted by rolling bearings.

From the load by the teeth, and using moment equilibrium, we calculated the response per bearings in "O" configuration, where the distance of the force acting to the bearing A is a=17.67 mm, and the effective bearing spread between the bearings A and B is b=53.5 mm. Sizes of response for each working cycle are shown in Table 1.

Equivalent load for tapered roller bearings according to ISO 281 [7] is calculated using the following equations:

$$P_e = F_r \qquad \text{for } F_a/F_r \leq e \tag{1}$$

$$P_e = 0.4\,F_r + YF_a \qquad \text{for } F_a/F_r > e \tag{2}$$

where e is the limit value of the ratio of the axial to the radial load, Y is the coefficient of dynamic axial load. Since the output shaft fit consists of a pair of identical tapered roller bearings, we will further consider only the more loaded bearing A, which is closer to the gearwheel.

Working cycle and response forces in bearings Table 1

Gear	Torque	F_{sG}	R_{Arad}	R_{Brad}	Time
#	Nm	N	N	N	hours
1	500	42 521.6	56 565.7	14 044.1	0.44
2	700	35 865.4	47 711.0	11 845.6	5.96
3	700	24 816.6	33 013.0	8 196.4	12
4	700	18 807.1	25 018.7	6 211.6	12.08
6	700	11 493.0	15 289.0	3 795.9	20.21
7	700	9 618.2	12 794.9	3 176.7	14.22
8	700	8 294.3	11 033.8	2 739.5	14.22
					79.13

Table 2 shows the calculation of equivalent load according to IS0 281 based on the external radial and axial dynamic loads. The individual preload values correspond to the levels of the axial load, then the center of the table presents in a matrix form the above-mentioned values of partial equivalent loads P_1, P_2, etc., calculated by the ratio of the radial and axial loads from equations 1 and 2. Subsequently, we calculated the resultant equivalent load P_e for a working cycle according to equation 3 for each preload value.

Since the gearbox operates in different modes, it is necessary to calculate the resultant equivalent load at the same revolutions (gearbox operates at the speed of 5000 rpm at all stages of the working cycle), using the following equation:

$$P_e = \sqrt[n]{P_1^n\left(\frac{t_1}{\sum t_i}\right) + P_2^n\left(\frac{t_2}{\sum t_i}\right) + \cdots} \quad (3)$$

Working cycle and equivalent load in the bearing A Table 2

Preload (mm)	0	0.05	0.1	0.15	0.2	0.25
R_{Arad}\ R_{Aax} (N)	0.0	4 506.8	9 735.3	15 275.9	21 029.4	26 946.7
56 565.7	56 565.7	56 565.7	56 565.7	56 565.7	62 200.6	73 335.9
47 711.0	47 711.0	47 711.0	47 711.0	47 831.4	58 658.7	69 794.1
33 013.0	33 013.0	33 013.0	33 013.0	41 952.2	52 779.5	63 914.9
25 018.7	25 018.7	25 018.7	28 327.9	38 754.4	49 581.7	60 717.1
15 289.0	15 289.0	15 289.0	24 436.0	34 862.6	45 689.9	56 825.3
12 794.9	12 794.9	13 599.2	23 438.4	33 864.9	44 692.2	55 827.6
11 033.8	11 033.8	12 894.7	22 733.9	33 160.5	43 987.8	55 123.2
Pe (N)	27 359.5	27 434.4	30 121.4	37 830.1	48 441.5	59 480.7

3. Calculation of optimal preload using FEM

We used CAD system PTC/Creo to create a 3D model of the gearbox output shaft fit. Subsequently, the model was imported into the software environment ANSYS/Workbench, where static structural analysis was created for the needs of analysis to determine the optimal preload. The model networking used standard elements from Ansys software library, the volume networking used the SOLID185 element and the contact pairs of rolling elements – raceways used the CONTA174 and TARGE170 elements. The entire model contained 2,635,001 elements and 807,815 nodes. Contacts between the rolling elements and raceways were defined as frictional with the coefficient of Coulomb friction 0.1. Bearing rings fitting in the housing and on the shaft was defined by frictionless contact with the respective value of interference of the inner ring on the shaft. Since the model of fitting is symmetrical, we defined the plane of symmetry x-y to simplify the calculation. Boundary conditions (Fig. 4) were defined at the face of the housing largest fit by removing degrees of freedom in the direction of the x and y axes, the radial load from the gearwheel on the output shaft was applied at the end fit of the shaft at the point of the gearwheel force in the y axis direction. Definition of preload was prescribed by displacement at the faces of the outer bearing rings uniformly against each other in the x axis direction.

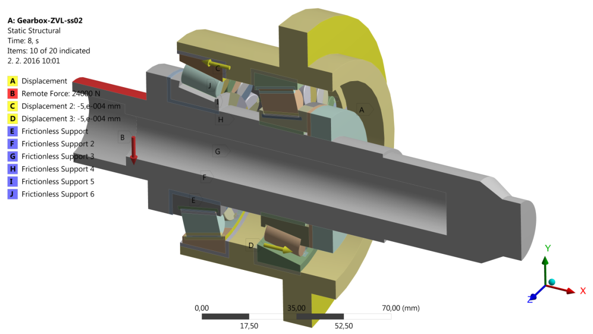


Fig. 4 Definition of boundary conditions and load

From the results of the static structural analysis, using the finite element method, we first conducted analysis of the load distribution on individual rolling elements depending on the size of the external load (force from the gearwheel) and on the size of preload between the pair of tapered roller bearings. Figure 5 shows the radial load distribution on the rolling elements at different preload values.

The analysis of the load indicates that the size of the forces acting between the rolling elements and the raceways depends on the size of preload at the same outer load, they are not constant and with increasing preload also increases the number of rolling elements that transfer this load.

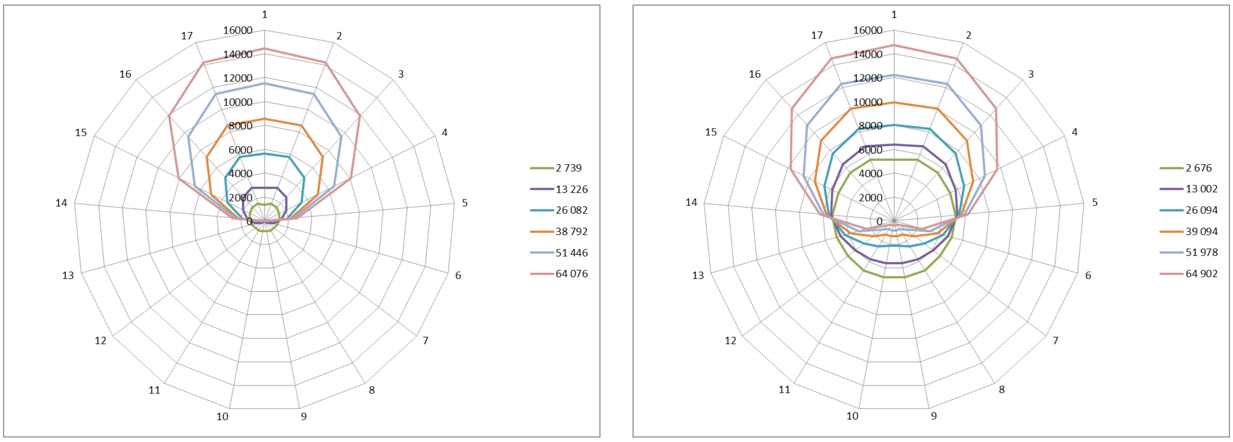


Fig. 5 Load distribution on the rolling elements depending on the size of the load, upper at 0.05 mm preload, lower at 0.25 mm preload

