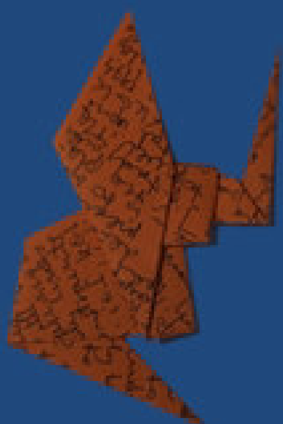
1st International Conference
Decision Making in Manufacturing and Services
&
XX Jubilee Symposium
Applications of Systems Theory



DMMS 2017 & ZTS XX

Zakopane, September 26-30, 2017

1st International Conference
Decision Making in Manufacturing and Services
&
XX Jubilee Symposium
Applications of Systems Theory

DMMS 2017 & ZTS XX
Conference Proceedings

AGH

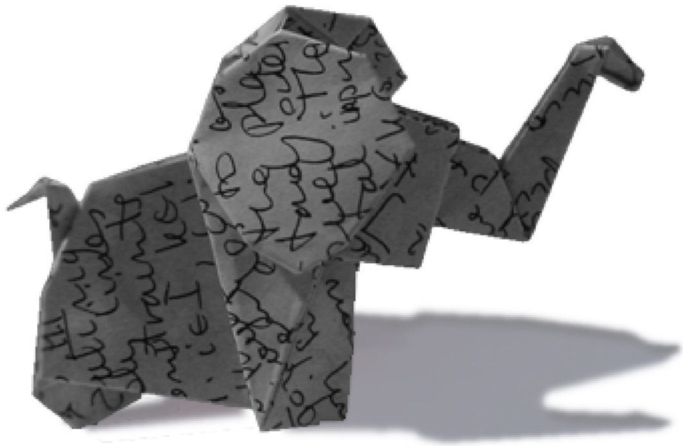
DMMS ZAKOPANE 2017

MAŁOPOLSKI URZĄD WOJEWÓDZKI

MAŁOPOLSKA

Ministry of Science and Higher Education
Republic of Poland

MINISTRY
OF INFRASTRUCTURE
AND CONSTRUCTION



DMMS 2017 & ZTS XX
Zakopane, September 26-30, 2017

Editor:

Tadeusz Sawik

Publisher:

Agencja Reklamowo-Wydawnicza „Ostoja", Cianowice, ul. Niebyła 17, 32-043 Skała

Organizer:

AGH University of Science and Technology, Poland

Cover design:

Joanna R. Marszewska

Technical editors:

Katarzyna Gdowska, Roger Książek

Each paper has been reviewed by at least two reviewers. The accepted versions of all papers are presented in the form which is delivered by authors. The Organizer isnotresponsible for statements advanced in papers or spelling and grammar irregularities.

Printed in: 80 copies

ISBN 978-83-62218-57-8

Scientific/Reviewing Committees

Tadeusz SAWIK

T.C. Edwin CHENG

Alexandre DOLGUI

Moshe DROR

Jan Tadeusz DUDA

Stanley B. GERSHWIN

Henryk GÓRECKI

Alain B. HAURIE

Marek HRYNIEWICZ

Dmitry IVANOV

Marianna JACYNA

Waldemar KACZMARCZYK

Rainer KOLISCH

Janusz KOWAL

Jacek B. KRAWCZYK

Marek KUBALE

Wiesław KUBIAK

Mary E. KURZ

Piotr ŁEBKOWSKI

Edward MICHLOWICZ

Jerzy MIKULIK

Wojciech MITKOWSKI

Seichi NAKAMORI

Peter NELSON

Chrissoleon PAPADOPOULOS

Joao Pedro PEDROSO

Erwin PESCH

Czesław SMUTNICKI

Moshe SNIEDOVICH

Kathryn E. STECKE

Janusz SZPYTKO

Ryszard TADEUSIEWICZ

Thomas M. TIRPAK

Manoj K. TIWARI

Eugeniusz TOCZYŁOWSKI

Jan WĘGLARZ

Ana VIANA

Jan WĘGLARZ

Margaret M. WIECEK

Kiyoshi YONEDA

W. Henk M. ZIJM

DMMS 2017 & ZTS XX
Organizing Committee

Chairman

Tadeusz SAWIK

Members

Katarzyna Gdowska

Roger Książek

Joanna Marszewska

Bartosz Sawik

Editor's foreword

As conference chair of the two conferences, this is my pleasant duty to present this joint publication, which contains the papers from the 1st International Conference on Decision Making in Manufacturing and Services – DMMS 2017 and XX Jubilee Symposium on Applications of Systems Theory – XX ZTS.

The conferences took place in Zakopane, Poland, from September 26–30, 2017. It was hosted by Department of Operations Research, Faculty of Management, AGH University of Science and Technology (Kraków, Poland).

The ZTS (Zastosowania Teorii Systemów) Symposium has been established jointly by Departments of Automatics of the two faculties of AGH University: Faculty of Mechanical Engineering and Robotics and Faculty of Electrical Engineering, Electronics and Automatics. The idea of a joint Symposium, was initiated by Edward Nawarecki and Tadeusz Sawik during the IFAC/IFORS Symposium on Comparison of Automatic Control and Operational Research Techniques Applied to Large Systems Analysis and Control, in Toulouse, 1979. ZTS Symposium was launched in 1979 and has been organized annually or biannually (first in Krościenko and then in Zakopane) with Stefan Ziemba as a honorary conference chair for many years and Ewa Kochan as a chair of the organizing committee. The symposium has become a well-known national scientific event that attracted most distinguished Polish researchers and academicians in the area of systems theory and engineering.

The launching of a new DMMS Conference in 2017 coincides with 10^{th} anniversary of the international journal ***Decision Making in Manufacturing and Services***, published by AGH University since 2007. One of the major roles of the DMMS journal is to provide a platform for bridging the gap between theory and implementation of decision-making. I hope that the future DMMS conferences will provide a recognized forum for the Decision Making community, including academic and industry researchers and decision makers working at the interface of research and implementation in manufacturing and services.

I would like to express sincere gratitude to all those without whom the handling and organizing of this conference would have been impossible. In particular, it has been a pleasure to work with Katarzyna Gdowska, Roger Książek et al. I thank especially to Joanna Marszewska for design of the DMMS Conference logo as well as for the cover design of this proceedings and the picture shown above. The conference logo is her artistic vision of “parzenica”, a traditional decorative motif embroidered on the folk costumes of Polish mountaineers.

On behalf of both scientific and organizing committees of the two conferences, I wish that all of you take the opportunity of attending DMMS 2017 & ZTS XX to renew friendships and forge new ones, and that, you are able to enjoy your stay in Zakopane, a capital of Polish Tatra Mountains.

Prof. Tadeusz Sawik, PhD, ScD
AGH University of Science and Technology
Faculty of Management
Kraków, Poland

TABLE OF CONTENTS

International Conference
on Decision Making
in Manufacturing and Services
2017

Invited Speakers

Small Graphs Can Cause Big Coloring Problems

Marek KUBALE[1]

[1] *Gdańsk University of Technology, Gdańsk, Poland*

Abstract: *On the whole, the performance of graph coloring heuristics is studied by giving asymptotic results. These are usually the performance guarantee and time complexity. Both functions tell us what one can expect at worst as the number of vertices n tends to infinity, but we really do not know what is going on at the other end of the scale, say if n<10. In 1990 Hansen and Kuplinsky introduced the concept of hard-to-color (HC) graphs. These are graphs that cannot be colored optimally by some approximation algorithms. The aim of studying such graphs is threefold. First, analyzing HC graphs makes it possible to obtain improved algorithms which avoid hard instances as far as possible. Secondly, it enables to search for small benchmarks, which is an indispensable tool in evaluating comprehensive families of graph coloring heuristics. Thirdly, it proves a more sensitive way of assessing their efficiency as compared to the performance guarantee (the larger HC graph is the better algorithm performs), since the overwhelming majority of coloring algorithms have asymptotically the same linear function of performance guarantee. In the talk we review the most popular on-line and off-line graph coloring algorithms from this point of view.*

Short bio: *Prof. Marek Kubale has worked for the ETI Faculty at Gdańsk University of Technology continuously since 1969. During almost 50 years he published over 150 papers, including 40 JCR articles. Moreover, he co-/authored 6 monographs/textbooks (e.g. _Graph Colorings_, American Mathematical Society, CM 352 (2004)). Prof. Kubale has promoted 21 doctors of philosophy (7 of them work as University Professors) and supervised more than 100 master of science projects. He conducted 22 KBN/NCN grants including prestigious project MAESTRO). He was a member of Editorial Board of two JCR journals (_Networks_ and _Disscissiones Mathematice - Graph Theory_). His Hirsch index is 12.*

Planning and Scheduling in Electronics Supply Chains Using Mixed Integer Programming

Tadeusz SAWIK[1]

[1] *AGH University of Science and Technology, Kraków, Poland*

Abstract: *This talk presents mixed integer programming (MIP) applications to production planning and scheduling in electronics supply chains. First, the short-term detailed scheduling of printed wiring boards assembly in surface mount technology lines is discussed. Then, the medium-term aggregate production planning in a production/assembly facility of consumer electronics supply chain is described, and finally coordinated aggregate planning and scheduling of manufacturing and supply of parts and production of finished products is presented. The decision variables are defined and MIP modelling frameworks provided. The two decision-making approaches are discussed and compared: integrated (simultaneous) approach, in which all required decisions are made simultaneously using a complex, large monolithic MIP model; and hierarchical (sequential) approach, in which the required decisions are made successively, using hierarchies of simpler and smaller-size MIP models. The applications of MIP approaches are illustrated with real-world examples from Motorola. Final conclusions highlight current research on stochastic MIP applications in supply chain disruption management.*

Short bio: *Tadeusz Sawik is a Professor of Industrial Engineering and Operations Research at AGH University of Science and Technology in Kraków, Poland. He has been Associate Dean for Research at School of Management, Chair of the Department of Computer Integrated Manufacturing and Chair of the Department of Operations Research and Information Technology. He received the MS degree in mechanical engineering, the PhD degree in control engineering and the ScD (habilitation) degree in operations research, all from AGH University. He has been a visiting professor in Germany, Japan, Sweden and Switzerland and he has served as a research advisor of Motorola for several years. Five times he received Scientific Excellence Award from the Minister of Science and Higher Education, one of the most prestigious award in Poland.*

He has published numerous books (including Production Planning and Scheduling in Flexible Assembly Systems, Springer, 1998, Scheduling in Supply Chains Using Mixed Integer Programming, Wiley, 2011 and Supply Chain Disruption Management Using Stochastic Mixed Integer Programming, Springer, 2018), and more than 150 individual articles in refereed journals. His current research interests are in the area of supply chain optimization, risk management, scheduling and integer programming. In the 50th volume anniversary issue of International Journal of Production Research (IJPR), a flagship journal in production research, he has been recognized as one of the top authors who have had the greatest impact on defining the knowledge represented in IJPR. He is the founding Editor-in-Chief of Decision Making in Manufacturing and Services (AGH University Press).

Newest Trends in Combinatorial Optimization

Czesław SMUTNCKI[1]

[1]*Wrocław University of Science and Technology, Wrocław, Poland*

Abstract: *Competitive markets generate lots of practical combinatorial optimization (CO) tasks with clearly defined tradeoff between computation time (cost) and profits obtainable from the implementation of the solution found. These are cases derived, among others, from: transportation/delivery systems, warehouses, workload balancing, scheduling, service, manufacturing, location, timetabling, assembly systems, supply chains, logistics, etc. Thus, this class of problems are considered as the most attractive for scientific researchers as well as for practitioners in the area of operations research. For scientist – because generates challenges for the theory and its applications; for practitioners – because they quickly change found solutions into real financial profits. Solving of these CO problems engage large computing resources in a relatively short period of time. The presentation provides the survey of newest trends in CO, especially: progress in modeling technology, development in methodology, variety of solution approaches, some algorithms, new theoretical results.*

Short bio: *Czeslaw Smutnicki received his PhD in Electronics at the Wroclaw University of Science and Technology in Poland, then DSc in Automation and Robotics at Warsaw University of Technology in Poland and Prof title by President of Poland. Currently, full Professor at the Faculty of Electronics, Wroclaw University of Science and Technology. Formerly Director of the Institute of Computer Engineering, Control and Robotics. Now the Head of Department of Computer Engineering and Dean of the Faculty of Electronics. His research interests include issues related to discrete optimization, manufacturing and transport systems, scheduling, logistics, warehousing and storage, computer algorithms. He is the author of a great deal of research studies published at national and international journals, conference proceedings, as book chapters, and a lot of books. He is also the reviewer of many scientific papers for journals from ISI master list and a member of scientific councils of many international conferences.*

International Conference on Decision Making in Manufacturing and Services 2017

Conference Proceedings

Packet Scheduling Algorithms for HSDPA

Mateusz RUSSAK[1], **Paulina RUSSAK**[2]

[1] *Wroclaw University of Science and Technology, Faculty of Electronics ul. Janiszewskiego 11/17, Wroclaw, Poland, mateusz.russak@pwr.edu.pl*

[2] *Wroclaw University of Science and Technology, Faculty of Electronics ul. Janiszewskiego 11/17, Wroclaw, Poland, paulina.russak@pwr.edu.pl*

Abstract: The paper provides certain novel scheduling algorithm for High Speed Downlink Packet Access (HSDPA) technology used in the Universal Terrestrial Radio Access Network (UTRAN). Proposed method combines two existing approaches in such a way that allows one to take into account at the same time several transmission parameters, namely Carrier to Interference (CI) ratio, actual delays and throughput. Quality of the algorithm was evaluated with the use of simulation.

Keywords: packet scheduling, HSPDA, scheduling algorithms, round robin

1. INTRODUCTION

Rapid growth of Internet usage, observed nowdays, generates increased demands addressed to the area of telecommunication networks. A result of these actions is the dynamic development of Universal Terrestrial Radio Access Network (UTRAN) technology. UTRAN in High Speed Downlink Packet Access (HSDPA) service provides communication between User Equipment (UE) device and a node of Base Station (nodeB) (see Wang, Huang, and Le (2009) for details). Thus, there are a few fundamental components of UTRAN, namely: 1. UE *(User Equipment)* station, 2. nodeB *(node of Base)* station, 3. RNC *(Radio Network Controller*). The signal is transmitted from User Equipment (UE) to Base Station (nodeB) and then to Radio Network Controller (RNC). From RNC signal goes to core network, (Korhonen (2003)). Obtaining correct signal from UE is a crucial element of the Access Network. The Access Network is responsible for receiving signal from all users which are in the cell scope. Failures in collecting packages cause signal losses and imply retransmissions. We expect from the Access Network to provide high data range while supporting the greatest number of users in cell scope in Transmission Time Interval (TTI).

Another very important aspect of UTRAN is the ability to adapt to changeable system conditions, perceived as the possibility of setting up new services requiring reconfiguration and network improvement. All these elements should guarantee stable Quality of Service (QoS). For example, in *Rel5,* the Radio Resource Management (RRM) unit was moved from RNC to nodeB enhancing needs for telecommunication network. This modification entailed reduction of TTI frame length from 10 ms to 2ms (see Wannstrom (2009), Rutkowski and Sobczak (2002)). Actually, RRM contains a lot of key methods responsible for receiving the signal. *Packet Scheduling* (PS) is a general term refering to one of these methods, (see Kolakowski and Cichocki (2007)), although no specification has been given yet.

The main role of packet scheduling algorithm is to select the user whose packets should be handled by base station in available time frame. The most important aspect of PS is increasing network throughput while maintaining good Quality of Service for each user at the same time. Up to now, PS algorithms proposed by 3rd Generation Partnership Project (3GPP) are commonly used: 1. Round Robin, 2. MAX C/I, 3. Proportional Fair. The main advantages of algorithms mentioned above are low complexity and high tolerance for mistakes in development part (see Aniba and Aissa (2004), Wannstrom (2009)), but in recent years, the focus of recent research has been put on the packet scheduling problem. The main goal is to increase data rate and decrease number of loss data. With this in mind, we should look at succeeding paragraphs.

2. PROBLEM

The main goal of this article is to analyze the problem of Packet Scheduling while supporting the proper QoS values and propose a new solution to solve described problem. In this system packet scheduler takes place between UE and nodeB (as is shown on Fig. 1) and it is responsible for resource allocation of radio users who are in the cell scope. Resources are allocated for TTI time *(Time Transmission Interval).* Selected case is a part of on-line system, which means there is no data known in advance, but all data appear on real time. This mode of action enforces continuous data recalculation by an algorithm.

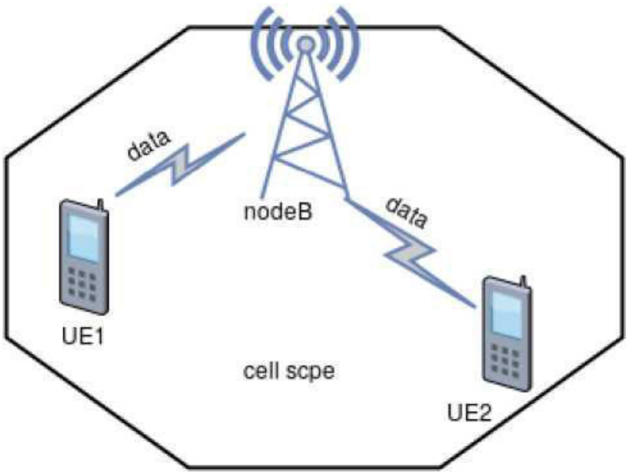


Figure 1: Main principles of tested system

Packet Scheduler plays a principal role in an access network. Furthermore the usage of correct scheduling algorithms increases throughput and it is very important in case of large concentrations of people. Some examples of that problem are offices during working hours and blocks of flats in the evening.

In this paper, we will consider the problem of scheduling packet in UMTS systems in dense population area. The main goal is to build an algorithm providing high-level of throughput, low-level of delays and lost packets while there are many UEs in the cell scope. For this purpose, we designed a simulator with the traffic model.

3. STATE OF KNOWLEDGE

Policy of Packet Scheduling can be considered as an optimization in on-line system. This problem occurs relatively often in the literature and is a permanent subject of study. Continuous research on this problem is very important for the development of telecommunication network. To begin with, take a look at the most commonly used approaches suggested by 3GPP organization:

- Round Robin (Wannstrom (2009)) – an algorithm that prioritizes packets based on the appearance time of the system; it is based on FIFO queue,
- MAX C/I (Wannstrom (2009)) – this algorithm is based on the Carrier to Interference ratio; user with the highest level of Carrier to Interference (C/I) ratio has the highest priority level in the actual TTI frame,
- Proportional Fairness (PF) (Tse (1999)) – priority level is calculated by maximizing possible UE data rate with historical value ratio.