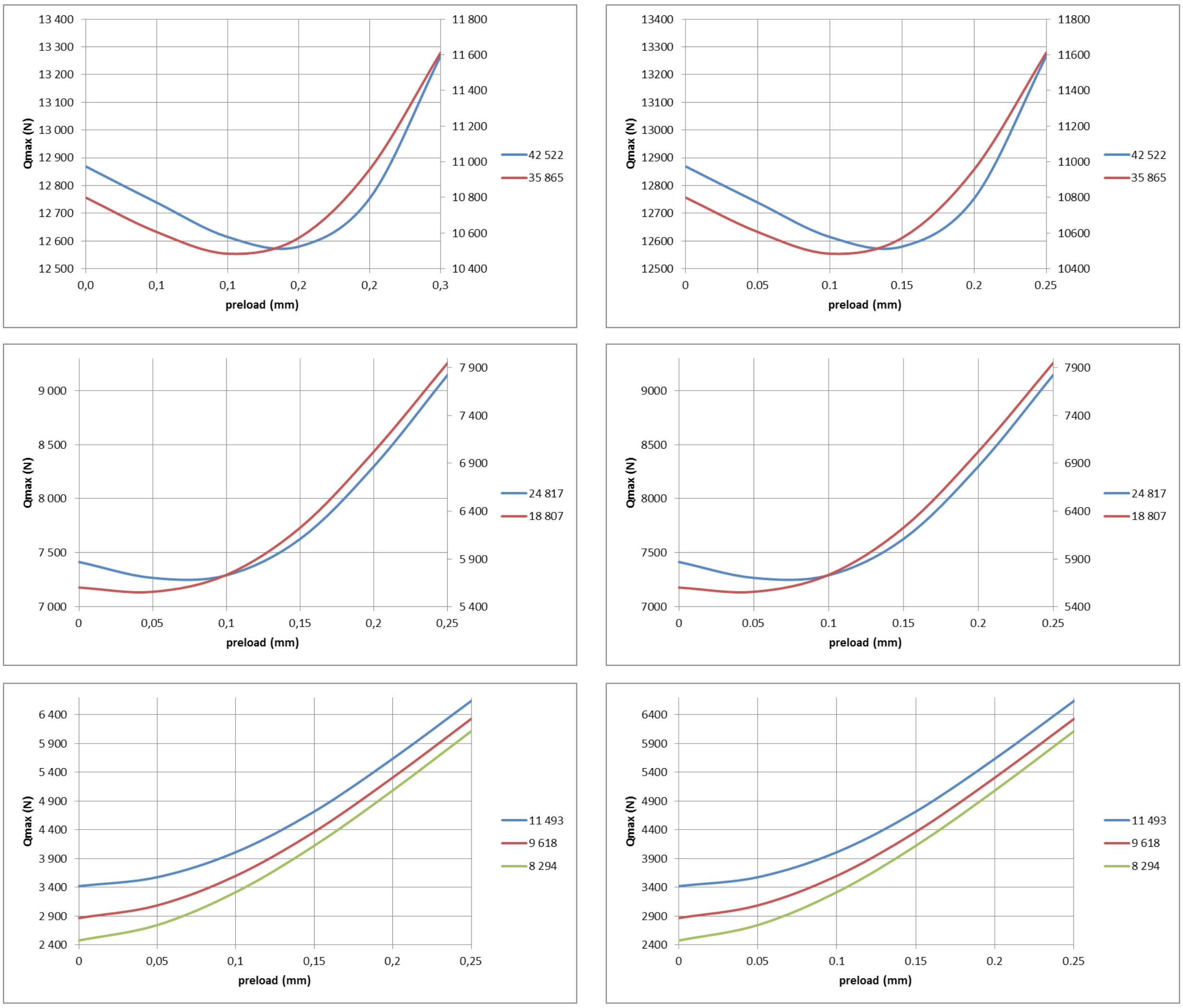


Fig. 6 Dependence of the maximum rolling element load on the size of preload at different sizes of outer load

Figure 6 shows the dependence of the maximum rolling element load on the size of preload at different sizes of outer load. The size of preload plays an important role in particular at large outer radial load.

Figure 7 shows the dependence of the maximum rolling element load on the size of outer load at various preload values. The courses show that it is not a linear dependence between the maximum rolling element load and outer load and that the course depends on the preload size between the tapered roller bearings that make up the fit on the gearbox output shaft.

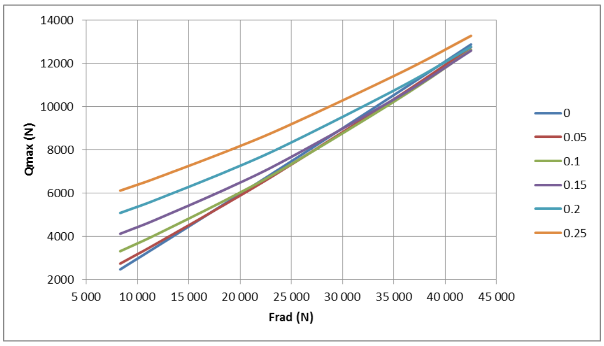


Fig. 7 Dependence of the maximum rolling element load on the size of outer load at various preload values

From the finite element analysis performed on the model of the output shaft fit, we analyzed the maximum orthogonal shear stress τ_{yz}, the depth z_0 at which this stress acts and at what length l_{ef} depending on the size of force of the maximum rolling element load Q_{max}. Calculation of service life of the inner ring of a tapered roller bearing 32008 according to the Lundberg-Palmgren theory [8 and 9], based on the calculation using the FEM is shown in Table 3.