The above mentioned solutions are the most commonly used in actual telecommunication systems. Although those algorithms do not provide the optimum solution in all cases, they are simple to implement and develop. However, researches are still looking for a better solution. In the literature we can find the following proposed approaches:

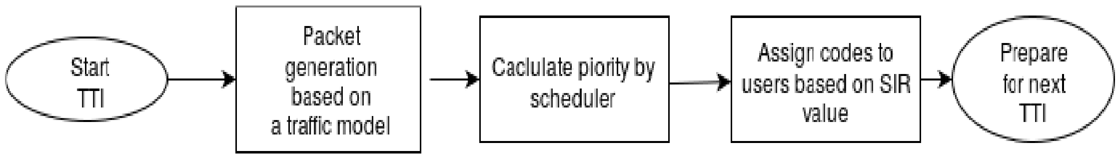


Figure 2: Flow of particular time frame in simulation

- Maximum Tolerable Delay (Yu, Kim, and Park (2007)) – this algorithm is based on maximum possible delay thereby reducing the number of expired packets. Compared to MAX C/I and PF this algorithm provides a low-level of signal retransmission.
- Hybrid Getetics Packet Scheduler (Abedi and Vadgama (2005)) – this approach provides combination algorithms proposed by 3GPP with genetic algorithm. The main profit of this solution is a low-level of delays.
- L-AWDF (Golaup, Holland, and Aghvami (2005)) – in this case algorithm deals with two parameters: packet delays and quality of UE channel. In this work and in related references it was observed that L-AWDF achieve shorter delays and less value of missing packets in comparison to MAX CI and PF.

All things considered, it seems reasonable to assume that the problem of Packet Scheduling can be solved in many ways. Furthermore, there are quite a few methods to calculate priority in telecommunication systems. The main goal of presenting solutions is to increase throughput and decrease delays. Research has revealed a problem of lost UEs with the bad quality signal. In Janevski and Jakimoski (2009) has also found that the important measure lays in the cell scope.

Even though the efficiency of those algorithms has improved in recent years, it was mainly achieved by implementing new ways of counting priority. It is possible to further upgrade the efficiency by combining existing algorithms.

4. SIMULATION MODEL

In this section, we will describe the simulation model and algorithms which were used to solve the Packet Scheduling problem in Universal Mobile Telecommunications System (UMTS) systems. General flow of one time frame is shown in Fig. 2. At the beginning of each TTI there is a generation packet based on traffic model from Section 4.3 then scheduler sets the priority to each user by using one of the algorithms described in Paragraph 4.1. It is also very important to assign codes to users based on SIR value (see for detail to Section 4.2 where simulation data and system conditions are shown).

4.1 Algorithms

UE sends packets in TTI frame to nodeB where the nodeB station is. The maximum size of packet depends on UE category, which determines maximum data rate. In this paper we will consider four scheduling algorithms as follows:

- Round Robin,
- MAX C/I,
- Maximum Tolerable Delay - TD (Yu et al. (2007)),
- new proposed solution - TDCI.

Algorithms mentioned above provide various degrees of fairness and their implementation will be described in detail below. Codes are assigned to terminals with the highest level of priority. An important implication of these findings is that all of those propositions have to work in on-line system and because of that, they are not very complex.

4.1.1 Round Robin

This is one of the three algorithms described in the 3GPP documentation (Wannstrom (2009)). This algorithm chooses the proper UE in order of appearance in the system and do not take into consideration any additional information about the terminal. This algorithm works like FIFO queue and thereby ensure handling of all packets.

4.1.2 MAX C/I

The MAX C/I algorithm is also presented by 3GPP, Suh (2007). The priority in this case is determined by Carrier to Interference ratio. This factor specifies the signal quality between nodeB and UE terminal. Higher C/I ratio means better signal which guarantees low data loss (see Scriba and Takawira (2009), Suh (2007)). The unit of measurement C/I ratio is decibel (dB).

4.1.3 TD

This scheduling algorithm sets the priority on the basis of packet size and waiting time ratio. The user with the highest priority will be handled faster. In this case the most important users are those with a high data rate but long waiting time. In our work we used the following formula for simulation purpose:

$$P_i[n] = \frac{L_i[n]}{T\max_i - W_i[n]} \tag{1}$$

where:

i UE number,

n time (measurement in time frame TTI),

$L_i[n]$ amounts of bits in buffer,

$Tmax_i$ maximum tolerable delay,

$W_i[n]$ waiting packet time in buffer.

First example of the algorithm mentioned above was presented in Yu et al. (2007). The most interesting matter of this issue is that in this scenario the buffer delay monitoring is used to prevent data loss and the occurrence of time-outs in the system. TD solution choosing the right priority value takes into account the possible UE data rate.

4.1.4 TDCI

A key limitation of this TD algorithm is that it does not take into consideration actual channel condition. To change that situation, we propose a custom modification of algorithms described above through a combination of two solutions – TD and MAX C/I. The new priority is given as follows:

$$P_i[n] = \frac{L_i[n] * CI_i}{T\max_i - W_i[n]} \tag{2}$$

where:

i, n, $L_i[n]$, $Tmax_i$, $W_i[n]$	like mentioned above
CI_i	C/I ratio

Those modifications include quality of signal between UE and nodeB, which is very important in case of high bandwidth. What is more, they take into account maximum tolerable delay which should prevent data loss in the system. To maximise data rate, calculations also include actual packet size in the buffer.

4.2. System simulation parameters

This section is concerned with the issue of simulation parameters. All system parameters are collected in Table 1. The simulation proceeds as follows: a given number of UE stations are entering into the cell scope of nodeB. Then terminals declare readiness to send packets to nodeB and the scheduling algorithm determines the users to which HS-DSCH (High Speed Downlink Shared Channel) channels should be assigned.

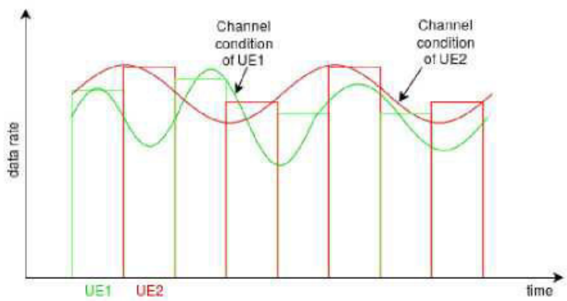


Figure 3: Relationship between channel condition and UE data rate.

Table 1: System simulation parameters

Parameter	Assumption	Comment
Cell	100m	Picocell
TTI frames	2ms	
terminals speed	0 km/h	terminals are not in move
min number of UE	500	
max number of UE	1200	

HS-DSCH is composed of 15 virtual channels (HS-PDSCH) represented by encryption codes. Those codes can be assigned to the service multi-users in a time frame depending on the SIR value as shown in Table 2. Based on received SIR value, scheduler assigns channel numbers to a proper terminal. In that way, even user with the highest priority, but in a poor channel condition can still guarantee the transmission of any information. Moreover, this user will get more channels when its own channel state gets better. The relationship between channel condition and UE data rate is given in Fig. 4.2. Higher SIR value causes better UE data rate (see for details Tanaka, Ishii, and Sao (2015)).

Table 2: SIR level

SIR level	Assigned channel numbers
$SIR \geq 27dB$	15
$27dB > SIR \geq 22dB$	9
$22dB > SIR \geq 16$	6
$16\text{dB} > SIR$	4

The data rate that achieves the maximum rate is based on UE category and computed for each user for each frame. The list of UE category parameters is given in Table 3 (see for details Holma and Toskala (2007)). Moreover, in this context, it is worthwhile to consider the large concentration of people who's not going anywhere e.g. offices during working hours. With those assumptions in mind all terminals stay in one place (terminal speed is assigned to 0 km/h) and data-only UE are just considered. Traffic network model is based on Pareto distribution and is described in Section 4.3. From another point of view, we will consider only picocells, which scope is less or equals 100 meters. This choice seems to be obvious when we have in mind offices during work hours or flats in afternoon. In this case, users are in small area and usually not moving.

4.3. Traffic Model

In this paper, we have also considered the selection of real time traffic network model. Traffic model parameters are presented in Table 4. The most interesting approach to the traffic model issue has been proposed by Yuan, Xu, Wu, and Ma (2014) and Yu et al. (2007). Based on mentioned works we applied the following distributions:

Table 3: UE category

UE CATEGORY	CODES	MIN TTI FRAME	DATA RATE (MBPS)
6	5	1	3.6
7	10	1	7.2
9	15	1	10.1
10	15	1	14

- packet call size – the size of every packet call is spread based on the Pareto distribution with the minimum size of k and maximum size of *m*. With this information in mind $f_p(x)$ is the probability destiny function is defined by:

$$f_p(x) = \begin{cases} \dfrac{a * k}{x^{\alpha+1}}, \beta = \left(\dfrac{k}{m}\right)^{\alpha}, \\ \beta, x = m \end{cases} \quad (3)$$

where α is set to 1.1, *k* value is predetermined to 4.5 kB and m is equal to 2 MB.

- packet arrival time – appearance time of packet is demonstrated by geometric distribution with an average value of 5 sec.

Based on the parameters mentioned above, we will make the assumption that the average value of the packet size is 8.5 kB Another key thing to define is a maximum tolerable delay for each packet and to simplify, we fixed this value to 100 ms.

Table 4: Traffic model parameters

Component	Distribution	Parameters
Packet calls size	Pareto	$\alpha = 1.1$, $k = 4.5kB$, $m = 2MB$ average packet call size: 8.5 kB
Packet arrival time	Geometric	average 5 sec
Max. tolerable delay	Fixed	100 ms

RESULTS

Simulation starts from 500 users in cell scope and proceeds for 60 seconds (18000 frames after which amount of terminals is increased by 100. Actual channel condition is determined by SIR values which (without loss of generality) are randomly generated. All algorithms are tested on the same environment. The results are shown in the paragraphs below.

Fig. 4 shows that the average packet data rate which is equivalent to a throughput in the system. It can be seen that RoundRobin is the slowest solution and does not exceed 4.5 Mbps for 1200 users in the simulation. From another point of view MAX CI, TD and TDCI achieved similar results, but the best channel throughput was reached by TDCI.

Fig. 5 shows that the average packet delay depends on the number of UEs in the cell scope. The analysis and simulation indicate that MAX CI algorithm caused the slightest delay lower than 50 ms for 1200 UEs. In fact, that requirement delay is fixed to 100 ms, the TDCI scheme packet loss rate is in the middle, but lower than TD and does not reach 80 ms for 1200 users.

The actual values of loss packets due to time-out are given in Fig. 6. It has been found out that proposed algorithm scheme decreases level of loss data in comparison to other described solutions which are RoundRobin, MAX CI and TD. Moreover, for 1000 users on the cell scope, value of loss data for TDCI algorithm is lower than 30 percent.

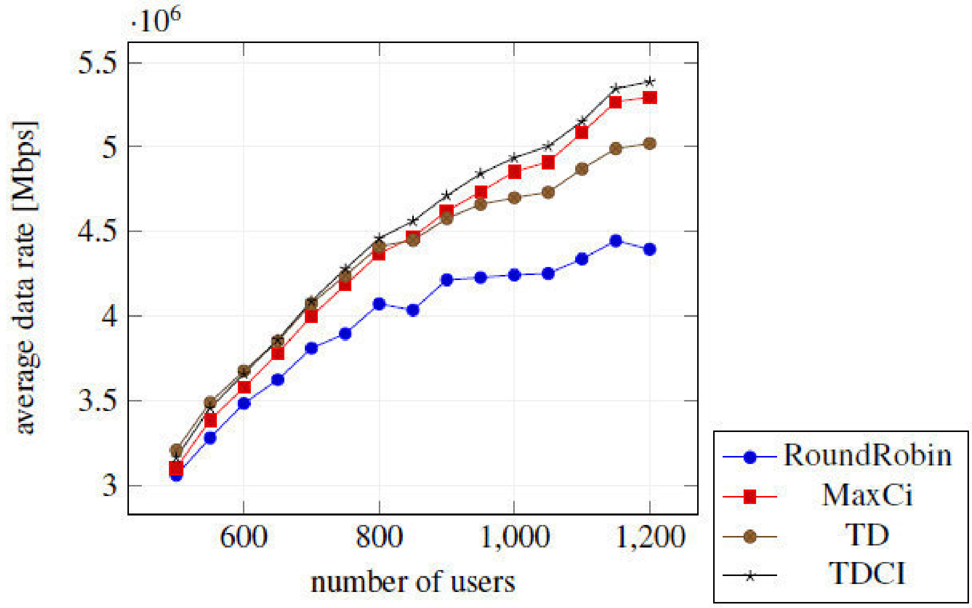


Figure 4: Average data rate for different number of users

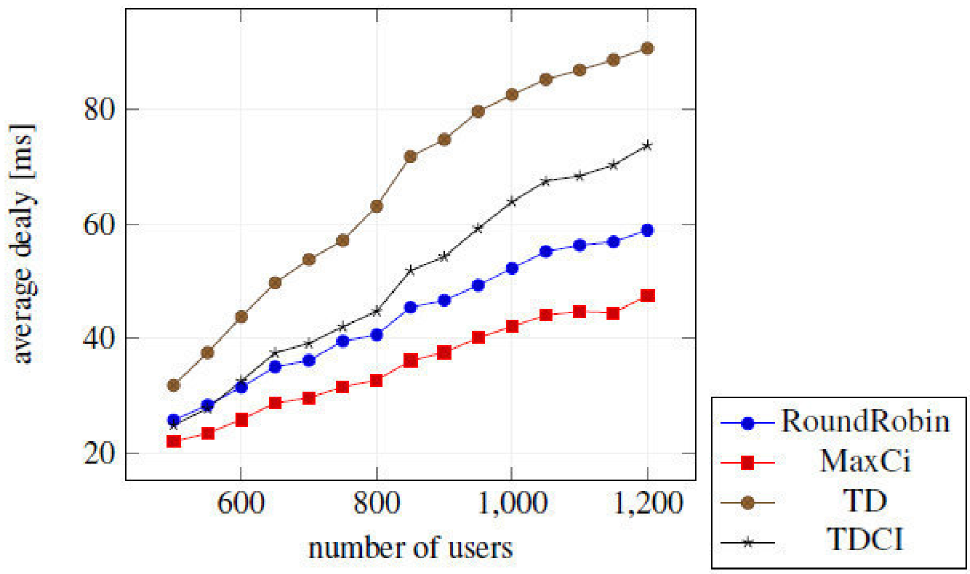


Figure 5: Average packet delay in system

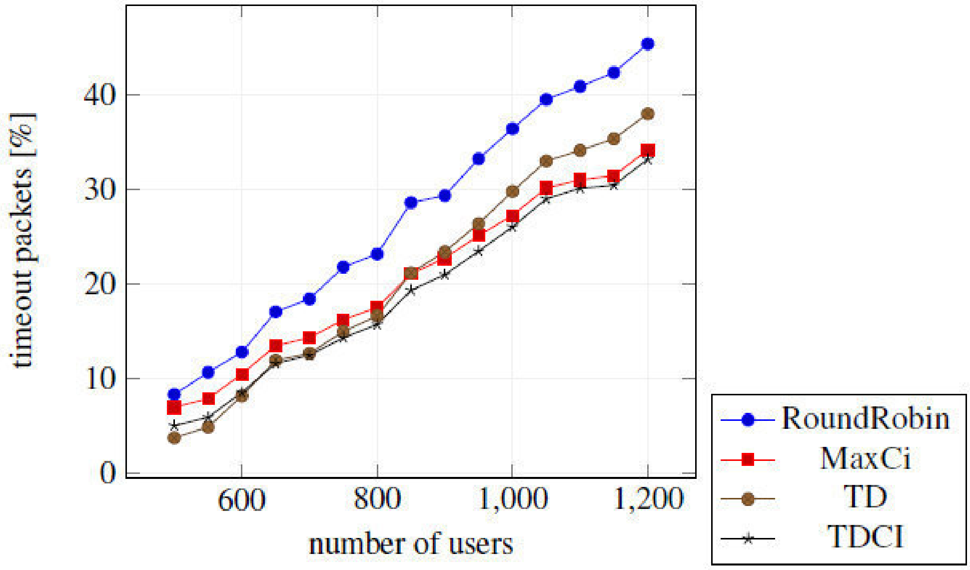


Figure 6: Percent of time-out packets

Diagrams described in paragraphs above show that TDCI algorithm provides the highest level of throughput and low levels of loss packets in comparison to other tested solutions. The proposed algorithm has advantages of Max CI and TD. CI part provides a high-level of data rate and TD piece ensures a low-level of loss data.

CONCLUSIONS

In this paper, we have proposed a Packet Scheduling algorithm which is a combination of two existing solutions. From the research that has been done, it is possible to conclude that proposed scheme can provide satisfactory throughput values and acceptable level of delays. Compared to any other algorithm, the number of loss packets is lower. Based on presented results, the proposed method can be readily used in practice. Still, further research will be required to validate presented method in LTE systems, where level of throughput value is much bigger than in HSDPA.

REFERENCES

[1] Abedi, S., & Vadgama, S. (2005). A genetic approach for downlink packet scheduling in HSDPA system. *Soft Computing,* 9(2), pp. 116–127.

[2] Aniba, G., & Aissa, S. (2004, Nov). Adaptive proportional fairness for packet scheduling in HSDPA. In *Global telecommunications conference, 2004. globecom '04. ieee* (Vol. 6, pp. 4033–4037).

[3] Golaup, A., Holland, O., & Aghvami, A. H. (2005, Sept). Concept and optimization of an effective packet scheduling algorithm for multimedia traffic over HSDPA. In *2005 ieee 16th international symposium on personal, indoor and mobile radio communications* (Vol. 3, pp. 1693–1697).

[4] Holma, H., & Toskala, A. (2007). Hsdpa/hsupafor umts- high speed radio access for mobile communications. *John Wiley and Sons.*

[5] Janevski, T., & Jakimoski, K. (2009). Comparative analysis of packet scheduling schemes for HSDPA cellular networks. *Telfor Journal,* 1(1), pp. 2–6.

[6] Kolakowski, J., & Cichocki, J. (2007). *UMTS System Telefonii Komorkowej Trzeciej Generacji.* Wydawnictwa Komunikacji i Lacznosci.