Calculation of the life of the inner ring according to the L-P theory

Table 3

Qmax	2b	lef	□b	zo	V	N	N	N
N	mm	mm	MPa	mm	mm3	cycles	mil. rev.	hours
356	0.158746	8.47	74	0.039686	51.81803	2.70E+10	2.74E+09	9.12E+09
388	0.16325	8.75	79	0.040813	55.03566	1.64E+10	1.66E+09	5.53E+09
690	0.196973	10.14	114	0.049243	76.94489	6.08E+08	6.15E+07	2.05E+08
2 759	0.309751	13.15	277	0.077438	156.9132	2.35E+05	2.38E+04	7.92E+04
5 721	0.393055	14.07	442	0.098264	213.1354	3.99E+03	4.04E+02	1.35E+03
8 666	0.450145	14.31	577	0.112536	248.1074	4.01E+02	4.05E+01	1.35E+02
11 607	0.49522	14.44	695	0.123805	275.6014	7.96E+01	8.05E+00	2.68E+01
14 565	0.533338	14.49	805	0.133334	297.7658	2.27E+01	2.30E+00	7.67E+00

Figure 8 shows the dependence of the service life of the inner ring of a taper roller bearing 32008 on the size of force of the maximum rolling element load calculated using the L-P theory. This dependence is used to calculate the life of a tapered roller bearing according to the maximum rolling element load, as it is the impact of preload that affects the size and distribution on the rolling elements according to the applied preload [10].

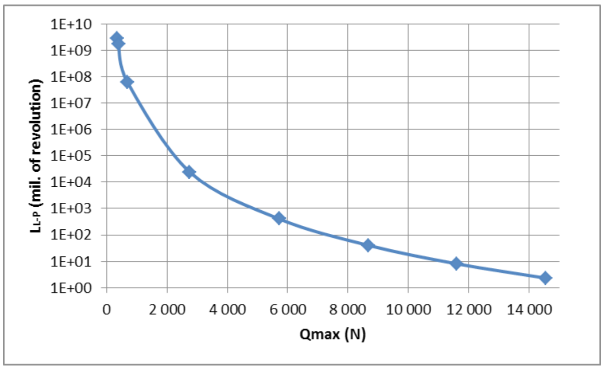


Fig. 8 Dependence of tapered roller bearing life on the maximum rolling element load

The analyzed values of the maximum rolling element load were used to calculate the values of the tapered roller bearing life for each working cycle at various preload values. Subsequently, using the ratio of required life (from the working cycle) and calculated life according to the L-P theory we determined the resulting life in hours (see Table 4 below).

Determination of optimal preload according to the L-P theory

Table 4

	preload (mm)					
	0	0.05	0.1	0.15	0.2	0.25
RArad (N)	Qmax (N)					
56 565.7	12 868.7	12 738.4	12 614.3	12 579.8	12 753.5	13 266.9
47 711.0	10 798.8	10 605.8	10 484.3	10 572.8	10 957.1	11 612.2
33 013.0	7 414.5	7 267.0	7 291.2	7 627.3	8 299.9	9 146.4
25 018.7	5 600.7	5 556.3	5 733.9	6 227.5	7 024.8	7 952.2
15 289.0	3 418.9	3 574.5	4 009.2	4 715.9	5 634.7	6 638.7
12 794.9	2 864.2	3 084.1	3 597.3	4 362.5	5 306.9	6 326.7
11 033.8	2 473.6	2 742.3	3 313.8	4 121.2	5 082.5	6 112.5

	preload (mm)					
	0	0.05	0.1	0.15	0.2	0.25
RArad (N)	LL-P (milion of revolution)					
56 565.7	4.377	4.636	4.899	4.975	4.605	3.687
47 711.0	11.760	13.017	13.891	13.248	10.834	7.810
33 013.0	97.853	109.582	107.549	83.425	51.819	29.980
25 018.7	475.476	497.282	416.494	261.534	132.644	65.950
15 289.0	7 673.72	5 972.43	3 127.771	1 252.914	459.549	182.406
12 794.9	20 809.18	13 715.84	5 761.758	1 943.529	644.148	239.243
11 033.8	47 537.83	26 591.70	9 149.841	2 678.003	821.707	290.497

durability	347.5	381.4	393.7	348.4	250.8	156.5

Figure 9 shows the dependence of the tapered roller bearing life on preload. Optimal preload was determined by calculation using finite-element analysis at 0.09 mm.

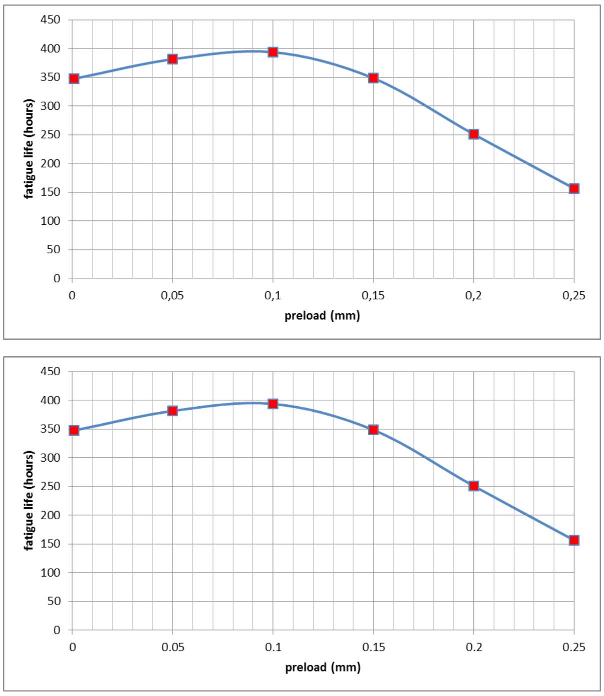


Fig. 9 Dependence of the bearing A life on preload

4. Conclusion

This paper shows the procedure that can be used to identify in an effective way the optimal preload with respect to the life of rolling bearings for a specific application defined by working cycle. The results of finite-element analysis showed its usefulness, and that analysis of contact and subsurface state of stress state and application of the Lundberg–Palmgren theory enable us to obtain a much more accurate estimate of the life of rolling bearings, compared to using the standard calculation method prescribed by ISO 281.

Figure 10 shows the dependence of the calculated bearing A life on the size of preload in the fit of the output shaft by all the three procedures.

As can be seen, calculation of equivalent load using the method according to ISO 281 does not allow determination of optimal preload with respect to the life of bearings used in the shaft fit. A more accurate method of determining the optimal preload is analytical, by calculating the maximum rolling element load from radial load distribution using iteration procedure by Stribeck. However, this procedure enables to determine with some precision the maximum rolling element load only with purely radial load. Once the bearing load is contributed to also by outer axial component, the calculation of load distribution is difficult and may not give a reliable result. The most equitable result of the calculated rolling bearing life and determination of optimal preload can be achieved using the finite element method. Finite-element analysis enables to apply to the model of fit any outer load, take into account all design and technological impacts, analysis of contact and subsurface state of stress leads to obtaining the most accurate calculation of the life applying the Lundberg-Palmgren theory. With the today's powerful hardware and sophisticated software tools we can perform these analyses in an efficient manner that does not need to be too time-consuming.