[7] Korhonen, J. (2003). *Introduction to 3G Mobile Communications.* 2nd Edition Artech House Inc. Rutkowski, D., & Sobczak, R. (2002). Uslugi w systemie UMTS. *Przeglad Telekomunikacyjny Wiadomosci Telekomunikacyjne*, nr 11–12, pp. 639–645.

[8] Scriba, S., & Takawira, F. (2009). A comparison of scheduling algorithms in HSDPA. School of Electrical, Electronic and Computer Engineering, University of KwaZulu-Natal.

[9] Suh, K. W. (2007). A generalized formulation of the protection ratio applicable to frequency coordination in digital radio relay networks. *Radio Science,* 42 (RS1007).

[10] Tanaka, S., Ishii, H., & Sao, T. (2015). Hsdpa throughput performances using an experimental HSDPA transmission system. In (Vol. 6).

[11] Tse, D. (1999). Forward-link multiuser diversity through rate adaptation and scheduling. Bell Labs presentation, New Jersey.

[12] Wang, Y. H., Huang, K. F., & Le, C. Y. (2009, May). Packet scheduling algorithm with QoS provision in HSDPA. In *2009 international conference on advanced information networking and applications* (pp. 616–623).

[13] Wannstrom, J. (2009). Physical layer aspects of UTRA High Speed Downlink Packet Access *(Tech. Rep.* No. 3G TR 25.848). 3rd Generation Partnership Project (3GPP).

[14] Yu, X., Kim, S. W., & Park, Y. W. (2007). Real-time traffic packet scheduling algorithm in HSDPA system considering the maximum tolerable delay and channel assignment. In Y.-H. Lee, H.-N. Kim, J. Kim, Y. Park, L. T. Yang, & S. W. Kim (Eds.), *Embedded software and systems: Third international conference, icess 2007, daegu, korea, may 14–16, 2007. proceedings* (pp. 540–547). Berlin, Heidelberg: Springer Berlin Heidelberg.

[15] Yuan, B., Xu, B., Wu, C., & Ma, Y. (2014). Mobile web user behavior modeling. In B. Benatallah, A. Bestavros, Y. Manolopoulos, A. Vakali, & Y. Zhang (Eds.), *Web information systems engineering – wise 2014: 15th international conference, thessaloniki, greece, october 12–14, 2014, proceedings, part I* (pp. 388–397). Cham: Springer International Publishing.

Service Oriented Scenarios for Support Enterprise Project Management

Henryk KRAWCZYK[1], Narek PARSAMYAN[2]

[1]*Gdansk University of Technology, Faculty of Faculty of Electronics Telecommunication and Informatics Gabriela Narutowicza 11/12, Gdansk, Poland, e-mail: hkrawk@pg.gda.pl*

[2]*Gdansk University of Technology, Information Centre of Tri-City Academic Computer Network Gabriela Narutowicza 11/12, Gdansk, Poland, e-mail: nparsamyan@task.gda.pl*

Abstract: In the paper, a service oriented architecture is used to describe various scenarios of human-oriented enterprises. They can be modeled by BPMN (Business Process Model and Notation) to show the basic information flow. That approach gives an opportunity for the right management through taking the best possible choice of services, or composition of services called service orchestration or service mashup. We explain the idea of such techniques, and present their usefulness for description of service based supply chain activities, first of all the decision-making process. The replacement of all human activities by IT services is not entirely possible; however, computer aid is desirable for an immediate flow of comprehensive data. It is also emphasized that the proposed method can be used to implement various enterprise scenarios.

Keywords service-oriented scenarios, BPMN, e-services

1. INTRODUCTION

Enterprise Project Management (EPM) [9] supports organizations in their development, according to the assumed goals and strategies describing how the goals can be achieved. Each such strategy involves two major phases: formulation and implementation. The first one is related to situation recognition, and guides developing policies. The second one refers to suitable actions to achieve the established goals. In the case of IT enterprises, which use IT infrastructure (hardware and software) to meet the demands of large organizations, we can distinguish an extra maintenance step. The aim of maintenance is to keep IT infrastructure to run reliable, and be less likely to break down, ensuring effective performance and minimizing risks. Such solutions have become an increasingly important part of modern organizations. It is necessary to describe how an organization can transform the processes in order to maximize business value, including economic profit and well-being in the long run. Wide utilization of SOA technology [5] causes that an organization can operate as a service-oriented one. In consequence changes in organization are enabled to better govern, manage, and secure IT services offered by e.g. computing clouds [6]. What is important to know is that different kinds of services are required in each phase of the organization lifecycle. For instance for the formulation phase, various IT services related to data gathering, transforming and adjusting should be collected. While in the implementation phase, different services supporting developing functionality are defined. In the maintenance phase, monitoring, and recovery services are largely used. All such services can be divided into two categories: manually and automatically performed services. The first of them will be called services, and the second one e-services. There is a dilemma how to choose and use the available set of services, and sometimes how to design and implement them to create the required functions of an organization. The organization where services play dominant roles can be called a service-oriented one. The proposed solution differs from the traditional, including those where the maturity IT management systems are such as ERP, MRP, and CRM. In comparison to a service-oriented approach, the traditional one is rather closed, and not able to easily adapt to changing conditions.

In the paper, business and management patterns are described by service-oriented scenarios. Such scenarios integrate people, information and appropriate applications (web services). What is important, according to service-oriented-architecture (SOA) required e-services are self-governing, and may call other services to implement complex tasks. We limit our consideration to the formulation phase where suitable implementation scenarios consisting of services and e-services are considered. It was shown how to model such scenarios using

BPMN (Business Process Model and Notation). Besides this, a popular kind of e-services integration is described, based on mashup technologies. It allows the building of complex suitable scenarios consisting of simple e-services in order to support different kinds of business enterprises. In section 2, enterprise modeling using BPMN is considered. Some aspects of selling organization resource management are analyzed. In section 3, enterprise as service architecture approach is taken into account, and well promising technologies called mashup are discussed in the case of decision-making applications. Section 4 describes the management problems of service oriented enterprises scenarios, and gives examples of how to manage such kinds of activities. In conclusion, the results of the paper are summarized, and some propositions for future investigations are suggested.

2. MODELING ORGANIZATION PROCESSES

To generate bigger profit from the organization of business activities, it is necessary to recognize the processes taking place in the organization, and to understand them. The processes are often cross-functional so could be used alternately, including both human services activities, and computer activities (e-services). Their proper relations can affect whether the organization is able to succeed. Generally, a business process is a sequence of actions designed to produce a product or service that is of value to the customer, whereas business process modeling is simply representing the process in graphical form. Of course, it is important to maintain the sequence logic, without it the model could be misinterpreting or incomprehensible. We can distinguish the most popular and official process modeling notations UML (Unified Modeling Language) and BPMN (Business Process Modeling Notation). The alternative way is not to use notation, but standards like BABOK (Business Analysis Body of Knowledge) [7], which is a guide for correct process modeling, including different knowledge and various techniques, published by International Institute of Business Analysis (IIBA).

Below is an example of process analysis and process modeling for a company selling its assortments in a stationary store. The Figure 1 showing the information flow among the company is organizational departments of logistics, warehouse, store, which operate in correlation with manager and customer. The presented relationship is strongly focused on logistic time optimization, and increasing customer satisfaction. Joint and close cooperation inside the organization will give more valuable and attractive products for a customer. The computer systems are distributed among organizational departments, and some coordination (management decisions) is needed for proper process executions. It cannot be done manually, so the well-known approaches client/server or peer to peer can be recommended.

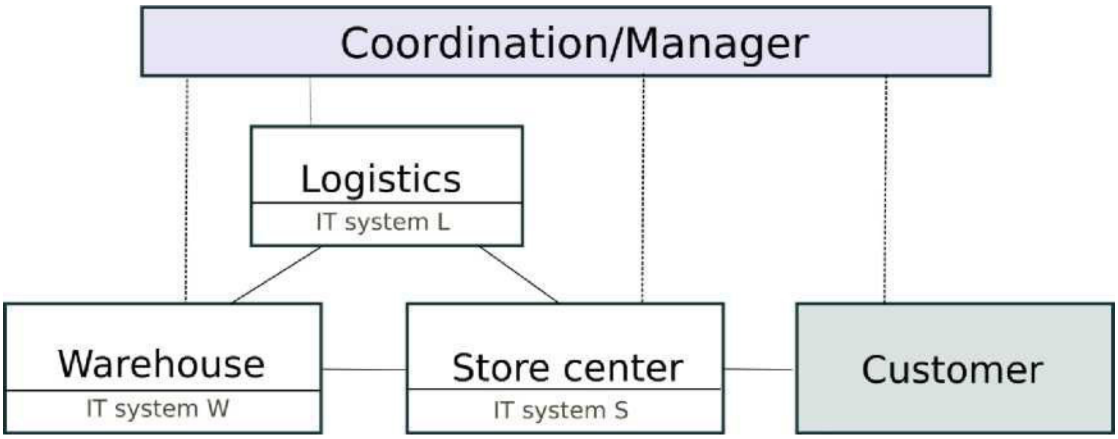


Figure 1: Information flow among departments of the selling company

Conducting a relationship analysis and current operations is an essential step in building a process model. Figure 2 shows a model of the selling company from Figure 1, created in a BPMN standard. BPMN describes a workflow with a set of graphical tools, and rules for their use (semantics and syntax notation), it allows us to see visual representation of the process, and related artefacts. Standard BPMN 2.0 was published by OMG (Object Management Group) in 2011. It is currently the most popular standard for describing business processes, used to accurately describe business processes at a technical level. Using BPMN we can describe three areas of organization: processes, procedures, policies (3P approach). Let us consider one scenario, to create supply streams of products (from the warehouse vis logistics to store) which are mostly sold in the store.

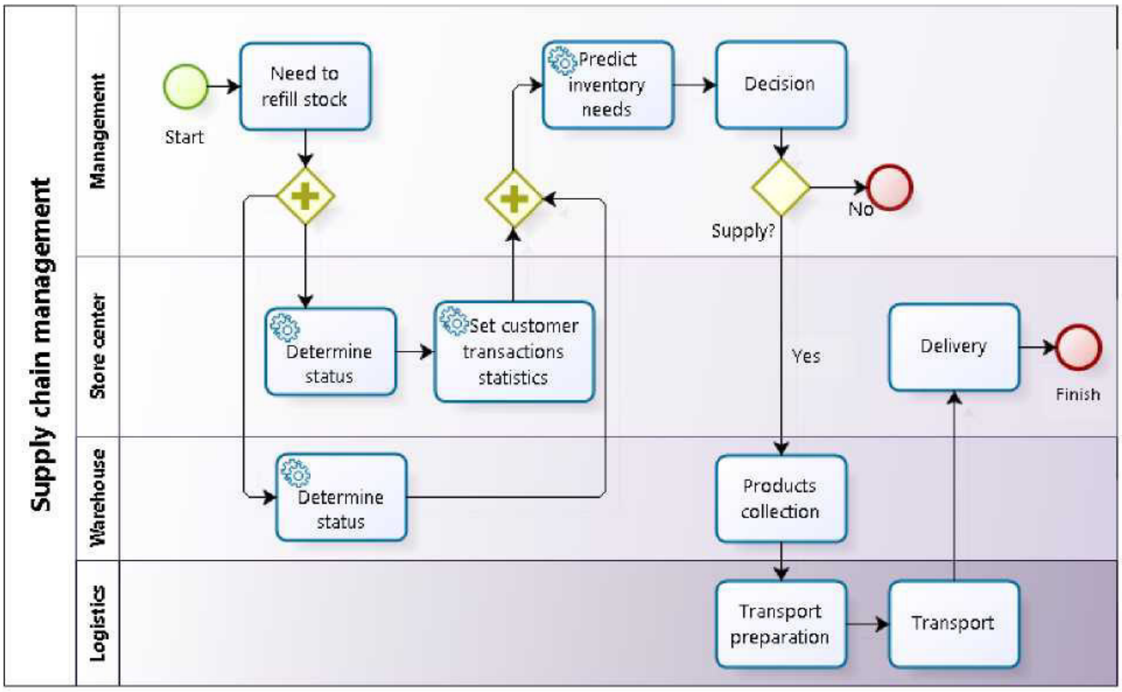


Figure 2: Process flow in a selling company expressed in BPMN

The process presented in Fig. 2 initiates the need for a stock refill, and forecasts the future demand for products offered in store centre. In the next step, with the support of information systems, a parallel analysis of supply taking place. The well prepared e-services determine the status of store and warehouse, and also shows the sales historical data (transaction statistics). They collect data and analyse them, so that the manager is able to predict future demand for products, and is also able to undertake the most optimal supply decision. If the manager makes an order, the information is dispatched to the warehouse department, where the workers collect the products to send to the store. When the product is ready to move, the information is sent to the logistics department, which is starting the transportation process. By supporting the logistics process of IT systems, entrepreneurs do not have to worry about the lack of assortment in the store. Such scenarios can be executed when more than one product was sold. This approach is Just In Time (JIT) methodology, which assumed that stocks should be available when the company needs them, and not earlier or later. Generally, the aim of using IT technology in logistics processes is to manage the inventory, eliminate waste, and ensure timely delivery of orders to the customer. In practice, it is much more complicated, and many IT subsystems can be enrolled in such scenarios [1].

Currently many small, specialized, external services providers offer different e-services, which can be used to create various business scenarios for different organizations. Moreover, they can be used as cloud services [4] and then they are more flexible, making them procure, integrate, useful and adaptive. The organizations decide which parts of their processes they can make themselves, and which parts can be made with external providers. Then, new methods to manage multi supplier e-services can be defined. Problems, such as consistency of standards (interfaces), run on-demands, security achievement could then be solved.

3. SERVICE-ORIENTED APPROACH

Let consider a service-oriented approach [8], where service management is across organizational boundaries, in comparison to traditional closed system management (ERP, MRP, CRM). The progress of running services can be monitored by extra services, performing within their operational organization environment. In such cases, some information about services, and their characteristics, is required. In other words, services should be more measurable and manageable. Several standard management frameworks are currently used, such as SNMP (Simple Network Management Protocol), WBEM (Web-Based Enterprise Management), WSMF (Web Service Management Framework) [9]. The latter provides support for discovering, inspecting, securing and invoking resources, management functions, infrastructure services and tool sets. The standard management interfaces enable higher-value utility functions, processes or applications. They concern aspects such as availability and performance management, optimization, capacity, planning, billing, configuration management, asset protection, problem determination and business analysis. They allow one to describe cooperation among departments for instance the whole selling organization. Service-oriented organization in accordance with SOA principles is shown in Figure 3.

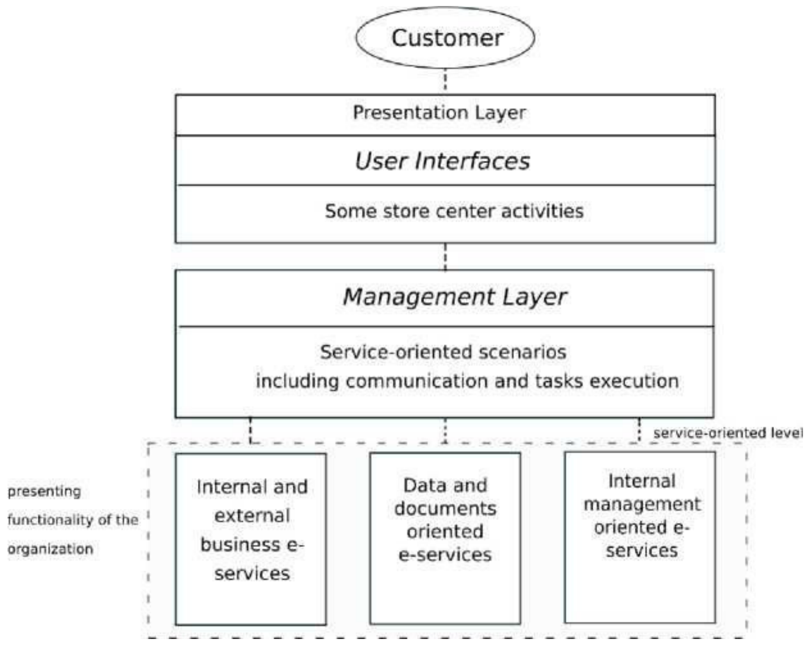


Figure 3: Architecture of service-based supply chain system

The business scenarios are interpreted by basic services: such as market analysis, shipping, order processing and inventory services. The management of services resources is through their interfaces and infrastructure services. They typically provide content for management, and managers interact with resources. It easy to note, that the scenario presented in Figure 2 can be easily implemented, managed, and executed.

Management composition of e-services can also be managed on both high and low levels. High level policies would address composition, including user's preferences to integrate into e-services, which is required to notify their performances, and when to track Web services execution progress. Low-level policies would include:

- how to exchange comprehensive information among Web services,
- how to deal with Web service unreliability,
- how to substitute e-services with other peers without disrupting the execution flow,
- how to suspend an e-service execution.

The risk that behavior might change the information, that would be intercepted. An open problem is also how to determinate the required set of all such e-services.

The Language WS-BPEL (Web Service Business Process Execution Language) can be used to describe and execute such management activities. Different services can be used to realize SLA (Service Level Agreement) and provide the suitable service management. An interactive monitoring and reporting can be also organized in that way.

4. SERVICE-ORIENTED INTEGRATED SCENARIOS

To create service-oriented enterprises, we should use different kinds of e-services altogether, in contrast to human-oriented services (services), and point out the available data sources for processing. There are two different approaches to composite e-services, orchestration, and choreography IT [3]. Orchestration refers to business scenarios that can interact with both the interface and external e-services. It determines service execution under the message transfer level. In general, it is managed by one extra unit, that coordinates the work of others. The specification called BPEL (Business Process Execution Language) following from BPMN (Business Process Model Notation), models the behavior of services in a business process interaction, very often by using workflows. First of all, it allows handling receiving, replying, invoking e-services. Choreography refers to peer-to-peer models without any coordinator. The collaborative protocols are described by WSCI (Web Services Choreography Interface) that allow one to define message exchanges between services. It supports message correlation, sequences rules, exception handling, transaction and dynamic collaboration, which we will consider as orchestrating e-services. The representative case study can be a system scenario in a selling company which built its selling propositions using lists of available supplier's offers. It is likely, that such a list is defined by managers who by watching over business operations, can identify all opportunities and serious problems as they occur. Thus, they can ensure that there are e-services supporting such business scenarios, which are performing in accordance with service level objectives. Of course, the buyer will decide about, acceptation, or to refuse the suggested proposals (list of products). Another shopping scenario is that a seller automatically informs suppliers about products that are the most popular for buyers. Then, suppliers can adapt their product-gathering strategies. It is possible to model the above-considered scenarios by BPMN (see section 2) and analyze their behavior by BPEL. Besides, special languages can be offered, e.g. WS-DL (Web Service Choreography Description Language), to describe such scenarios [3].