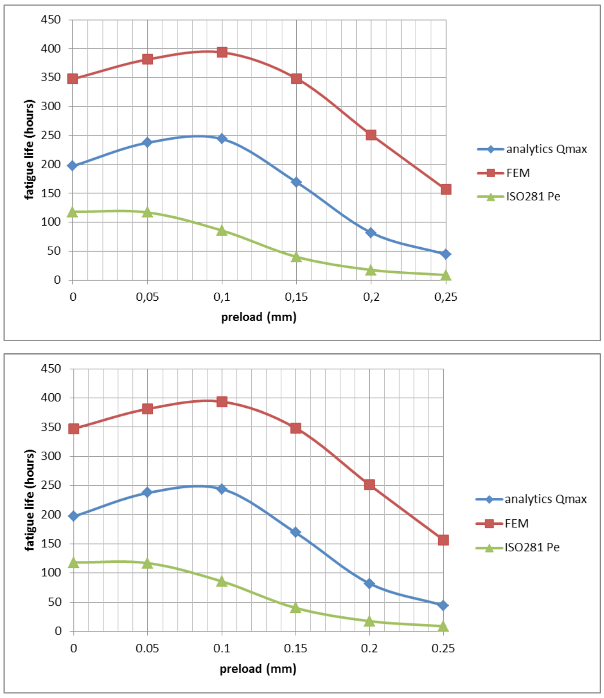


Fig. 10 Dependence of the bearing A life on the size of preload calculated by various procedures

Acknowledgements

The research is supported by the following institutions with project Slovak Research and Development Agency under the contract No. APVV-0419-11 - Adaptation of modern computer-simulation methods to the development of rolling bearings and their verification in real conditions.

The research is supported by the following institutions with project Slovak Research and Development Agency under the contract No. APVV-14-0508 - Development of new methods for the design of special large-size slewing rings.

The research is supported by the following institutions with project Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the contract No. V-1/0396/14 - Research of the influence of construction and technological parameters of rolling bearings on the service life thereof.0,

References

[1] HARRIS, T. A., KOTZALAS, M. N.: *Rolling Bearing Analysis, 5th ed.: Essential Concepts of Bearing Technology,* Taylor & Francis, 2007, ISBN 0-8493-7183-X.

[2] HARRIS, T. A., KOTZALAS, M. N.: *Rolling Bearing Analysis, 5th ed.: Advanced Concepts of Bearing Technology,* Taylor & Francis, 2007, ISBN 0-8493-7182-1.

[3] BRANDLEIN, J., ESCHMANN, P., HASBARGER, L., WEIGARD, K.: *Ball and Roller Bearings,* Theory, Design and Application, 3rd ed., John Wiley & Sons, 1999. ISBN: 0 471 98452 6.

[4] SKF: Main Catalogue, SKF Group, 2007.

[5] Timken Products Catalogue, 2015.

[6] LEMAITRE, J., DESMORAT, R.: *Engineering Damage Mechanics: Ductile, Creep, Fatigue and Brittle Failures,* Springer, 2005, ISBN 3-540-21503-4

[7] ISO 281:2007 *Rolling Bearings - Dynamic Load Ratings and Rating Life*, Intern. Standards.

[8] OSWALD, F. B., ZARETSKY, E. V., POPLAWSKI, J. V.: *Effect of Internal Clearance on Load Distribution and Life of Radially Loaded Ball and Roller Bearings,* NASA/TM-2012-217115.

[9] SADEGHI, F., JALALAHMADI, B., SLACK, T. S., RAJE, N.: A Review of Rolling Contact Fatigue, *J. of Tribology*, 2009, vol. 131/041403-1.

[10] DAMBAUGH, G. F.: *Fatigue Considerations of High Strength Rolling Bearing Steels,* 2006.

Milan Malcho - Jozef Jandacka - Richard Lenhard - Stanislav Gavlas - Tadeas Ochodek *

NUMERICAL SIMULATION OF HEAT TRANSPORT FROM THE ELECTRIC ARC FURNACE TO THE HEAT RECOVERY SYSTEM

The paper presents a numerical model of heat transport from melting charge of the electric arc furnace designed for ferroalloy production, into a cooled roof of the electric arc furnace from which waste heat is led to the heating system of a company. The model considers the influence of an asymmetric flue gas removal and geometry of electrode arrangement. The model was developed implementing CFD - the finite volume method in ANSYS - Fluent program code. Outputs of the simulation model served as a basis for dimensioning heat exchanger circuits designed for waste heat recovery to provide a company heating.

Keywords: *Numerical model of heat transport, electric arc furnace, waste heat recovery, ferroalloys.*

1. Introduction

Energy supply, energy price development together with legislative measures concerning ecology are factors which expressly encourage to pursue such technical solutions which lead to savings of waste low- and medium potential heat. This source of heat energy can advantageously be recovered in suitable facilities and used as secondary energy sources particularly in those technologies which produce it.

Secondary energy sources generally occur as a consequence of fuel and energy consumption in technological processes, in which they are not used anymore because their parameters do not fit the original technology. However, they can be an energy source for other equipment in which they can completely or partly substitute either fuel or energy.

One way of reducing costs of heat energy is its recovery from waste heat through heat exchangers of various constructions which depend on a given application.

2. Determining the parameters of working medium in electric arc furnaces (EAF)

The processes of ferroalloy production are reduction processes during which the main element is reduced by a reductant from its oxide to the metallic state. Together with the present iron it then forms alloy – ferroalloy [1]. Reduction processes take place at high temperatures and the required heat is supplied in the form of electric energy in the electric arc furnace (Figs. 1 and 2).

As can be seen from Sankey diagram (Fig. 3), technological processes of ferroalloy production in electric arc furnaces (EAF) show a massive flow of waste heat [2] and are potentially usable as secondary energy sources.