Management of such a distributed environment requires reliable and efficient extra services for both scenarios, business development and management. Configuring services into discrete logical components aligned directly with business execution and management functions can create a new architecture of the organization. We can orchestrate various services into many different configurations to support business processes. Management services, similarly to execution services, involve collections of services that communicate with each other in order to pass the same data on to coordinate the same activities (see Fig. 3).

A mashup is an approach that uses content from more than one source to create new services displayed in a single graphical interface [2]. For example, a seller could combine supplier offers with opinions of buyers to create new services belonging to the group of recommendation ones. In the simplest example, the proper kind of services can be prepared. Moreover, the mashup makes existing data more useful, using it to communicate, for visualizations and execution. What is important, mashup composition tools are usually simple enough to be used by an end-user without programming skills. Only some choice of GUI widgets, and access to various services is required. We can distinguish three kinds of mashups: enterprise (business scenarios), consumer (behaviors) and data (collection) ones. Enterprise mashup defines scenarios that integrate their own resources services and data with external available services, focusing data into a unique representation. In such a way they allow for collaboration among businesses and developers. Customer mashup combines some data from multiplate public access to them by browsers. Data mashup concentrates on focusing transformation of similar types of data, and transforms them into a single representation. In Figure 4, according to the system model presented in Figure 2, there is shown the relationship between different kinds of services, in order to create one complex service with integrated data. Such an approach is called a mashup solution.

The services shown in Figure 4 that support the process from Figure 2 and Figure 3 are described in more detail below:

Service A: Product description. Service provides particular information about the product, e.g. name, function, price, producer.

Service B: Customer transaction. Service provides a set of statistics containing specific information about product sales history, e.g. transaction date, number of sold products, delivery time.

Service C: General status. Service provides information about the numbers (levels) of particular product in store and warehouse.

Service D: Predict inventory needs. The result of data combination from A, B and C services, e.g. sales statistics, delivery time and store inventory level, can be represented in the mashup service.

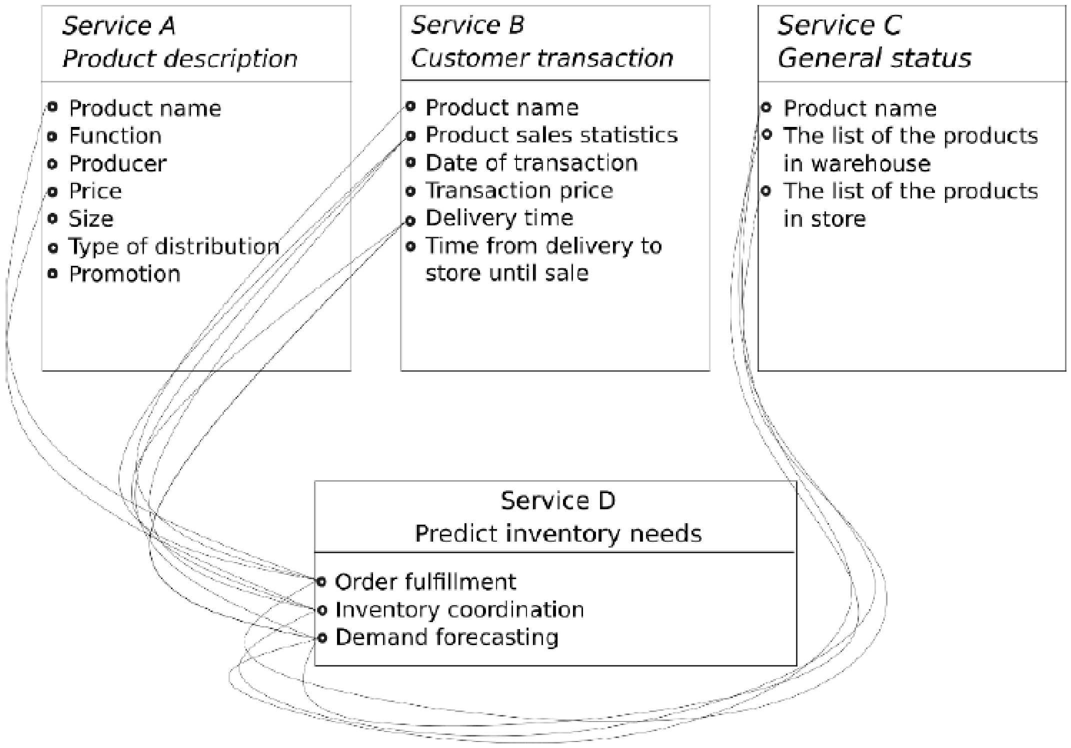


Figure 4: Mashup service for selling company

The architecture of mashup is also divided into three levels. The presentation layer support user's interaction by technology like HTML, CSS, JavaScript. Service level represents main functionality by using API services: XML-RPC, SOAP, REST. Data level manages orders, receiving and storing information using XML technologies. Data level compares to different sources of data; is presented and plays an important role when IoT (Internet of Things) technology is used. In Figure 5 the example of a services scenario is given, where *A*, *B, C* services support decision-making process in a selling organization. However, the manager, based on these services, should analyze and predict inventory needs and take a responsible decision. To make just-intime the proper decision, extra services should be added. However, the final decision belongs to the manager of the considered organization.

Due to mashup technology, different services can be reused within other services to provide higher business flexibility. Moreover, supported data services can be also unified and integrated. Then the management tasks are needed, to coordinate and manage execution of all services supporting the organization activities. It corresponds to the service-oriented scenario shown in Figure 5. Due to use of e-services, all decisions can be taken in a shorter time, and can be more efficient, because inventory maintenance cost will be significantly decreased. In a similar way, we can improve other different organizational activities. However, we need to create new e-services which support or execute required human tasks.

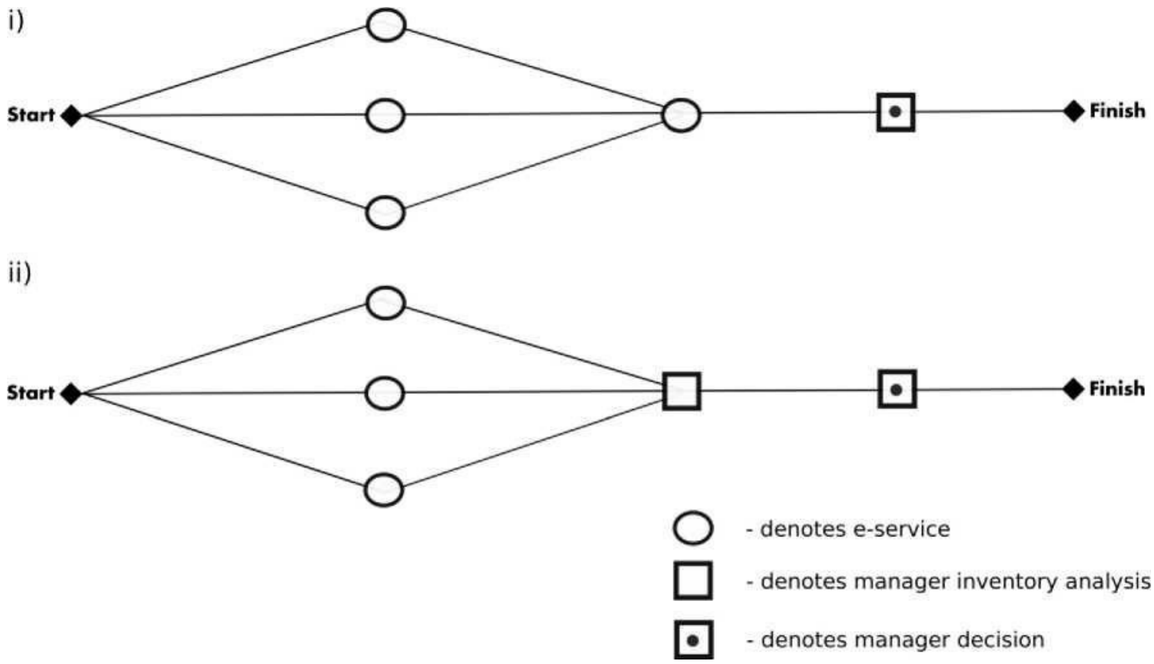


Figure 5: Example of two management services scenarios:
i) based on e-services ii) mashup technology.

5. CONCLUSIONS

In the paper service-oriented processing for business organizations is presented. Besides, how to model and implement such structures is given. The new flexible service-oriented architecture is proposed, and it was shown how to describe and manage it using Web technology. Moreover, propositions of the mashup approach are also discussed, to create complex e-services and interpret them into service-oriented scenarios, which can be also easily modified and managed. It allows you to create business-oriented systems on the base of cloud computing technologies. The need for modeling has also been underlined, the example of using standard BPMN was presented and described. The proposed modeling approach gives an opportunity for better understanding of the whole organization processes, and formulating a chosen activities as e-services. In the case of complex services (consisting of other services) a mashup approach can be successfully used. In the paper we concentrate on technological aspects, which allows you to achieve general approach for description and implementation in various kinds of organizations. Additionally, implementation details of the proposed solution can be analyzed and evaluated. Moreover, other aspects of the enterprise projects management, related to both human activities (briefly, services) can be taken into consideration.

6. REFERENCES

[1] Bardi, J., Coyle, J., Langley, J. (2010) *Zarządzanie logistyczne,* Warszawa: Polskie Wydawnictwo Ekonomiczne.

[2] Braga D., Ceri S., Daniel F., Martinenghi D., (2008) *Mashing Up Search Services,* (pp. 16–23) IEEE Internet Computing 12(5).

[3] Decker G., Maria Zaha J., Dumas M., (2006) *Execution Semantics for Service Choreographies,* (pp. 163–177), WS-FM.

[4] Erbes J., Graupner S.,Motahari-Nezhad H.R, (2012) *The Future of Enterprise IT in the Cloud,* (pp. 66–72) IEEE Internet Computer, vol. 45.

[5] Erl, T., (2005) Service-Oriented Architecture (SOA) Concepts, Technology and Design, *The Prentice.*

[6] Hitt, M., (2013) *Management (2nd Editon),* Pearson Education Limited.

[7] *Schedlbauer, M. (2010)* The Art of Business Process Modeling: The Business Analyst's Guide to Process, Modeling with UML & BPMN., USA: The Cathris Group.

[8] Talal, H. Noor., Quan, Z. Sheng., Bouguettaya, A. (2014). *Analysis of Web-Scale Cloud Services,* (pp. 55–61). IEEE Internet Computing 18(4).

[9] Weinong W., Zhang Y., Xiaoli L., (2005) *Policy Lifecycle Model for Systems Management,* (pp. 50-54) IT Professional, vol. 7.

Decision-Making in Virtual Software Teams Using Cloud Platforms

Beata KRAWCZYK-BRYLKA[1], Henryk KRAWCZYK[2]

[1]*Gdansk University of Technology, Faculty of Management and Economics, G. Narutowicza Street 11/12, 80-233 Gdansk, Poland, bkrawczy@zie.pg.gda.pl*

[2]*Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics G. Narutowicza Street 11/12, 80-233 Gdansk, Poland, hkrawk@pg. gda.pl*

Abstract: Software development projects are usually realized by traditional or virtual IT teams using computing clouds. Team collaboration requires decision-making regarding essential aspects of a project progress. The article concerns methods of decision-making process in the case of traditional and virtual teams' work. The research results conducted in a group of IT specialists are presented, and to analyze their preferences in decision-making methods, four possible cases were tested: hierarchical choice, team members' choice (consensus or voting) and external expert decision. The additional value of the research is the recognition of IT specialists' attitudes towards computing cloud usability in IT projects.

Keywords: software development, IT teams, virtual collaboration, cloud computing, decision-making

1. INTRODUCTION

A traditional team is a group of people linked in a common purpose, where its members have complementary skills. That allows them to maximize their strengths, and minimize their weaknesses, in order to improve their performance, and in consequence to solve their problems or to finalize successfully their tasks. In the case of IT teams, the main task is the software development of specified applications. A virtual team is also a group of individuals who work together, their tasks are interdependent, inseparable, and leading to a common goal, but they are geographically dispersed and their communication depends on ICT platforms (Dâvidekovâ, Hvorecky, 2017). However, it is not easy to evaluate how virtual a team is. In general, there are three predominant factors that contribute to virtuality. The first one is the possibility of IT technology, as was mentioned above. The second one is how big is the distance between team members, in geographical dispersion and in consequence in time zones. The third one is organizational and cultural diversity which affects team members' relationships and team output.

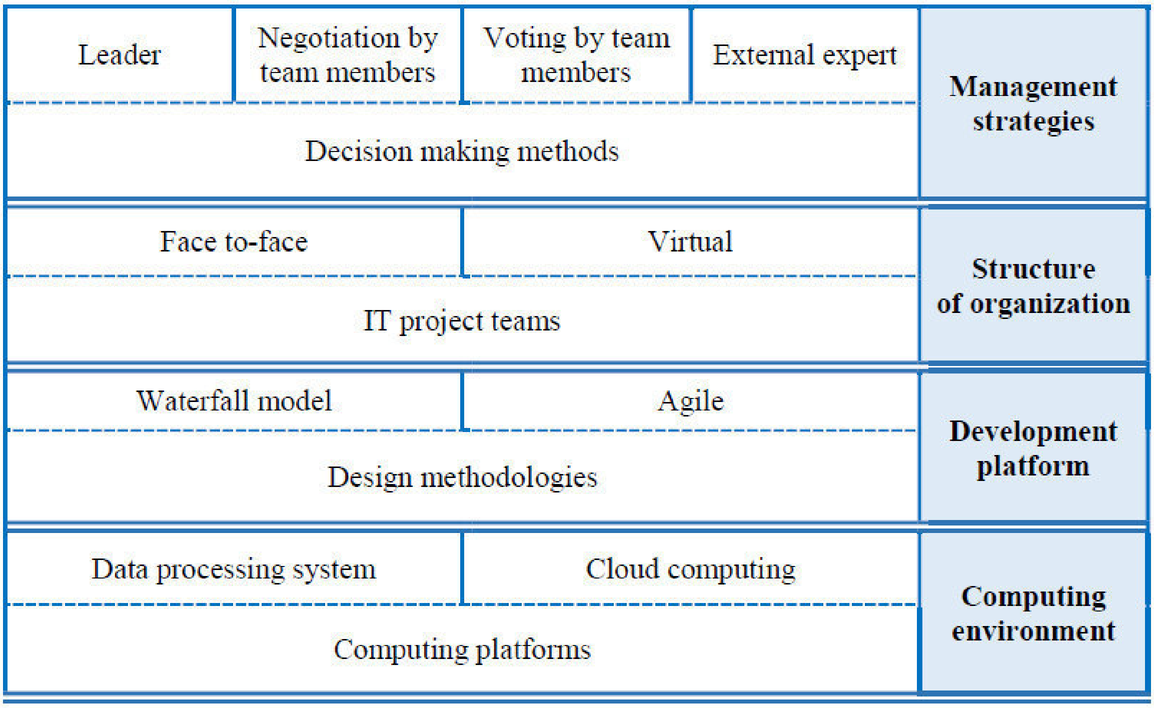


Figure 1: The main aspects analyzed in the paper.

Traditional or so-called face to face teams (ftf), and virtual ones, are still under research and examination. It is a known fact that virtual teams generate more ideas compared to traditional ones. However, virtual teams require a longer time to reach a decision. When comparing performance of ftf and virtual teams the opinions about them are mixed. Some empirical research shows that real teams are better (Schmidt, 2001), others that virtual teams, and other research confirms that they are on the same level. Progress in communication media (e.g. 3D virtual environments) brings closer these two kinds of teams (Bourgault, Dronin, Homel, 2008). We will concentrate only on decision-making processes accompanying work of such teams. We consider the same decisions, taken in both kinds of teams, and compare them to distinguish differences. We cannot analyze why some of them make correct decisions, while other teams reach a wrong decision. There are many causes of these, such as: quality and experience of teams, their leader skills type, conditions of working environment, social and surrounding context and current circumstances, existing relations among members or working atmosphere and also organizational and technical support. During the last decades researchers have found many conditions under which groups make better decisions than individuals. However, in the case of decision-making in virtual teams many problems are open and need more research.

In the paper we limit our consideration to some aspects of ftf and virtual IT teams, as is shown in Fig. 1. Some elements of computing environments, development platforms, structures of organization and management strategies can be analyzed. More than 50 IT teams have been tested by a specially prepared questionnaire, and different methods of making decisions in the above context are analyzed.

2.PROJECT PLANNING, DEVELOPMENT AND MANAGEMENT

The aim of project management is to achieve all of the assumed project goals with the specified constraints. We can distinguish two basic strategies of project development. One concentrates on the planning and design stage, where the project is defined clearly, and the required resources (including human, financial and technical ones) are strictly determined (see Fig. 2i). The second one describes Execution & Improvements, where the project is consequently realized, and in the case of some problems, properly modified (see Fig. 2ii).

Figure 2: The Project development stages: Planning&Design (i) and Execution&Improvements (ii)

Proper selection of the project team members and other resources in the first stage of project development significantly impacts on the efficient realization of the second stage. In the literature there are many propositions about how to organize the best conditions for project realization. There are many possibilities to choose suitable development platforms offering designing tools and the computing environments for software design, execution and testing (Philips, 2004).

Currently, the most popular technology has become cloud computing (Vasan, 2011). It offers various kinds of services assigned to three layers IaaS (Infrastructure as a Service), PaaS (Platform as a Service) and SaaS (Software as a Service). We can arrange, in comparison to traditional data processing platforms, more flexible and much cheaper developing possibilities on the PaaS level (Cohen, 2013).

The Execution and Improvement stage of software development is much more difficult to implement. The best practice and/or well-known principles must be kept in mind during this stage. Moreover, a team (during the work) should be aligned to the business goals, and should consider all changes in real time. One of the well-known development methods is the PRINCE 2 approach, which provides a method for managing projects with a clearly defined waterfall framework (Pollice, Augustine, Lowe, Madhur, 2004). Each activity is defined as a process with key inputs and outputs and with goals to be carried out to deliver a project's outcomes. Besides, it provides a platform for the cooperation of all participants of the project, and defines their roles and responsibilities. Also, there are some other international standards (e.g. ISO/IEC/IEEE) that present guidance on project development and management. One of the main purposes of development methods is to accelerate product delivery and reduce production costs. These requirements stimulate a new kind of software development called agile principles. Most agile development methods break development work into small parts called increments, that are planned and implemented in short time frames (e.g. 1–4 weeks). Each product of such an iteration is evaluated and improved and after acceptation it is integrated with earlier accepted parts. The most popular of such methodology includes such frameworks as Kanban and Scrum (Shalloway, Beaver, Trott, 2009). Such methods, in contrast to waterfall models, focus more on producing working software, and less on documentation. Evaluation reports are saying that agile software development helps teams to deliver software faster, improves their ability to manage changes suggested by customers and leads to an increase of team productivity. It seems to be a preferable approach for virtual project teams.

3. DECISION-MAKING MODEL

A team, to do its work professionally should be well organized, truly engaged in doing the work, and effectively managed. Teamwork process involves a series of interactions, information exchanges, questioning, answering, creating solutions and solving problems. There are two kinds of such problems: task-oriented, that corresponds to actions directly related to project purposes, and another, called team-oriented, referring to personal conflicts, unproductive communication, low members' motivation, etc. In both situations, the proper decisions should be made, which will eliminate or reduce negative influences on team output. In the paper we focus on task-oriented problems, but the well-known obstacles related to team relations and team climate must not be forgotten. The most important ones are, for example:

- Lack of team identity
- Low participation in a team
- Poor communication
- Ineffective team leadership
- Obstructive conflict
- Group thinking
- Absence of creativity

- Ineffective decision-making.

The above obstacles are typical for traditional teams but can also affect virtual team collaboration. In the case of virtual teams, some other challenges appear like: potential misunderstandings and misinterpretation caused by non-verbal communication gaps, lower team cohesion and trust, cultural diversity problems, social isolation or difficulties with brain-storming processes (Zofi, 2012).

It is possible to define some pattern actions to eliminate the above obstacles. One of them is to complete all the stages of the decision-making process that lead to the best problem solutions. It refers also to task-oriented problems, that in software teams can be oriented on:

- Choosing the platforms adequate to the project needs,
- Pointing out design methodology that is well-known and satisfactory for the team,
- Finding the research results to solve met difficulties.

They are often supplemented by the problems related to cloud computing usage in software development teams.

There are some decision-making models known, like Tuckman' s model composed of five stages: forming, storming, norming, performing and adjourning, or the Fisher model containing steps like: orientation, conflict, emergence and reinforcement (Turban, Liang, Wu, 2011). The new suggested model, matching software team decisions, is shown in Fig.3. It is assumed that decision-making consists of four main steps called: problem, knowledge, alternatives and solution.

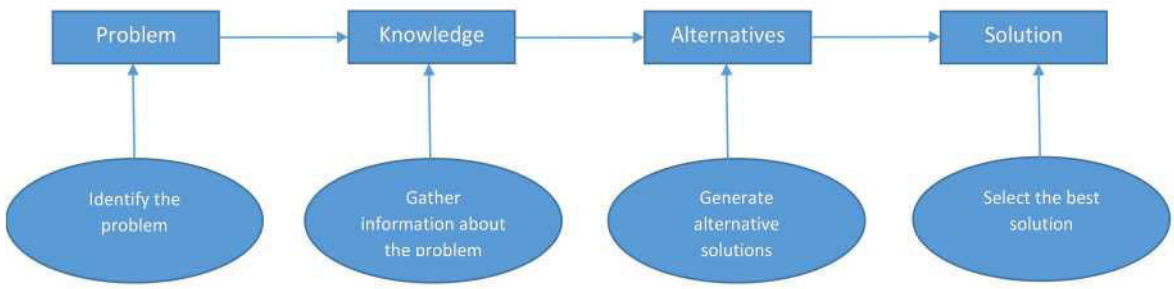


Figure 3: Four step sequence of decision making

First of all, we should identify the problem and recognize the reasons for its occurrence. It is required to evaluate how the team is achieving its results, and to understand what kind of problem it is: task- or team-oriented. Next, the information explaining the problem context, essence and possible elimination methods. To find the best solution, several alternative solutions should be defined, analyzed and compared by the assumed criteria. The last step is to select the best resolution which next can be implemented. Going effectively through all the stages can be time-consuming and sometimes needs to be triggered by a team leader or a coordinator. Such a person can be also especially useful at the last stage when the final decision is made. There are four ways of decision-making used at that phase:

1. By the leader of the team, who gather information from members, and uses it to make decisions
2. By consensus, where every member of the team must agree to adopt a proposed decision. If it is not possible, new alternatives are formulated and presented for evaluation.
3. By voting, where members of the team discuss the decision and then vote, the team accept the choice if more than 50 % of members accepted it.
4. By extra expert, where nobody from the team is responsible to recognize how useful the solution is, and to solve it professionally.

It was discovered in previous research that the decision paths are influenced by two main factors: task characteristics and group internal structures that define work relationships in the team (Wei, Heckmann, Crowston, Li, 2011). Because of virtual team specificity, it seems probable that the way of making a final decision also differs in ftf and virtual teams. Collaboration technology used by a virtual team can encourage users to categorize information, to judge its importance, and leads to better quality decisions (McNamara, 2008). When asynchronous communication ICT is used, team members have much more independence in developing and testing their own solutions, before sharing them with the other members. At the same time they are less tended to discuss all of their ideas of problem solutions (Wei, Heckmann, Crowston, Li, 2011). Virtual collaboration process structure supported by a communication channel leads to equal information voicing in the team, which stimulates the negotiation process and focus on analyzed alternatives, not on personal preferences (Cordes, 2016). Because virtual team members commitment and team cohesion are the important challenges, the consensus decision model is fully recommended (Falkowski, Troutman, 2005).

Does it mean the decision-making methods are different in virtual and ftf teams? Answering the above question was the main goal of the research presented in the next part of the paper. Another aim was to test IT specialists' attitudes towards cloud computing. In other words, we concentrate on recognition of how project virtual teams are making decisions working in cloud computing environments, and what kind of design methodologies are more preferable.

4. RESEARCH RESULTS

In order to illustrate how decisions are made in traditional and virtual teams, a questionnaire was distributed in a group of IT specialists. More than 50 of them decided to answer the request. Most of them are men: 48/54 persons, and 55% of respondents have experiences in virtual collaboration. Their age structure is presented in Table 1.

Table 1: Age of respondents

age	
18-24	17%
25-34	74%
35-44	9%

Respondents were asked to point out the kind of methods they use while working in software projects and about 65% of these IT specialists matched agile solutions (traditional waterfall method was pointed out by 24%). The answers to question about which computing and development platforms are more favorable were much more diverse, and opinions about them are presented in Figure 4.

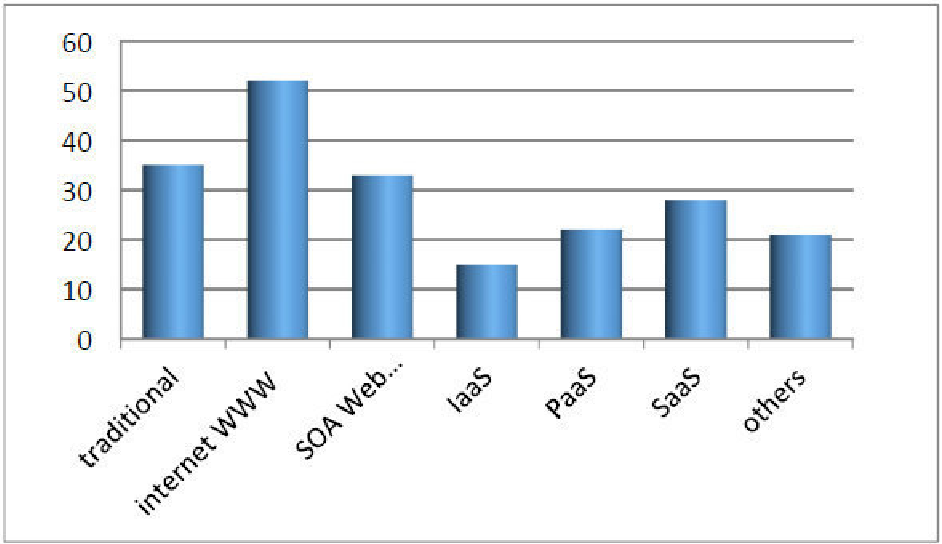


Figure 4: Computing platforms used by IT specialists (% of respondents)

More than 43% of the tested IT specialists use clouds in their professional tasks, and 77% of them would use cloud computing in realized projects if possible. Figure 5 shows respondents' preferences towards cloud computing usage, depending on their previous experiences.

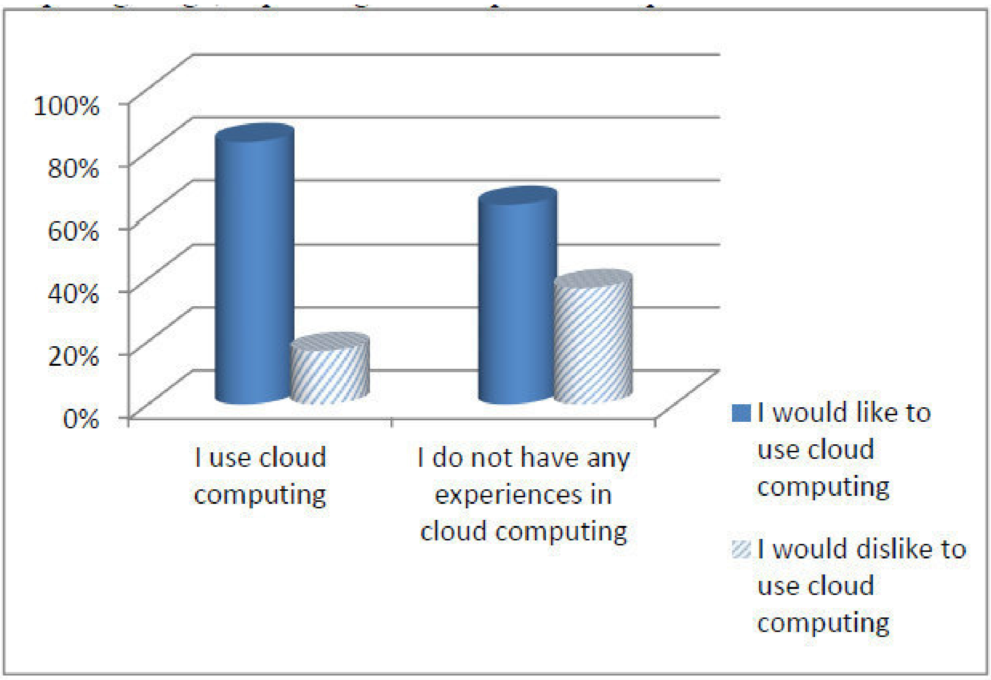


Figure 5: Cloud computing experiences and preferences

It can be assumed that IT specialists' attitudes towards cloud computing are positive, most of them would like to use it if possible. Previous experiences in cloud computing raise the preference to use it in future projects. Respondents were also asked about the reasons for their acceptance or rejection of cloud computing solutions. Pros and cons were categorized into a few groups that are presented in Table 2.

Table 2: pros and cons of cloud computing

pros	cons
Calculation speed	No need to use
Scalability	Security issues
Software expense elimination	Loss of control
Security	Knowledge gap
Convenience	

The most often-mentioned arguments against cloud computing were the lack of need and security issues. IT specialists (who are against) do not trust clouds, and were also afraid to lose control over programming code. An interesting argument was the knowledge gap – respondents said they do not have enough information to choose or access cloud computing solutions. The pros part was dominated by time and scalability arguments. There were also some reasons based on clouds' accessibility, and comfort connected with cloud computing solutions from the programmers' point of view.

The next two figures show in what way decisions are taken in the case of task-oriented problems in case of software development projects. Fig 6. presents differences in decision-making in the case of cloud computing chosen as a computing environment. Respondents were asked to choose typical ways of making decisions while they were working as traditional (ftf) and virtual teams (v).

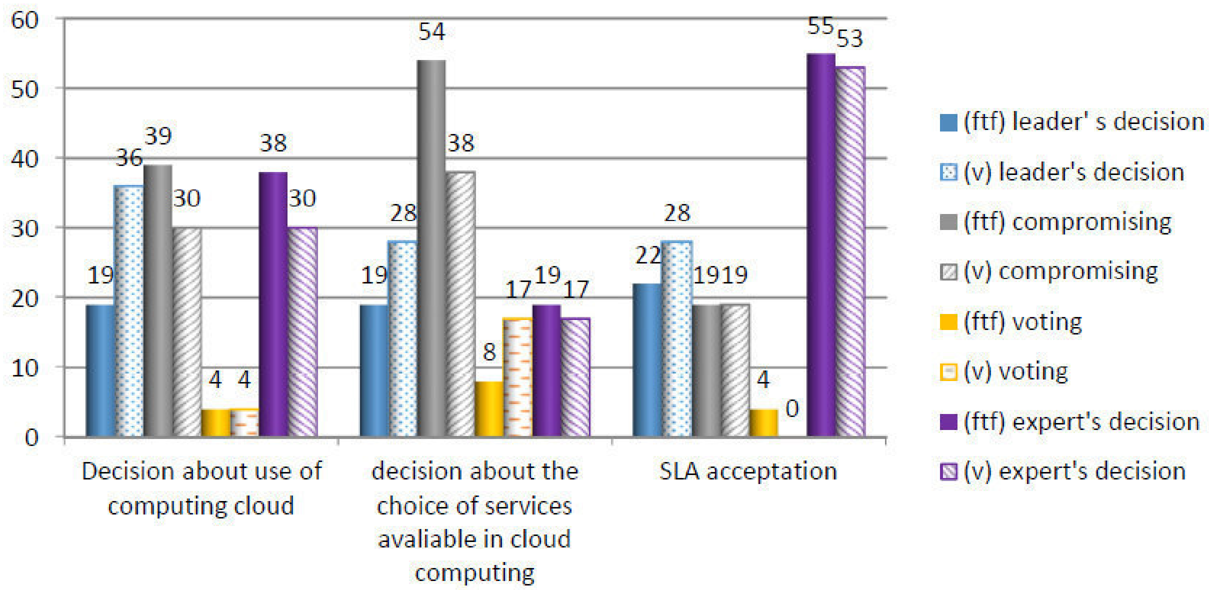


Figure 6: The way of decision-making in the case of choosing cloud computing as the development platform (% of respondents)

There were three types of decision assessed: about use of a cloud, services choice and SLA acceptation. Respondents could point to one, of four, ways of decision-making. The most important results are:

- In the case of the decision referring to cloud computing usage, and services choice, the most popular in ftf teams was compromising; in virtual teams, the leader was usually responsible for the decision of cloud computing usage and members were rather empowered to make the decision about the services choice,
- SLA acceptation was typically decided by the expert in traditional and virtual teams,
- Expert's decision is also required in some cases of answering the question as to if the cloud computing should be used,
- Leader's decisions were much more popular in the virtual team case for each kind of decision,
- The least popular team decision method was voting, but it was a bit more often in virtual teams in the case of services choice.

Fig 7. shows differences in decision-making in the case of cloud computing chosen as development platforms. There are also results presented for both: face to face teams, and virtual teams.

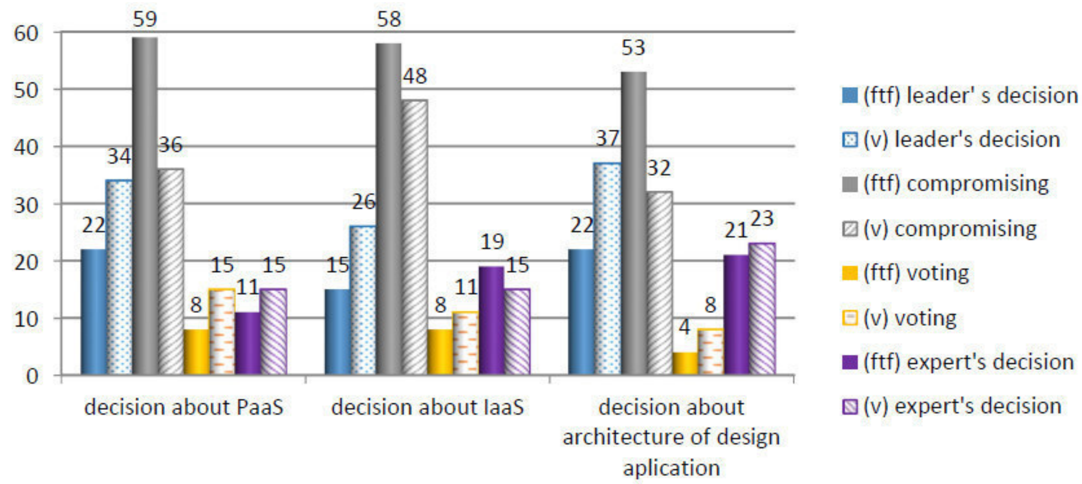


Figure 7: The way of decision - making in the case of the choice of cloud computing as a development platform (% of respondents).

It is visible that the most important way of making decisions about cloud computing as the development platform, in traditional teams, was compromising. The leader's opinions were much more powerful in virtual teams, especially in the case of both the choice of PaaS platforms, and architecture of designed applications. Decisions about choice of IaaS by virtual teams as application running environments were more often taken as compromising, but the role of leader seems to be stronger in the case of virtual teams than in ftf teams. Moreover, voting was the least preferable way in each kind of decision making, but it is possible to notice that this method is more often pointed to in virtual collaboration cases.

5. FINAL REMARKS

It is well-known fact, that decision - making is a very important operation in strategic management. There are plenty of factors that have big influence on such process. In the paper some of them, related to modern computing environment (i.e. cloud computing) were discussed. Moreover, we considered two kinds of software teams: traditional and virtual ones to analyse the most popular ways of decision making in each of them. The achieved results show as follows:

- The virtual environment, which is said to be a great chance for self-managed teams of experts, stimulates teams to rely on leaders' opinions more than on the discussion and compromise way. It can be the result of virtual teams' low trust and low cohesion (especially in Polish conditions). In consequence the different ways of building trust and commitment should be implemented in these kinds of teams, to promote team discussions and information sharing,
- The voting is fortunately perceived as the worst way of decision-making. It is important in the case of the development of different services supporting virtual team decisions. Disadvantages of this way of making decisions such as: possible team division, lack of satisfaction and commitment in the losers' group, should be better understood and eliminated,
- The influence of team virtualization on its results obtained in cloud computing environments need to be deeply analyzed. It should be confirmed that software teams make better decisions, and make their collaboration more effective thanks to cloud computing.