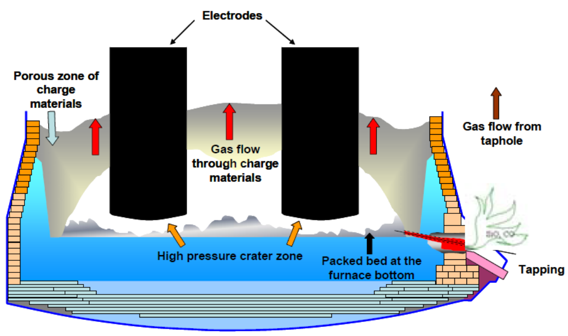


Fig. 1 Schematic of submerged arc furnace used for ferrosilicon production [1]

* [1]Milan Malcho, [1]Jozef Jandacka, [1]Richard Lenhard, [1]Stanislav Gavlas, [2]Tadeas Ochodek
[1]Department of Power Engineering, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]Energy Research Center, VSB – Technical University of Ostrava, Ostrava - Poruba, Czech Republic
E-mail: milan.malcho@fstroj.uniza.sk

Fig. 2 ARC furnace used for ferroalloy production in OFZ, a.s. Istebne, plant Siroka

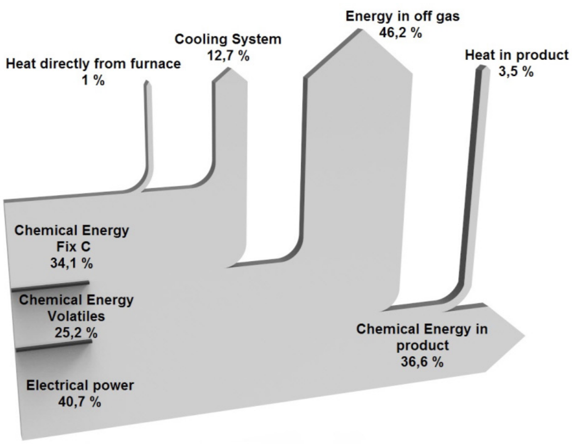


Fig. 3 A Sankey diagram of an electric arc furnace [2]

Ferroalloy furnaces work at relatively high temperatures. In ferrosilicon production the temperature of the furnace jacket is about 400**°C** and in the case of technological failures it achieves as much as 900 - 1000 °C. When ferrosilicon is produced in electric arc furnaces (EAF) there are silicon ashes which are formed by condensation of silicon dioxide vapour radiating from the EAF. Ashes are very fine particles of regular spherical shape and smooth surface. Silicon ash is adhesive, features excellent heat insulating properties and increases heat resistance on the inner wall of the furnace roof.

The main task of the cooled roof is to carry away part of radiation and convection heat from the process of ferroalloy melting (Fig. 3) and recover it through the recovery equipment to the energy system of a company. Subsidiary effect of the equipment is to reduce the temperature of flue gases and to improve the workplace (reduction of average radiation temperature) in the vicinity of the furnace.

The roof is an auxiliary technological device suspended above the furnace on the manufacturing hall construction. It is of rotary shape with a perpendicular axis. There are three perpendicular electrodes installed in an axis of the furnace and roof. Current for the electrodes is supplied from three sides through holes in the roof.

3. Numerical model of heat transport in the electric arc furnace

To determine heat flows to individual parts of the EAF roof and, consequently, dimension the primary circuits of heat transfer medium, it was necessary to develop a numerical model of heat transport from melting charge and from flue gases to the heat exchanger walls positioned in the furnace roof. Numerical simulation of transfers in EAF in a stationary regime was performed in ANSYS Fluent 16.2 code taking into consideration a radiation element of heat transport.

3D geometry of an EAF simplified model was created in the program DesignModeler R16.2. The modelled system consists of four types of flue gas/water heat exchangers placed one above the other in four zones, of three electrodes and a part of flue gas removal without heat transfer surfaces. The total volume of an EAF simplified model was considered as fluid.

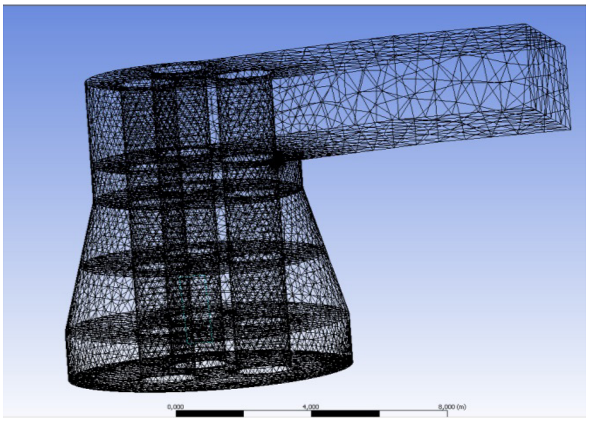


Fig. 4 Mesh generated in program Meshing R16.2

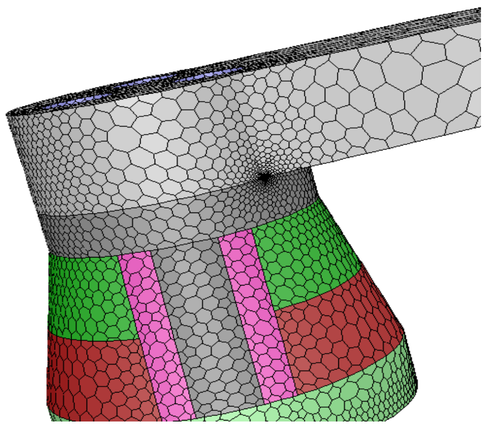


Fig. 5 Converted to polyhedral mesh

Computer mesh for 3D simplified model of EAF was developed in the program Meshing R16.2 (Fig. 4). The mesh itself contains 332,564 cells. The mesh was built with the patch conforming method by means of which the cells of tetrahedron type were developed (Fig. 5) using the Patch Conforming algorithm. Advanced meshing functions Curvature and Proximity were used for meshing.

The achieved quality of mesh – skewness was 0.84; from the point of view of mesh assessment, this mesh is in the range of acceptability. To improve it and to speed up the numerical simulation, the mesh was further arranged through conversion to polyhedral mesh. This polyhedral mesh achieved the mesh quality – skewness of 0.75 and is assessed as good. This arrangement contributed to a faster conversion of the task.

In the numerical simulation, we chose the turbulence model k-ε with setting "realisable" that shows very good results in solving similar tasks [3, 4 and 5]. To solve the radiation we used a radiation model Ordinates (DO) with resolution 5x5. In the equation for radiation transfer (1) [6], this model solves radiation intensity which depends on location and direction. Radiation heat transfer is counted in every quadrant or octant. It solves transport equations in a similar manner as when solving convection and energy balance according to:

$$\nabla \cdot [I(\mathbf{r},\mathbf{s}) \cdot \mathbf{s}] + (a + \sigma_s) I(\mathbf{r},\mathbf{s}) = an^2 \frac{\sigma T^4}{\pi} + \frac{\sigma_s}{4\pi} \int_0^{4\pi} I(\mathbf{r},\mathbf{s}') \Phi(\mathbf{s} \cdot \mathbf{s}') d\Omega' , \quad (1)$$

where
r – position vector
s – directional vector
s´ - directional dispersion vector
s - distance
a – absorption coefficient
n – refractive index
σ_s –coefficient of dispersion
σ - Stefan-Boltzmann constant (5.669 x 10^{-8} $Wm^{-2}K^{-4}$)
I – intensity of radiation which depends on position and direction
T – local temperature
Φ – phase function
Ω´ - angle.

The simulation itself was performed in the program ANSYS Fluent 16.2. in which the boundary conditions were set according to Table 1. Flue gases with a defined composition were taken as a working medium. Walls (wall - Table 1) of the model were developed as simplified and in simulation the number of layers, their thickness, temperature and emissivity were given (Fig. 6).

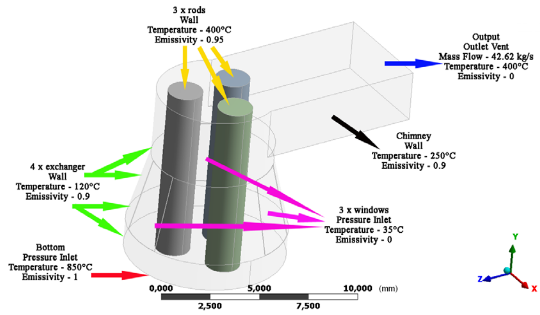


Fig. 6 Input parameters for a simulation model of EAF

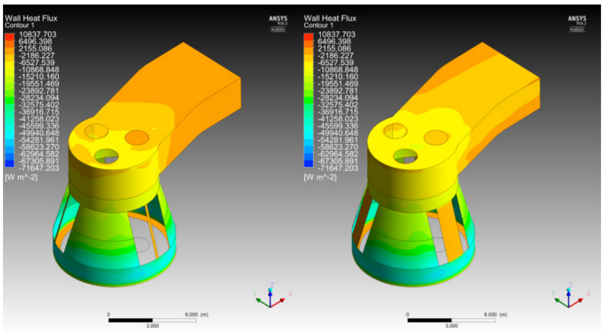


Fig. 7 Comparison of influence of the size of short mesh hole on the air intake and jacket temperature

The numerical simulation was performed in standard operational conditions with a stationary heat transport by radiation from the melting layer under chosen input parameters presented in Table 1. The change of input cross-section to the EAF roof through a hole for a short mesh was determined through the reduction of its cross-section according to the number and cross-sections of cables and cooling pipes (Fig. 7). For the modelling of the hole for a short mesh we restricted the diameter of the hole

Boundary conditions

Table 1

Name of boundary	Input parameters	Mass flow	Temperature	Emissivity
	Type of boundary	kg/s	°C	
bottom	pressure-inlet	-	850	1
3 x windows	pressure-inlet	-	35	0
output	outlet-vent	42.62	400	0
4 x exchangers	wall	-	120	0.9
chimney	wall	-	250	0.9
3 x rods	wall	-	400	0.95

with a continuous band covering 50% of the input cross section among roof segments. The influence of the hole size on the volume flow of intake air and, thus also on the temperature of the jacket of flue gas removal unit can be seen in Fig. 7.

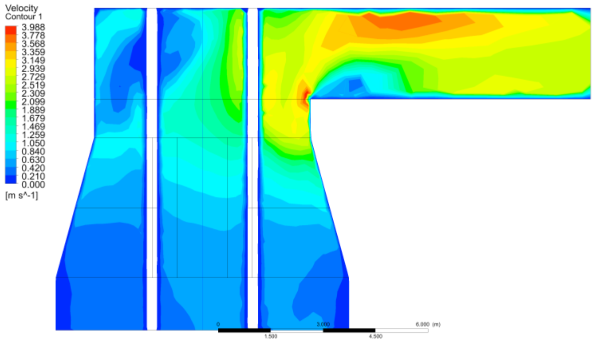


Fig. 8 Velocity field in a longitudinal section of flue gas removal unit

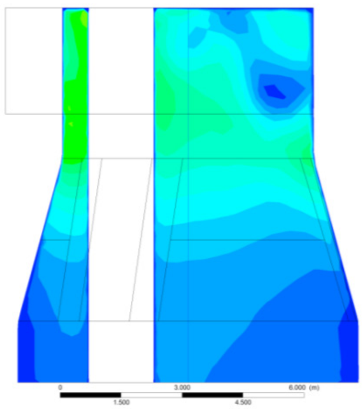


Fig. 9 Velocity field in vertical section to flue gas removal unit

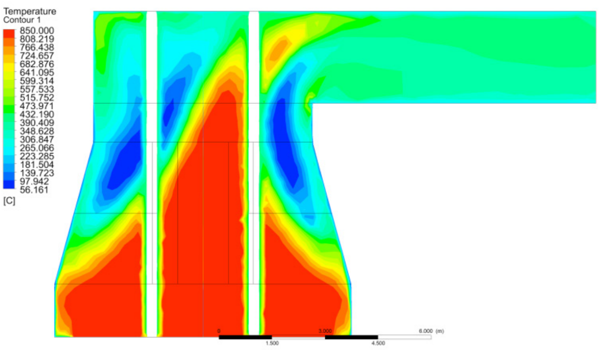


Fig. 10 Thermal field in a longitudinal section against flue gas removal unit

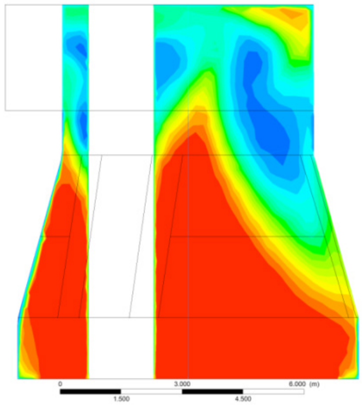


Fig. 11 Thermal field in vertical section to the axis of flue gas removal unit

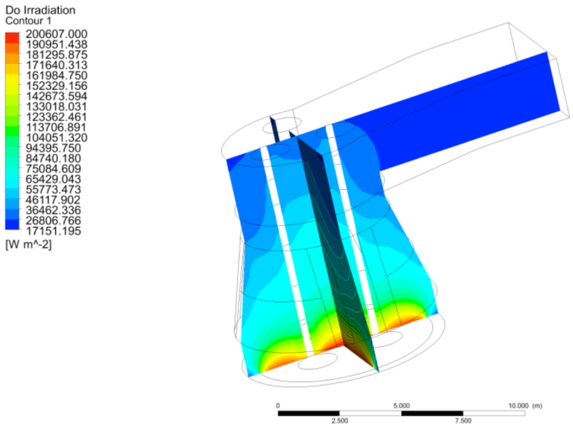


Fig. 12 Radiation density shown in two mutually perpendicular planes

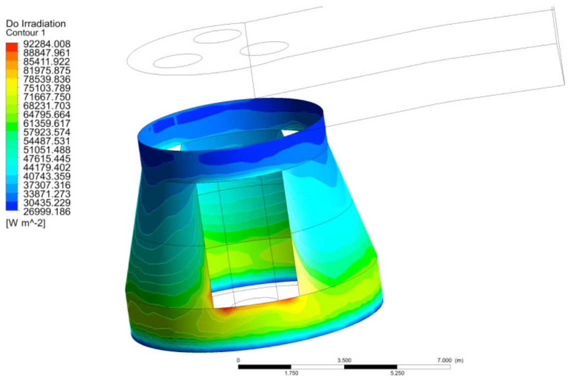


Fig. 13 Radiation flow to the roof wall

The result of simulation (Figs. 8, 9, 10, 11, 12 and 13) is the computation of radiation and convection element of heat transport to the four zones of the EAF roof (Fig. 12, 13 and 14).