The obtained results are initial knowledge, and further research is necessary. Other aspects of software teams characteristic as well as the roles of leaders or available technologies can be taken into account. IT specialists' attitudes towards cloud computing were also widely analyzed, and based on the present research cloud computing seems to be widely accepted in many research groups. One of the reasons for negative attitudes can be the knowledge gap, which is astonishing in a group of IT specialists. That suggests the need for cloud computing pros and cons promotion in such groups, because the lists generated by respondents do not cover the ones given by cloud computing experts. It emphasizes that there is a need also for improving education in the area of cloud computing.

REFERENCES

[1] Bourgault M., Dronin N., Homel E. (2008). Decision-making with distributed project teams: anexploration of formalization and autonomy as determinants. *Project Management Journal* vol. 39 No 3, pp. 97–110.

[2] Cohen B. (2013). PaaS: New Opportunities for cloud Application Development. *IEEE computer*, vol. 46, No. 9, pp. 97–100.

[3] Cordes S. (2016). Virtual team learning: The role of collaboration process and technology affordance in team decision making. Knowledge Management & E-Learning, Vol. 8, No. 4, pp. 602–627.

[4] Davidekova M., Hvorecky J. (2017). *ICT Collaboration Tools for Virtual Teams in Terms of the SECI Model,* iJEP, Vol. 7, No. 1, pp. 95–116.

[5] Falkowski G., Troutmau S. (2005). Remote Control: A Practitioner's Guide to Managing Virtual Teams. *An IHRIM Press Book, USA.*

[6] Jeffrey B. Schmidt (2001). New Product Development Decision-Making Effectiveness: Comparing Individuals, Face-To-Face Teams, and Virtual Teams. Decision Sciences Vol. 32 No. 4, 575 - 600.
[7] McNamara K., Dennis A.R., Cart T.A. (2008). It's the Thought that Counts: The Mediating Effects of Information Processing in Virtual Team Decision Making. Information Systems Management Vol. 25 , Iss. 1, pp. 20–32.
[8] *Phillips J. (2004).* IT Project Management. On Track from start to Finish 2e. *The McGraw-Hill Companies.*
[9] Pollice P., Augustine L., Lowe Ch., Madhur J. (2004). Software Development for Small Teams: A RUP-Centric Approach, The Addison-Wesley Object Technology Series.
[10] Shalloway A., Beaver G., Trott J.R. (2009). *Lean- Agile Software Development.* Achieving enterprise Agility. Addison – Weseley Professional.
[11] Turban E., Liang T., Wu S.P.J, (2010). A Framework for Adopting Collaboration 2.0 Tools for Virtual Group Decision Making, *Group Decision Making*, No. 20, pp. 137–154.
[12] Vasan R. (2011). *A Venture Perspective on Cloud Computing,* IEEE Computer, vol. 44, No. 3, pp. 60–62.
[13] Wei K., Heckmann R., Crowston K., Li Q. (2011). Decision-Making Processes and Team Performance in Self-Organizing Virtual Teams: The Case of Free/Libre Open Source Software Development Teams. Retrieved from: https://crowston.syr.edu/sites/crowston.syr.edu/files/decision%20making%20paper_final.pdf
[14] Zofi Y. (2012). A Manager's Guide to Virtual Teams. Amacom, USA.

Strategic Allocation of Customers to Personnel in Large Scale Periodic Routing Problems

Tomasz ŚLIWINSKI[1]

[1]*Warsaw University of Technology, Institute of Control and Computation Engineering Nowowiejska 15/19, Warsaw, Poland, e-mail: tsliwins@elka.pw.edu.pl*

Abstract: Periodic routing and scheduling is of utmost importance in many industries with mobile personnel working in the field: sales representatives, service technicians, suppliers, etc. The resulting optimization problems are of large scale and complexity, mostly due to discrete, combinatoric nature of the systems and due to complicated, nonuniform constraints. In many cases the long-term stability of the customer to personnel allocation is required, leading to the decomposition of the major problem to single employee subproblems. The paper deals with allocation of customers to personnel, taking into account diverse system requirements and constraints, possible traveling schedules and expected operational costs. The difficulty of the problem lies in its large scale and constraints complexity as well as in troublesome objective evaluation for the given allocation. The general solution concept is presented and its usefulness is supported by the results of computational experiments.

Keywords: mobile personnel management, personnel allocation, vehicle routing and scheduling, time dependent, time windows, periodic

1. INTRODUCTION

In many industries periodic visits to the customers are the major part of the business. Field workers or sales representatives are committed to visiting customers on the regular basis in specific time intervals. Their work schedule must conform to many requirements and regulations, like labor law, time windows of the customers, customer-dependent time intervals, etc. It should also take into account various goals, with travel and labor costs being the most important ones.

In many cases personal relations are an important business factor. It is then allowed or even required that the same salesperson visits given customer for a long time. This leads to the decomposition of the main problem to many subproblems, one for each salesperson. The difficulty of such an approach lies in nondeterministic nature of the future schedules, as it is not possible to predict exact requirements, time schedules and routes in the future. Another stumbling block is the size of the allocation problem - dozens of salespersons and hundreds or even thousands of customers.

In the following paper simulation based approach to allocation of customers to salespersons is presented. It is based on the construction heuristics for fast building of the hypothetical future schedules combined with the allocation schema. The paper is organized as follows. First, the detailed problem for the single salesperson is presented. Then, the main subject of customers to salespersons allocation is considered and the optimization procedure is demonstrated. Finally, the results of computational experiments are presented.

2. PERIODIC ROUTING AND SCHEDULING FOR SINGLE SALESPERSON

The problem consists in finding the route and the schedule of visits for a given set of customers. Let $i \in I$ denote a set of visits to be scheduled over the given scheduling horizon consisting of days $d \in D$. Each visit takes place in the customer's location (geographical point). The visits at the given point should be periodic, which can be modeled by specifying day ranges when the visit has to take place or with some penalty for deviation from the planned date. Customers can have defined the time windows for each day of the planning horizon. Due to the business requirements, only uncapacitated problem was considered.

The total travel time depends on the selected route (sequence of visits) and the duration of each visit. The travel times and costs between points can be asymmetric and time-dependent, which means they depend on the direction and on the time of the day. It is assumed here, their changes are given on the hourly basis. The above problem can be seen as the Time Dependent Asymmetric Periodic Traveling Salesman Problem with Multiple Time Windows.

Due to the business requirements, it was also necessary to include basic labor law regulations. Different measures of work time were achieved and various regulations were fulfilled with additional parameters being rather simple extensions of the classical TSP problem: the number of the basic working hours (without implying over hours) for each *d* and the unit cost, the maximum total working hours for each day *d* (hard constraint) and the over hours unit cost. The above ones, could also be set on a weekly basis.

The formal mixed integer linear programming model for the above problem can be found in Ogryczak (2016).

The Asymmetric TSP with Time Windows have been studied intensively in the literature (Gendreau, 1998; Ascheuer, 2000; Ascheuer, 2001; Focacci, 2002) and respective optimization algorithms have been proposed. Extending it by adding time dependent travel times and costs (ATDTSPTW) makes the problem much more difficult. Most approaches focus on using different heuristic procedures (Hurkala, 2015). The exact solution to ATDTSPTW proposed in (Albiach, 2008) is based on the graph transformation into an ATSP for which standard optimization procedure can be used. Unfortunately, adding periodicity to the above problem prevents us from applying this approach. The reason is the size of the problem (for example, 14 day of the scheduling horizon with moderate 20 visits a day is similar to TDTSP with 280 nodes, depending on the periodicity settings and specific settings on time intervals between consecutive visits at the same point (penalty for deviation from given intervals). On the other hand, PTSP with Multiple Time Windows problem is recognized in the literature. For example in Tricoire (2010) authors utilize heuristic approach with VNS meta-heuristic to drive the changes of the current solution. Unfortunately, authors do not take into account data and rules, that are important from the business perspective, that is time dependent travel times, advanced interval definitions and various labor law regulations and impeding costs.

For that reason, a multi stage heuristic approach was proposed. It consists of 4 stages: construction heuristic to find initial solution, Variable Neighborhood Search (VNS) algorithm to improve global solution for the whole horizon, different algorithms to improve solutions separately for each day (VNS, simulated annealing, threshold accepting) and, finally, specialized, greedy heuristics to improve final solution (Hurkala, 2015; Ogryczak, 2016).

2.1 Finding optimal schedule for the given route

In classical TSP problems, given the route (sequence of visits) it is relatively easy to compute the travel cost and time. Introduction of time dependent travel parameters and time windows requires for a given route specialized approach to compute the detailed schedule and the resulting total travel cost and time.

The problem was addressed in the literature (Savelsbergh, 1992; Jong, 1996). An exact algorithm of finding the minimum route duration for time dependent travel with multiple time windows was presented in Hurkala (2015). After the preprocessing phase which eliminates some of the time windows, the algorithm iteratively reviews schedules and chooses ones with the shortest duration. The procedure actively eliminates the dominated solutions, substantially reducing the number of the computations.

Approximate approaches are also possible. The simple yet very efficient one, probes different travel starting times. It is assumed, the total travel duration and cost depend only on the time of the travel start, and the resulting schedule is computed by the simulation of the travel over the successive visits (points) on the route. This simple procedure was chosen as the main algorithm of finding optimal schedule in the allocation procedure. Our experiments showed it performs exceptionally well with the real-world data reaching performance two orders of magnitude better than achieved by the exact approach (Hurkala, 2015).

3. LONG TERM ALLOCATION OF CUSTOMERS TO PERSONNEL

In a standard mTSP (Multiple Traveling Salesman Problem) and VRP (Vehicle Routing Problem) the allocation of customers to personnel is made on a daily basis leading to high variability of actual field workers visiting each customer. This may be highly undesirable in some industries, where it easy to choose from many competitors and personal relations help to bind the customer to the company for longer period. It has to be noted, that even if poorly optimized allocation can be later improved, with each replaced salesperson, the business relation with the client must be renewed. That is why the long term allocation in the problem under consideration and in other VRP related problems often is an important requirement for the real world applications.

The main stumbling block in the long term allocation is the problem size and uncertainty of the future schedules. That is why the proposed procedure utilizes relatively simple construction heuristic in a long term system simulation.

The approach consists of generating random visits based on their required properties (visit frequency/interval settings, duration), ordering of visits and, finally, construction based scheduling and subsequent change of the initial allocation.

The algorithm's objective is to minimize the long term expected system operation cost, including time dependent travel costs and labor costs.

3.1 Input data

It is crucial that as many as possible different input data and parameters are considered, so that the resulting simulated schedules are a good approximate of the real ones. Below is the list of all indices and input data utilized by the allocation algorithm.

Indices:

$d \in \{1,\ldots,\lvert D\rvert\}$	days of the scheduling horizon
$w \in W$	weeks in the scheduling horizon
$p \in P$	customers
$i \in I$	visits
I_p	set of visits to the customer p
$p(i)$	customer visited in visit i
Θ_i^d	set of time windows for customer $p(i)$ on day d
Λ_{ij}^d	set of traffic time zones for path $(p(i), p(j))$ on day d

Parameters:

b_q^d	starting visit (depot) for the salesperson *q* on day *d*
f_q^d	final visit (depot) for the salesperson *q* on day *d*
$r_i = [d_i^{\min}, d_i^{\max}]$	possible day range for visit *i*
$\tau_i^{dk} = [a_i^{dk}, b_i^{dk}]$	*k* -th time window for point *p*(*i*) on day d; $k \in \Theta_i^d$
$\sigma_{ij}^{dn} = [c_{ij}^{dn}, d_{ij}^{dn}]$	*n*-th traffic time zone for path (*p*(*i*), *p*(*j*)) on day *d*
s_p	duration of each visit for customer *p*
t_{ij}^{dn}	travel time on path (*p*(*i*), *p*(*j*)) on day *d* in traffic time zone *n*
c_{ij}^{dn}	travel cost on path (*p*(*i*), *p*(*j*)) on day *d* in traffic time zone *n*
T_b^d, T_m^d	basic and maximum working hours on day *d*
W_b^d, W_m^d	unit cost for basic working hours and for overtime hours on day *d*
T_b^w, T_m^w	basic and maximum working hours in week *w*
W_0^w	unit cost for overtime hours in week *w*
a_p	required interval between subsequent visits to customer *p*
$P_p^+(\delta)$	piecewise linear non decreasing penalty function of excess δ over the interval a_p
$P_p^-(\delta)$	piecewise linear non decreasing penalty function of shortage δ to the interval a_p
w_p	customer importance/weight
Q_p^0	initial set of salespersons allowed to visit customer *p* – initial allocation

As can be seen, all the parameters can have different values for each day of the scheduling horizon. In particular one can introduce different labor costs depending on the week day and/or season. Similar applies to customers' time windows - they can be set differently for each day of the week.

Unfortunately, such detailed parameter setting may lead to very high memory footprint of the application, for example, when the travel times and costs are considered. That is why in the experiments single travel cost and times matrix was utilized for all days of the scheduling horizon. The matrix was computed as the average of the data for all week days. In future implementations it is possible, however, to utilize the information on the traffic changes over all days of the scheduling horizon.

Two ways of setting time intervals between consecutive visits to the same point are possible – one with the day range r_i when the visit should take place and the other with the penalty functions ($P_p^+(\delta)$ and $P_p^-(\delta)$) for the deviation from the specified interval. Both ways can be applied simultaneously. This approach was required by the business partner and allows very flexible setting of the intervals (note, the day ranges of the consecutive visits to the same point can overlap). One should note that the day ranges r_i are input parameters of the single salesperson problem, and in the allocation problem are generated based on a_p parameters (see next section).

3.2 Generating random visits

Important step of the algorithm is the generation of possible future visits to all customers. The visits are generated based on the input data, the a_p parameter in particular. For each customer p a number of visits has to be generated, and the day range r_i for each visit **i** should comply with the a_p parameter. Let d_b^0 be a random number in the range [0, a_p -l] generated with the uniform distribution, once for each customer. Then, day ranges for all subsequent visits to the same customer (numbered with k = (0,..., k^{max})) are computed $[d_p^0 + ka_p - a_p / 2, d_p^0 + ka_p + a_p / 2]$. Certainly the range size relative to a_p can be set individually for each customer, but for our computational tests the constant value was chosen. All other visit parameters, like visit duration s_p or penalty functions $P_p^+(\delta)$ and $P_p^-(\delta)$, are taken from the input parameters of the allocation algorithm.

The randomly generated values d_p^0 and the resulting day ranges r_i may affect created schedules and the final allocation. That is why the allocation algorithm should be repeated with different values of d_p^0.

3.3 Visits ordering

The construction heuristics applied for similar scheduling problems require initial ordering of operations/visits. We used the following properties to enforce the order of the visits:

1. value of d_i^{max} (latest possible visit day), earliest first,
2. penalty for possible delay on day d_i^{max}, highest first,
3. customer importance w_p, highest first.

The properties are enumerated according to their significance (with 1. being the most important one). The visits are ordered lexicographically, which means that only if the more significant properties are equal, the less significant property is considered.

3.4 Simplified scheduling

As the scheduling problem can be of enormous size (tens of salespersons, thousands of customers), fast and effective optimization procedure is required. It has to be sensitive to all problem properties, with geographical location of customers and frequency of visits being the most important ones.

The applied greedy construction heuristic tries to allocate each visit from the ordered list in the best possible place of the schedule created separately for each salesperson. In detail, let Q_p be the set of salespersons allowed to visit customer p Taking the first visit i from the ordered list, the algorithm examines all salespersons from the set $Q_{p(i)}$ and tries to insert the visit in the partially created schedules of each salesperson. Only days from the day range r_i are considered.

The visit is inserted temporarily in all possible places without changing the sequence of the previously scheduled visits. Each time the optimal schedule for the given route is computed.

Finally, the visit is inserted in a place (salesperson and the point its partial schedule) where the total cost increase is the smallest (Figure 1).

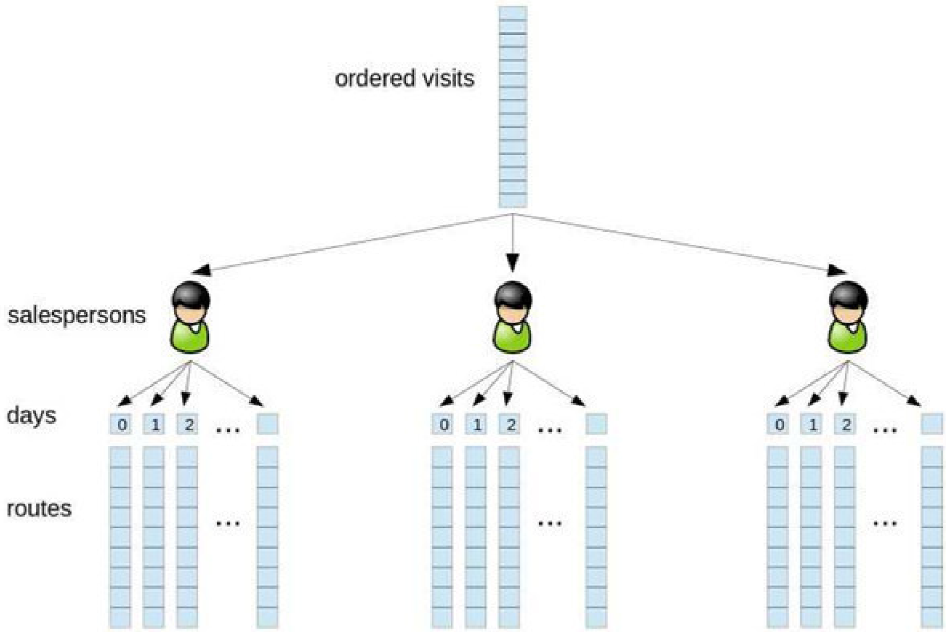


Figure 1: Illustration of the simplified scheduling.

One should keep in mind that the route of the salesperson q constructed on day d should begin with designated visit b_q^d and end with f_q^d, thus different starting and final depots on each day are possible.

3.5 The allocation algorithm

The results of the simplified scheduling algorithm can be very unpredictable, in particular when different seed for random number generation is used for visits creation. One can anticipate, however, the statistical properties of the schedules created in many runs will depend on the properties of the long term allocation, when customer location and visiting frequencies are considered. Based on this, the following algorithm of the long term allocation is proposed.