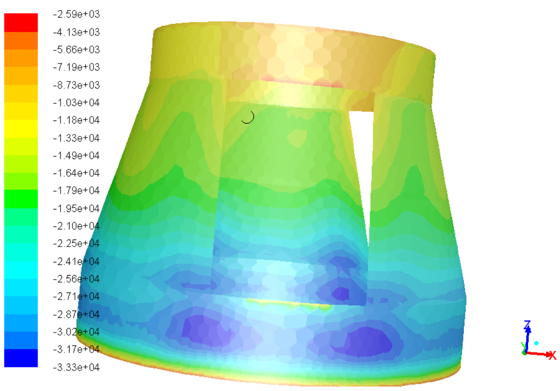


Fig. 14 The overall heat flow to the furnace roof through radiation and convection

The first zone comprises a cylindrical surface on which heat exchangers are placed marked as lifting screens and heat exchangers below the short mesh;

The second zone consists of heat exchangers in the bottom tapered part of the roof;

The third zone consists of heat exchangers in the upper tapered part of the roof;

The fourth zone consists of heat exchangers on the cylinder ring in the upper part of the EAF roof.

The computed mean weighted overall density of heat flow to individual zones which are free from flue ash is then as presented in Table 2.

Heat flow densities and thermal output of individual zones of EAF roof

Table 2

Part of EAF roof	Average density of heat flow in a zone (kW/m^2)	Calculated overall heat flow to the zones (kW)
1st zone	**26.684**	**1 055.706**
2nd zone	**22.873**	**752.109**
3rd zone	**16.119**	**412.052**
4th zone	**11.557**	**247.383**

The simulations of heat transport in the EAF via a simplified geometry of an unobstructed furnace jacket showed that it is possible, after having implemented the proposed arrangements, to carry off the total thermal output of about 2.467 MW from the EAF furnace roof to the heat transfer medium (warm water). Heat exchangers under a short mesh and those on the lifting screens seem to be the most powerful.

4. Constructional solution of the heat recovery system

When designing heat exchangers, attention is paid to the following facts so that they would:

- absorb mostly the radiation component of heat which prevails in the roof area,
- enable creating the largest possible area exposed to radiation and capturing the most possible amount of radiated heat [7],
- avoid reduction of the roof space and enable free flow of flue gases via the roof to the removal unit,
- have simple shapes without edges and corners, if possible, to avoid dust capturing,
- provide continuous functioning and technological process in the furnace,
- be resistant against failures, be controllable and repairable,
- be simple from production and material points of view.

The ring is the vital part of the cooled furnace roof. It is a welded construction from steel sheets and rolled profiles. The cooler segments are screwed to the lower flange of the ring and the flue removal unit is welded to the upper edge. On the ring periphery there are plate heat exchangers.

In order to achieve the required thermal output of about 2.5 MW it was necessary to add heat transfer surfaces to all usable surface of the furnace roof.

On the basis of calculated heat flows, we dimensioned pipes [8], heat exchangers and tube-type construction and developed a 3D model of the overall assembly of the EAF cooled roof.

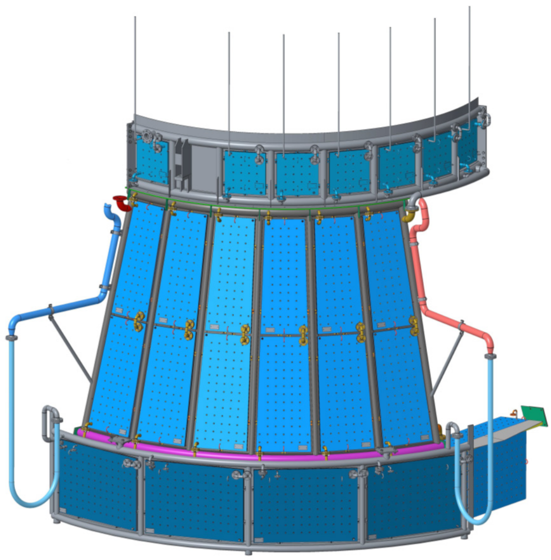


Fig. 15 A segment of pressure vessel of the EAF cooled roof

The furnace roof is designed as a pressure vessel divided into three independent segments (Fig. 15, 16 and 17). The segments consist of a set of plate heat exchangers and tube-type construction. Each segment is an independent unit of pressure vessels with a maximum pressure of 0.2 MPa and maximum temperature of 110 °C.

Useful thermal output demand at the furnace performance of 18 MW_e is 2 -2.5 MW_t. This heat is recovered from waste heat originating during ferrosilicon production and is led to a heat exchanger station and, consequently used to heat water in a company. In the heat exchange station water in the heating circuit of a plant is heated to the temperature of max. 90°C. When the heat from EAF roof is not usable for heating system, its degradation is provided in the cooling circuit of the company by means of individual heat exchangers placed in the heat exchanger station. The complete system of measurement, control and management is provided by the computer.

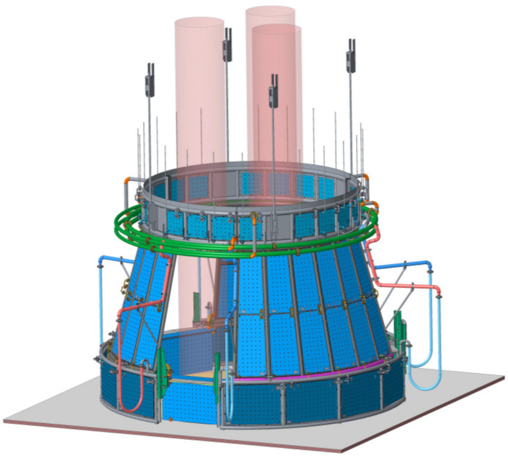


Fig. 16 The overall 3D model

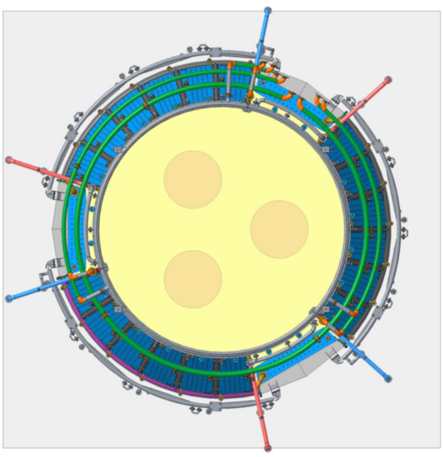


Fig. 17 Overall 3D model - plan view

5. Conclusion

The analysis of recovery of heat from electric arc furnaces showed that it is not a trivial problem. On the contrary, it requires a certain degree of experience from the field of recovery of heat from technological processes. Its solution should have a complex access not only to ways and forms of heat acquisition but also to possibilities of its utilisation. Highly efficient tools seem to be extensive program packages (code Ansys Fluent, ProEngineering), which enable - with the help of powerful computers - to analyse states also in such complicated thermal systems as electric arc furnaces are. Present experience also showed that it is necessary to design such technical solutions which accept the most relevant problem of heat recovery from EAF - clogging of heat transfer surfaces of recuperation exchangers by flue ash featuring very specific characteristic (perfect adhesion to uneven surfaces). The mentioned characteristic significantly impairs the use of waste heat from EAF for production of electric energy either by means of water vapour or ORC cycle. Current equipment requires a stable stationary relatively clean source of heat of high potential.