1. Initialization, input data preparation including initial allocation (set Q_p^0, k:= 0).
2. Generation of random visits and simplified scheduling for long time horizon. Only person $q \in Q_p^k$ can be scheduled for a visit to customer p.
3. The resulting schedule is analyzed to determine a person in the set Q_p^k with the smallest number of visits to customer p. If $|Q_p^k| > 1$, the person is removed from the set Q_p^k.
4. If $|Q_p^k| = 1; \forall p = 1$, then STOP, otherwise $k := k + 1$, go to 2.

In each iteration the algorithm restricts the sets of salespersons allowed to visit the customers. The decision is based on the properties of the generated schedule (total number of the visits to the given customer). The random visits generation in the simplified scheduling is performed in each iteration with different random seed. The number of iterations equals $max\ (Q_p^0\ -1)$, as each time only one person is removed from each set.

3.6 Possible extensions/modifications

The algorithm is very flexible and allows application in various real world situations. For example, it can be easily modified to allocate customers to a new salesperson or rearrange the allocation according to some new requirements. The stopping criterion can also be extended to allow limited (but greater than 1) number of employees visiting single client. This should increase the total system performance, but at the cost a weaker customer relations.

The algorithm makes possible to include personal travel characteristics – time and costs (in the experiments uniform data were used for all salespersons). The data should be available after longer period of collecting real travel data.

One of the concerns with the algorithm is uneven load allocation between salespersons. It is partially limited with maximum daily and weekly working hours. If better load balancing is required, it can easily be incorporated into the simplified scheduling algorithm, extending cost criterion with some over-/uneven load costs.

4. Computational experiments

4.1 Data generation

Three main parameters control the generation of random data: number of salespersons, horizon length and geographical area (all of Warsaw or only Ochota district). Different combinations of those parameters were tested in the computational experiments.

Labor costs and regulations are set the same for all salespersons (maximum 10 hours of working time a day and 40 hours a week). The number of customers is calculated so that approximate estimated daily travel time should be shorter than 7 hours. This resulted in average 24 customers per salesperson per day for the district Ochota, and 15 customers per salesperson per day for the whole city. The remaining input data were generated as follows.

- Time windows for customers (generated uniformly in the time range 7–20 with 30% probability, 714 with 25% probability, 12–20 with 25% probability, double 7–11 and 14–20 with 20% probability).
- Customer geographical coordinates were selected from a publicly available data of small businesses to cover selected area (all of Warsaw or only Ochota district).
- Travel time and costs are based on the real travel data acquired from the commercial provider.
- Separations between visits for each customer (a number generated uniformly in the range 2–5 with 30% probability, 5–15 with 40% probability and 10–25 with 30% probability).

- Being late penalty: piecewise linear function with 2 segments defined by points (0, separation/4, separation) with gradient 40PLN/separation, for the first segment and 160pln/separation for the second segment.
- Being early penalty: piecewise linear function with 3 segments defined by points (–separation, – separation/2, separation/4, 0) with gradient 160PLN/separation, for the first segment, 40PLN/separation for the second segment and 0 for the third segment.
- Day ranges for each visit are generated as described in section 3.2.
- All visits have the same duration: 7.5 minutes.

4.2 Simplified allocation algorithm

We wanted to compare the proposed allocation algorithm with any clustering heuristic available for the VRP problems (Marshall, 1981; Bramel, 1995; Ryan, 1993; Reanaud, 1996). Unfortunately, most of them were designed for relatively small problem instances (up to 20 vehicles and 200 customers), and in the problems considered here 5000 customers must be allocated to 20 salespersons. That is why a simplified approach had to be constructed. The proposed algorithm creates for each salesperson a single long route in which each customer appears only once. The input data are substantially simplified: no time dependent travel time/costs, no time windows, no visits intervals (although, the visits frequency is used for the customer ordering), no labor costs and regulations, assuming single depot. The algorithm consists of the following steps.

1. The preprocessing phase – the customers are ordered according to their visit frequencies (first the ones visited more often). For each salesperson a single initial customer is selected using standard MacQueen initialization method for the k-means algorithm (MacQueen, 1967) and initial route from the depot to the single customer is created.
2. The algorithm takes the first customer from the ordered list and tries to insert it in the partially created route of each salesperson. The route and the insertion position are chosen with the least cost/time increase. No detailed schedule has to be found for the given route.
3. The algorithm stops when all customers are allocated.

The simplified algorithm is similar to the proposed allocation schema, as also here the route construction is used to assign the customer to the salesperson. The major difference is that no visits frequency is used and the allocation is based on once created route.

4.3 The results

The test results cover different numbers of salespersons and different lengths of the scheduling horizon. All the tests were performed on a computer with 2.4GHz i7 series Intel processor. The values are the average of 10 tests. In Table 1. computing times are presented. As one can see the algorithm performs acceptable, although the exponential complexity in a function of the salesperson number can be seen. The horizon's length affects linearly the solution times.

To evaluate the allocation algorithm one needs a quality measure reflecting its desired properties. In the considered situation, the best quality measure seemed to be the total cost of the exact schedules computed separately for each salesperson and his allocated customers, over the whole scheduling horizon. The proposed allocation algorithm (its quality measure) is compared to the fully random allocation, and to the simplified allocation described above. The relative improvements are presented in Tables 2 and 3.

Table 1 : Computing times [s]

	Horizon length							
Area	All Warsaw				Ochota district			
Salespersons	30	60	120	180	30	60	120	180
5	0.3	0.5	0.8	2.1	3.8	5.2	9.0	15.3
10	3.2	6.8	13.2	12.3	25.3	38.3	85.3	134.5
20	40	58	130	252	192.3	385.3	815.4	1252.3

Table 2: Relative improvement over fully random allocation [%]

	Horizon length							
Area	All Warsaw				Ochota district			
Salespersons	30	60	120	180	30	60	120	180
5	1.3	5.3	2.4	10.0	3.8	4.8	8.3	4.5
10	5.0	7.3	11.9	16.6	9.3	13.0	19.1	20.0
20	7.7	14.8	23.0	25.5	10.3	17.8	23.3	24.0

As can be seen, the procedure performs best, showing an advantage of using all the information available. Also the length of the scheduling horizon affects the quality of the allocation, most probably due to the better (more data) estimation of the expected quality measure of the final schedule.

5. CONCLUSION

The presented long term customer to salesperson allocation for the periodic routing problems offers simple yet effective way of optimizing expected overall system costs. The algorithm combines simple construction heuristics for the schedule simulation with the iterative allocation procedure to effectively compute final allocation. The computational experiments show, the algorithm can be applied to large scale problems with various real world constraints and requirements.

6. ACKNOWLEDGEMENT

This research was financed by the European Union through the European Regional Development Fund under the Operational Programme “Innovative Economy” for the years 2007-2013; Priority 1 Research and development of modern technologies under the project POIG.01.03.01-14-076/12: “Decision Support System for Large-Scale Periodic Vehicle Routing and Scheduling Problems with Complex Constraints.”

7. REFERENCES

[1] Albiach, J., Sanchis, J., and Soler, D.M. (2008). An asymmetric tsp with time windows and with time-dependent travel times and costs: An exact solution through a graph transformation. *European Journal* of Operational Research*,* 189, pp. 789–802.

[2] Ascheuer, N., Fischetti, M. and Groetschel, M. (2000). A polyhedral study of the asymmetric traveling salesman problem with time windows. *Networks***,** 36, pp. 69–79.

[3] Ascheuer, N., Fischetti, M., and Groetschel, M (2001). Solving the asymmetric travelling salesman problem with time windows by branch-and-cut. *Mathematical Programming Series A***,** 90, pp. 475–506.

[4] Bramel, J. and Simchi-Levi, D. (1995). A location based heuristic for general routing problems. *Operations Research*, 43(4), pp. 649–660.

[5] Fisher, M.L., and Jaikumar, R. (1981). A generalized assignment heuristic for vehicle routing. *Networks***,** 11(2), pp. 109–124.

[6] Focacci, F., Lodi, A. and Milano, M. (2002). A hybrid exact algorithm for the tsptw. *INFORMS Journal on Computing***,** 14(4), pp. 403–417.

[7] Gendreau, M., Hertz, A., Laporte, G. and Stan, M. (1998). A generalized insertion heuristic for the traveling salesman problem with time windows. *Operations Research***,** 46(3), pp. 330–335.

[8] Hurkala, J. (2015). Minimum route duration algorithm for traveling salesman. *Vehicle Routing and Logistics Optimization***,** June 8–10, Vienna, Austria.

[9] Hurkala, J. (2015) Time-dependent traveling salesman problem with multiple time windows. *Annals of Computer Science and Information Systems***,** 6, pp. 71–78.

[10] Jong, C., Kant, G. and van Vliet, A. (1996). On finding minimal route duration in the vehicle routing problem with multiple time windows. *Tech. Rep.*

[11] MacQueen. J.B. (1967). Some methods for classication and analysis of multi-variate observation. In Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability*,* volume 1, pp. 281–297. University of California Press.

[12] Ogryczak, W., Sliwinski, T., Hurkala, J. Kaleta. M., Kozlowski, B. and Palka, P. (2016). Planning and management of mobile personnel tasks with time-dependent routing problems. In *Proceedings of the BOS 2016 conference* [in press].

[13] Reanaud, J., Boctor, F.F., and Laporte, G. (1996). An improved petal heuristic for the vehicle routing problem. *Journal of Operational Research Society*, 47, pp. 329–336.

[14] Ryan, M., Hjorring, D.C. and Glover, F. (1993). Extensions of the petal method for vehicle routing. *Journal of Operational Research Society*, 44, pp. 289–296.

[15] Savelsbergh, M.W.P. (1992). The vehicle routing problem with time windows: Minimizing route duration. *ORSA Journal on Computing***,** 4(2): pp. 146–154.

[16] Tricoire, F., Romauch, M., Doerner, K.F. and Hartl, R.F.(2010). Heuristics for the multi-period orienteering problem with multiple time windows. Computers & Operations Research**,** 37, pp. 351–367.

No-Wait & No-Idle Open Shop Minimum Makespan Scheduling with Bioperational Jobs

Michał MAŁAFIEJSKI[1], Krzysztof OCETKIEWICZ[2], Krzysztof PASTUSZAK[3]

[1]*Gdansk University of Technology, Faculty of Electronics, Telecommunication and Informatics, G.Narutowicza 11/12, Gdańsk, Poland, e-mail: michal@animima.org*

[2]*Gdansk University of Technology, Faculty of Electronics, Telecommunication and Informatics, G.Narutowicza 11/12, Gdańsk, Poland, e-mail: krzysztof.ocetkiewicz@eti.pg.gda.pl*

[3]*Gdansk University of Technology, Faculty of Electronics, Telecommunication and Informatics, G.Narutowicza 11/12, Gdańsk, Poland, e-mail: krzysztof_pastuszak@yahoo.com*

Abstract: In the open shop scheduling with bioperational jobs each job consists of two unit operations with a delay between the end of the first operation and the beginning of the second one. No-wait requirement enforces that the delay between operations is equal to 0. No-idle means that there is no idle time on any machine. We model this problem by the interval incidentor (1, 1)-coloring (IIR(1, 1)-coloring) of a graph with the minimum number of colors which was introduced and researched extensively by Pyatkin and Vizing. An incidentor is a pair (v, e), where v is a vertex and e is an edge incident to v. In the incidentor coloring of a graph the colors of incidentors at the same vertex must differ. The interval incidentor (1, 1)-coloring is a restriction of the incidentor coloring by two additional requirements:colors at any vertex form an interval of integers and the colors of incidentors of the same edge differ exactly by one. In the paper we proposed the polynomial time algorithm solving the problem of IIR(1, 1)-coloring for graphs with degree bounded by 4, i.e., we solved the problem of minimum makespan open shop scheduling of bioperational jobs with no-wait & no-idle requirements with the restriction that each machine handles at most 4 job.

Keywords: graph theory, interval coloring, consecutive coloring, one-sided interval, incidentor, shop scheduling

1. OPEN SHOP SCHEDULING WITH BIOPERATIONAL JOBS

The open shop scheduling problem is given by a set of machines and jobs and their restrictions. Let $M = \{M_1, M_2... M_m\}$ be a set of machines, and $J = \{J_1, J_2... J_n\}$ be a set of jobs (also called tasks). Each task consists of distinct operations $J_j = \{O_{1j}, O_{2j}, ...\}$, each of them being assigned to a distinct machine. Each machine may process at most one operation at any given time. At most one operation of each task may be processed at any given time. Operations within a task may be processed in any order. Each operation has a certain processing time assigned, denoted by p_{ij} for O_{ij} operation of task J_j, executed by machine M_i. The problem of open shop scheduling is one of the classic scheduling theory problems, first introduced by Gonzalez and Sahni in 1976 [Gonzalez 1976]. In the three-field notation, first introduced in 1982 by Graham et al.[Graham 1982], open shop is denoted by O in the α field.

Additional constraints are often imposed on the open shop scheduling problems [Giaro 2003], i.e., restricted delays between execution of subsequent operations within a job, availability of resources, and restrictions on the space of considered instances, i.e., limited number of machines [Giaro 2003] or UET only operations. In general, open shop scheduling is NP-hard, even when restricted to UET only operations [Giaro 2003]. Introducing restrictions on the space of the instances may allow us to construct polynomial time exact algorithms for certain subclasses of the open shop scheduling problem. In the open shop with bioperational tasks each job consists of exactly two operations. An example could be a scenario with read and write mode operations, which cannot be executed concurrently, e.g., in databases. We will denote this constraint on the number of operations within a job J_j by $op_j = 2$ in the β field. Assuming UET only operations, we will denote restriction on the maximum load per machine, i.e., the number of operations executed by machine, by load $\leq k$, for load no greater than k. An instance of open shop problem with bioperational tasks may be modeled by a graph, where machines are represented by vertices and bioperational jobs by edges.

Common constraints on feasible solution space include no-wait and no-idle restrictions [Giaro 2003]. No-wait requirement enforces that the delay between operations within a job is equal to 0. No-idle means that there is no idle time on any machine, once the machine started working. We will denote no-wait and no-idle restriction by NWI in the β field in three-field notation.

Open shop with bioperational tasks and UET operations, and with no-wait and no-idle restrictions is considered in this paper. The problem may be solved by computing an incidentor coloring of a graph that models the instance of a problem.

2. INCIDENTOR COLORING

The notion of incidentors was first introduced by A.A. Zykov in 60s [Pyatkin 1997, Pyatkin 2002]. The incidentor coloring model however remained unknown until the 90s, when A.V. Pyatkin and V.G. Vizing focused their research on the model and its applications in the network transmission and scheduling theory [Pyatkin 2002, Pyatkin 2006, Vizing 2009, Vizing 2012, Vizing 2007, Vizing 2014, Pyatkin 2015]. The incidentor coloring model arises from scheduling data transmission in the communication networks [Pyatkin 1997, Pyatkin 2002, Pyatkin 1999] and was first [Pyatkin 2002] introduced in 1997 paper [Pyatkin 1997] by Pyatkin and later described in greater detail in his Ph.D. Dissertation [Pyatkin 1999].

Let G = (V, E) be a graph. Incidentor is a pair (v, e), where e ∊ E(G) is an edge incident to v ∊ V(G). Each edge $e = \{v_i, v_j\}$, where v_i, v_j ∊ V(G), may be represented by a pair of incidentors: (v_i, e) and (v_j, e). Two incidentors (v_i, e) and (v_j , f) are called mated if e = f, where e, f ∊ E(G), and are called adjacent if $v_i = v_j$ [Vizing 2000].

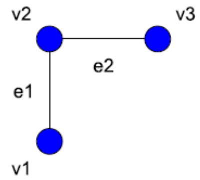


Figure 1: There are four incidentors in the graph: (v_1, e_1), (v_2, e_1), (v_2, e_2), (v_3, e_2). Incidentors (v_1, e_1) and (v_2, e_1) are mated. Incidentors (v_2, e_1) and (v_2, e_2) are adjacent.

The problem of incidentor coloring of graph is to assign the color to each incidentor, such that for each pair of adjacent incidentors their assigned colors differ. Additional constraints on the colors of mated incidentors may be introduced [Pyatkin 2013, Pyatkin 2004].

These constraints provide utility, that may be used to model the restrictions on time delays between the subsequently executed operations within a job. For any edge e ∊ E(G) we may specify lower and upper bounds on the difference between the numbers of colors used. In general, any edge may have different bounds assigned, however most of the research focused on the graphs with the same constraints applied to each edge. The coloring c may be called IR(k, l)-coloring, if for k ≤ l, k, l ∊ Z, for any pair, (u, e) and (v, e), where v, u ∊ V (G), e = {u, v} ∊ E(G), the following holds: k ≤ | c(u,e) – c(v,e) | ≤ l.

It may be also called IR(k, l)-coloring. The minimum number of colors sufficient for IR(k, l)-coloring of a graph G may be denoted by $\chi^{IR}_{(k,\,l)}(G)$. In particular, IR(0, 0)-coloring is equivalent to the edge coloring. In the IR(1, 1)-coloring the colors of mated incidentors form intervals of integers, hence it may be used to solve the no-wait open shop problem – O | UET, $op_j = 2$, NW | C_{max}. The IR(1, 1)-coloring always exists. If the degree of a graph G is even (2|Δ(G)), then Δ(G) colors are always sufficient [Pyatkin 2004, Hanson 1998]. If the degree of a graph G is odd, then the problem of deciding, whether Δ(G) colors are sufficient, remains open [Pyatkin 2004]. For any graph, Δ(G) + 1 colors are sufficient to construct IR(1, 1)-coloring. For any k ≥Δ(G) − 1, $\chi^{IR}_{(k,k)}(G) = \chi^{IR}_{(k,\infty)}(G)$ [Vizing 2003, Małafiejska 2016].

For any k, $\chi^{IR}_{(k,\ \Delta(G)-1)}(G) = \chi^{IR}_{(k,\infty)}(G) = \max\{\Delta(G), \lceil\frac{\Delta(G)}{2}\rceil + k\}$ [Vizing 2003]. If k ≥ Δ(G) 2, $\chi^{IR}_{(k,k)}(G) = \lceil\frac{\Delta(G)}{2}\rceil + k$ [Vizing 2005]. For a graph G, Δ (G) colors are sufficient to find both IR(0, 1) and IR(0, ∞) colorings of G [Pyatkin 2004, Melynikov 2000]. In interval incidentor coloring, denoted later by IIR coloring, colors of adjacent incidentors must form interval of integers.