Great attention should be also paid to the high quality of control of the equipment designed for heat recovery which has to flexibly react to a non-stationary character of thermal loading of heat exchangers on the furnace roof resulting from a manner of clogging or from melting the surface of charge in the furnace.

Last but not least, there is a question of using considerable thermal benefits from hot water. This type of heat energy should, first of all, cover the requirements for heating (both of spaces and water). It is necessary to find such an auxiliary technology of waste heat utilisation which will provide optimal effect from invested means.

Acknowledgements

This paper has been elaborated in the framework of the project APVV-0577-10 Cooling of power electronic systems by cooling cycles without mechanical drive (COPES).

This contribution is the result of the project implementation: Slovak Infrastructure for High Performance Computing, ITMS code 26230120002, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] KADKHODABEIGI, M., TVEIT, H., JOHANSEN, S. T.: *Modeling the Tapping of Silicon Melt from the Submerged Arc Furnaces,* 9th Intern. Conference on CFD in the Minerals and Process Industries, CSIRO, Melbourne, 2009.

[2] HJARTARSON, H., PALSSON, H., SAEVARSDOTTIR, G.: *Waste Heat Utilization from a Submerged Arc Furnace Producing Ferrosilicon.* 12th Intern. Ferroalloys Congress Sustainable Future, Helsinki, Finland 2010.
[3] LAZIC, L., VARGA, A., KIZEK, J.: Analysis of Combustion Characteristics in a Aluminium Melting Furnace. *Metalurgija*, vol. 44, pp. 195-199, 2005.
[4] MALCHO, M., JANDACKA, J. OCHODEK, T., KOLONICNY, J.: Correlations for Heat Transport by Natural Convection of Cylindrical Surfaces Situated Above the Other. *Communication - Scientific Letters of the University of Zilina,* vol. 17, No. 3, pp. 4-11, 2015.
[5] KAPJOR, A. HUZVAR, J., FTOREK, B., SMATANOVA, H.: *Analysis of the Heat Transfer from Horizontal Pipes at Natural Convection,* proc. of the intern conference XIX Application of experimental and numerical methods in fluid mechanics and energetics 2014, Melville : AIP Publishing, pp. 58-622014. ISBN 978-0-7354-1244-6.
[6] INCROPERA, F. P., DEWITT, D. P, BERGMAN, T. L, LAVINE, A. S.: *Fundamentals of Heat and Mass Transfer,* 6st ed. : John Wiley : New York, 2007.
[7] BRESTOVIC, T., CARNOGURSKA, M. PYSZKO, R., KUBIK, M. : *Effect of Radiation on Heat Exchange in Finned Heat Transfer Surfaces,* XVIII Intern. scientific conference The application of experimental and numerical methods in fluid mechanics and energy. EDIS: University of Zilina, 2012, pp. 19-26. ISBN 978-80-554-0516-2.
[8] SKOCILASOVA, B., SKOCILAS, J.: *Determination of Pressure Drop Coefficient by CFD Simulation.* Proc. of AIP Conference, vol. 1608, pp. 206-210, 2014.

Uncited references or further readings

[9] KAPJOR, A., GRESSAK, T., HUZVAR, J., VANTUCH, M.: *Heat Transfer by Thermovision and CFD Methods.* Proc. of AIP Conference, vol. 1608, pp. 103-107, 2014.
[10] KNOUREK, J., MATAS, R., PROKES, O., TENKRAT, D.: *Numerical Simulations of Natural Gas Flow in Pipe System with Flowmeters*. EPJ Web of Conferences, vol. 67, 2013.

COMMUNICATIONS – Scientific Letters of the University of Zilina Writer's Guidelines

1. Submitted papers must be unpublished and must not be currently under review for any other publication.
2. Submitted manuscripts should not exceed 8 pages including figures and graphs (in Microsoft WORD – format A4, Times Roman size 12, page margins 2.5 cm).
3. Manuscripts written in good English must include abstract and keywords also written in English. The abstract should not exceed 10 lines.
4. Submission should be sent by e-mail – as an attachment – to one of the following addresses: komunikacie@uniza.sk or holesa@uniza.sk (or on CD to the following address: Zilinska univerzita, OVaV – Komunikacie, Univerzitna 1, SK – 010 26 Zilina, Slovakia).
5. Uncommon abbreviations must be defined the first time they are used in the text.
6. Figures, graphs and diagrams, if not processed in Microsoft WORD, must be sent in electronic form (as JPG, GIF, TIF, TTF or BMP files) or drawn in high contrast on white paper. Photographs for publication must be either contrastive or on a slide.
7. The numbered reference citation within text should be enclosed in square brackets. The reference list should appear at the end of the article (in compliance with ISO 690).
8. The numbered references (in square brackets), figures, tables and graphs must be also included in text – in numerical order.
9. The author's exact mailing address, full names, E-mail address, telephone or fax number, the name and address of the organization and workplace (also written in English) must be enclosed.
10. The editorial board will assess the submitted paper in its following session. If the manuscript is accepted for publication, it will be sent to peer review and language correction. After reviewing and incorporating the editor's comments, the final draft (before printing) will be sent to authors for final review and minor adjustments
11. Submission deadlines are: September 30, December 31, March 31 and June 30.

COMMUNICATIONS

SCIENTIFIC LETTERS OF THE UNIVERSITY OF ZILINA
VOLUME 18

Address of the editorial office:
Zilinská univerzita
Office for Science and Research
(OVaV)
Univerzitna 1
SK 010 26 Zilina
Slovakia

E-mail: komunikacie@uniza.sk

Each paper was reviewed by two reviewers.

Journal is excerpted in **COMPENDEX** and **SCOPUS**.

It is published by the University of Zilina in
EDIS – Publishing Institution of Zilina University
Registered No: EV 3672/09
ISSN 1335-4205

Published quarterly

Single issues of the journal can be found on:
http://www.uniza.sk/komunikacie

ICO 00397 563
March 2016