Chromatic scheduling is one of the approaches used in solving open shop scheduling problems [Giaro 2003]. The general idea behind the chromatic scheduling is first to create a graph model of the instance of a problem, then choose the adequate graph coloring model, construct a feasible coloring and construct the schedule from the coloring. Open shop with bioperational tasks may be solved by constructing a graph model, later called incidentor scheduling graph, and then computing incidentor coloring of the graph. Machines are modeled by vertices in the graph. Tasks are modeled by edges between vertices corresponding to the machines, on which constituent operations are executed. Operations are modeled by incidentors. Colors of incidentors correspond to time windows, in which relevant operations are executed. With an additional no-idle restriction, this problem may be modeled by interval incidentor (IIR) coloring. No-wait restriction is modeled by IR(1, 1)-coloring. No-wait and no-idle restriction is modeled by the IIR(1, 1)-coloring. Restrictions of the instance space of open shop scheduling problem, that limit the maximum degree of scheduling graph, may allow construction of polynomial time algorithms, even if the problem is NP-hard in general.

3. INTERVAL INCIDENTOR COLORING OF GRAPHS WITH Δ BOUNDED BY 4

In this section the linear time algorithm for construction of IIR(1, 1)-coloring of graphs with Δ bounded by 4, using 4 colors, is presented. This algorithm may be used to solve the no-wait and no-idle open shop scheduling problem with bioperational jobs and at most 4 operations per machine. The problem can be described in the three-field notation by: $O \mid UET, op_j = 2, load \leq 4, NWI \mid C_{max}$.

Let $G = (V, E)$, with $\Delta(G) \leq 4$, be an incidentor scheduling graph. Consider an edge $e \in E(G)$ from vertex u to vertex v, where $u, v \in V(G)$. The following partial orientation of graph G:

- every vertex of degree 4 is adjacent to exactly two incoming arcs,
- every vertex of degree 3 is adjacent to exactly one incoming arc,
- every vertex of degree 3 or 4 is adjacent to exactly two outgoing arcs or undirected edges,

we call a *legal* partial orientation of graph G. An example of a legal partial orientation is shown in Figure 2.

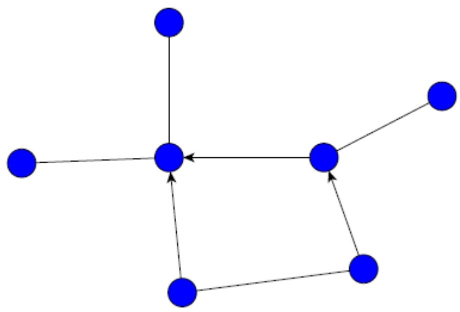


Figure 2: Example of partial orientation D.

Lemma 1. It is possible to construct a legal partial orientation D of the incidentor scheduling graph G, $\Delta(G) \leq 4$, in polynomial time.

Proof. We first create a partial orientation of a graph with some vertices split, then we join split vertices again and obtain a partial orientation D of original graph G. Only vertices of degree 3 or 4 will be split and only when one of the following conditions occurs:

- vertex of degree 3 is adjacent to exactly one incoming arc;
- vertex of degree 4 is adjacent to exactly two incoming arc.

No vertices of degree 3 in G may have more than one adjacent incoming arc, and no vertices of degree 4 in G may have more than two adjacent incoming arcs. We call vertices with these configurations legal. Only legal vertices may be split. Legal configurations of arcs and edges are shown in Fig. 3.

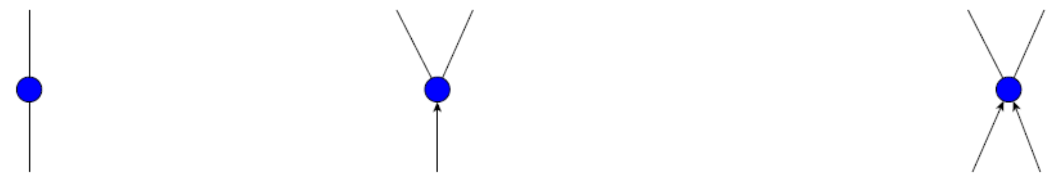


Figure 3: Straight lines represent edges or outgoing arcs. The ”legal” combinations of the arcs and edges: left – vertices of degree 2, middle – vertices of degree 3, right – vertices of degree 4.

In the splitting process, vertex v is split into two vertices: one incident only to the incoming arcs (one, if in G, d(v) = 3, and two, if in G, d(v) = 4), and one incident only to two outgoing arcs or undirected edges. The splitting process is shown in Fig. 4.

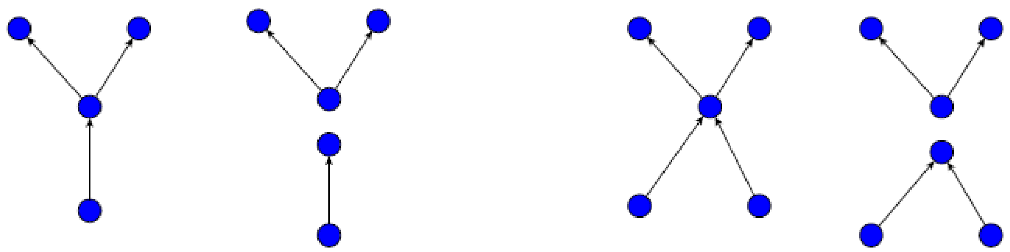


Figure 4: The splitting of vertices of degrees 3 (left) and 4 (right).

The following algorithm may be used to construct the partial orientation D:

In a loop:

1. Detect an undirected cycle C in a subgraph induced by vertices of degree 3 and 4 in the remaining graph.
2. Transform the cycle into a directed cycle.
3. Split vertices of degrees 3 and 4 in the remaining graph, if their configuration is already legal.

Once no more cycles can be found, in a loop:

4. Detect a path of maximal length in the remaining graph, so that a vertex of degree 1 in the remaining graph is the end of the path.
5. Transform the path into a directed path, starting in a vertex of degree 1.
6. Split vertices of degrees 3 and 4 in the remaining graph, if their configuration is already legal.

An example of contraction is shown in Fig. 5.

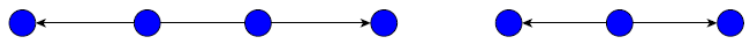


Figure 5: Contraction and reconstruction of the edges in undirected paths: left – original path, right – contracted path.

The algorithm returns a decomposition of G into paths and cycles. After joining split vertices again, we obtain a partial orientation D of original graph G.

Since Δ(G) is bounded by 4, it can be approximated by a constant. Both cycles and paths can be found in linear time, hence construction of D can be done in O(|E|) time.

Each vertex of degree 4 in G is an inner vertex in two paths or cycles. Each vertex of degree 3 is the end of one path, and the inner vertex in one path or cycle. As a result of the construction scheme, there are only the following structures in the decomposition :

- directed cycles, with two incoming arcs and two outgoing arcs, alternately,
- directed paths, with two incoming arcs and two outgoing arcs, alternately,
- undirected cycles,
- undirected paths.

Structures occurring in a decomposition are shown in Fig. 6. Boths ends of each path are vertices of degree 1 in the returned decomposition, of degree either 1, or 3 in G.

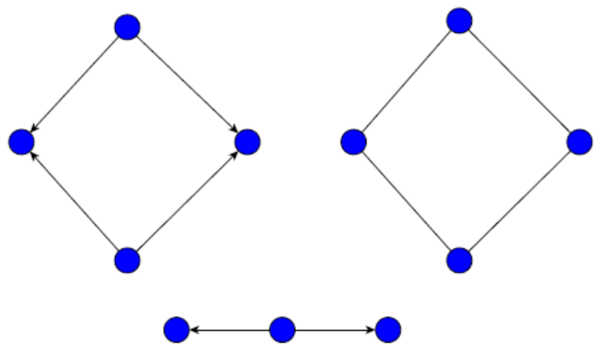


Figure 6: The structures obtained after the decomposition of the graph: left – directed cycle, right – undirected cycles, bottom – paths.

Theorem 1. For any graph G with Δ(G) ≤ 4 there is IR(1, 1)-coloring using at most 4 colors. For any graph G with Δ(G) ≤ 4 there is IIR(1, 1)-coloring using at most 4 colors.

Proof. According to Lemma 1, there always exists partial orientation of G. Incidentors of the incoming arcs may have only colors 1 or 4 assigned. Incidentors of undirected edges or outgoing arcs may assume only colors 2 or 3. Coloring of partial orientation D using 4 colors can be constructed using the following scheme:

1. Construct partial orientation of G. Do not join split vertices yet. Contract undirected paths into single vertices. An example of contraction is shown in Fig. 5.
2. Color undirected cycles using colors 2 and 3, alternately.
3. Color directed paths using colors 2, 1, 4, 3 (or 3, 4, 1, 2), repeatedly.
4. Restore contracted paths and color them, then join split vertices, thus obtaining colored partial orientation D of original graph G.

Transform arcs back into edges, while preserving the coloring from D, thus obtaining IR(1,1) coloring of G.

Lemma 2. IR(1, 1) coloring of G constructed by algorithm from Theorem 1, is also a feasible IIR(1, 1) coloring of G.

Proof. Let us remind, that in the decomposition only the structures shown in Fig. 6 occur, and that each vertex of degree 3 or 4 in G is split into two vertices, one with incoming arcs only, and one with outgoing arc or undirected edges only. Thus, for each pair of split vertices, after joining them again, a vertex with one or two incoming arcs and exactly two outgoing arcs or undirected edges forms. Since incidentors of incoming arcs are always colored with 1 or 4, and incidentors of undirected edges and outgoing arcs are always colored with 2 or 3 (and there are always two of them) colors of adjacent incidentors always form interval: $\{1, 2, 3\}$ or $\{2, 3, 4\}$ if the degree of the vertex was equal to 3, and $\{1, 2, 3, 4\}$ if it was equal to 4. The original structure of G is restored after joining again each pair of split vertices and its incidentor coloring is a feasible IIR(1, 1) 4-coloring of G

Hence the theorem follows:

Theorem 2. Schedule of makespan 4 for the problem O|UET,NWI, $op_j = 2$, load $\leq 4|C_{max}$ always exists and it can be obtained in linear time.

4. SUMMARY

Presented linear time algorithm may be used for construction of schedule of makespan 4 for the problem O|UET,NWI, $op_j = 2$, load $\leq 4|C_{max}$. The problem of no-wait & no-idle open shop scheduling with bioperational jobs with at most k operations per machine remains open for $k \geq 5$. We conjecture, that for $k = 5$ the problem is polynomial, and for $k \geq 7$ is NP-hard. Another open problem is O|UET,NWI, $op_j \leq 3$, load $\leq 4|C_{max}$. Giaro proved both the sufficient and necessary conditions for the existence of schedule with makespan 4, however the problem of existence of schedule with makespan $C_{max} \in \{5,6\}$ remains open.

5. REFERENCES

[1] Giaro, K. (2003). *Task scheduling by graph coloring (in Polish).* D.Sc. Dissertation. Gdansk University of Technology.

[2] Giaro, K. (1997). *The complexity of consecutive D-coloring of bipartite graphs: 4 is easy, 5 is hard.* Ars Combinatoria 47, pp. 287-300.

[3] Gonzalez, T., Sahni, S. (1976). *Open shop scheduling to minimize finish time*. Journal of the ACM, pp. 665–679.

[4] Graham, R., Lawler, E., Lenstra, J., Rinnooy Kan, A. (1979). *Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey.* Proceedings of the Advanced Research Institute on Discrete Optimization and Systems Applications of the Systems Science Panel of NATO and of the Discrete Optimization Symposium., pp. 287–326.

[5] Hanson, D., Loten, C. O. M., Toft, B. (1998). *On interval coloring of biregular bipartite graphs.* Ars ombinatoria, pp. 23–32.

[6] Małafiejska, A. (2016). *Wybrane problemy i modele końcówkowego kolorowania grafów (in Polish).* Raport Tech. WETI 3/2016, Politechnika Gdańska (2016), pp. 1–19.

[7] Melynikov, L., Pyatkin, A., Vizing, V. (2000). *On the (k, l)-coloring of incidentors (in Russian).* Diskretn. Anal. Issled. Oper. 1, pp. 29–37.

[8] Pyatkin, A. V. (1997a). *Some optimization problems of scheduling the transmission of messages in a local communication network.* Operations Research and Discrete Analysis, pp. 227–232, 1997.

[9] Pyatkin, A. (1997b). *The incidentor coloring problems and their applications, Ph.D. Dissertation (in Russian).* Institute of Mathematics SB RAS, Novosibirsk.

[10] Pyatkin, A. V. (2002). *The incidentor coloring of multigraphs and its applications.* Electronic Notes in Discrete Mathematics, pp. 103–104.

[11] Pyatkin, A. (2004). U*pper and lower bounds for an incidentor (k,l)-chromatic number (in Russian).* Diskretn. Anal. Issled. Oper., 1, pp. 93–102.

[12] Pyatkin, A., Vizing, V. (2006). *Incidentor coloring of weighted multigraphs.* Discrete Applied Mathematics, 120, pp. 209–217.

[13] Pyatkin, A. (2013). *Incidentor coloring: Methods and results.* Graph Theory and Interactions, Durham.

[14] Pyatkin, A. (2015). *On an Interval (1,1)-Coloring of Incidentors of Interval Colorable Graphs.* Journal of Applied and Industrial Mathematics, 9, pp. 271–278.

[15] Vizing, V. (2000). *On Incidentor Coloring in a Partially Directed Multigraph.* Fuzzy Sets and Systems, pp. 141–147.

[16] Vizing, V. (2003). *Interval colorings of the incidentors of an undirected multigraph (in Russian).* Diskretn. Anal. Issled. Oper. pp. 14–40.

[17] Vizing, V. (2005). *On the (p, q)-coloring of incidentors of an undirected multigraph (in Russian).* Diskretn. Anal. Issled. Oper. 1, pp. 23–39.

[18] Vizing, V. (2007). *On Bounds for the Incidentor Chromatic Number of a Directed Weighted Multigraph.* Journal of Applied and Industrial Mathematics, 1, pp. 504–508.

[19] Vizing, V. (2009). *On Incidentor Coloring in a Hypergraph.* Journal of Applied and Industrial Mathematics, 3, pp. 144–147.

[20] Vizing, V. (2012). *Multicoloring the Incidentors of a Weighted Undirected Multigraph.* Journal of Applied and Industrial Mathematics, 6, pp. 514–521.

[21] Vizing, V. (2014). *Multicoloring the Incidentors of a Weighted Directed Multigraph.* Journal of Applied and Industrial Mathematics, 8, pp. 604–608.

The Concept of Industry 4.0 Related Manufacturing Technology Maturity Model (Manutech Maturity Model,Mtmm)

Jaroslaw GRACEL[1], Piotr LEBKOWSKI[2]

[1]*AGH University of Science and Technology, Faculty of Management Mickiewicza 30, Krakow, Poland, e-mail: jaroslaw.gracel@gmail.com*

[2]*AGH University of Science and Technology, Faculty of Management Mickiewicza 30, Krakow, Poland, e-mail: plebkows@zarz.agh.edu.pl*

Abstract: The main objective of this article is to describe Industry 4.0 and key manufacturing-technology-related technological and business challenges for manufacturing companies. The groups especially interested in the industry 4.0 implementation are operations, technical and production directors responsible for operational excellence of manufacturing plants, strategic development and business continuity. Based on latest Industryand manufacturing technology market research, factories located in Poland are less technologically advanced than its counterparts in West European plants; special attention should be paid to understanding the differences and related risks. Accordingly, development of the model for assessing the current level of maturity for manufacturing technology related to Industry 4.0 initiative becomes a relevant research task. In the article, key Industry 4.0 related technological areas will be described, e.g. cloud computing, cobots (robots collaborating with humans), industrial internet of things, additive manufacturing, big data, systems integrations, mobile robotics and cybersecurity. Based on the extensive research into international references and industrial consulting experiences in industrial business consulting conducted in Polish manufacturing companies, the manufacturing technology ManuTech Maturity Model (MTMM) concept related to Industry will be developed and presented. A substantial and innovative part of the article will be devoted to the adjustment ofproposed maturity model to specific features of the Polish industrial and manufacturing sector. This will be relevant due to noticeable *differences in the level of technological advancement between Western and Eastern Europe sectors.*

Keywords: Industry 4.0, manutech maturity model, mtmm, smart factory, manufacturing technology, technology maturity, maturity model.

1. INTRODUCTION

In the last three decades manufacturing technologies, such as production automation system, industrial robotics, manufacturing execution system (MES), CAx systems, and enterprise resource planning systems (ERP) have been a relevant part of manufacturing businesses, but their role has been only supportive. In the recently launched Fourth Industrial Revolution (known as Industry 4.0 in the manufacturing sector), its role will be elevated to the strategic level. Based on the research conducted by PwC (Geissbauer, Vedso, & Schrauf, 2016), it is estimated that the share of investments in Industry 4.0 solutions will account for more than 50% of planned capital investments for the next five years. The German industry will thus invest a total of EUR 40 billion in Industry 4.0 every year until 2020. Applying the same investment level to the European industrial sector, the annual investments will be as high as EUR 140 billion. While countries like Germany with mature economies can afford certain level of investment in Industry 4.0 initiatives, economies of Central and Easter Europe (like Poland's) will be exposed to significantly greater risks. The key risk (Cyfrowa Polska, 2016) is connected with the digitization gap between Western European countries and Poland. Measured with Digitization Index (Cyfrowa Polska, 2016), the gap for the economy as a whole is 34%, but for the "advanced manufacturing" sector it is 45% and the "simple manufacturing" sector 78% (!). Therefore, formal Industry

Manufacturing Technology Maturity Model should be developed. The role of the model is to help top management answer the critical questions, such as: "What is the current level of technological advancement of the factory?" or "How the manufacturing technologies should be deployed to ensure effective execution of a new Industry 4.0 strategy or new business models?" This paper is structured as follows. Section 1 contains Industry 4.0 definitions, its impact on manufacturing technology and business/technical challenges related. Section 2 introduces definition of maturity models and presents comparison of the existing Industry 4.0 and manufacturing technology related maturity models. Section 3 contains the concept of the new ManuTech Maturity Model (or MTMM) with conclusions.

1.1 Industry 4.0 – definitions

The first three industrial revolutions came about as a result of mechanization, electricity and IT. Now, the introduction of the Internet of Things and Services into the manufacturing environment is ushering in the fourth industrial revolution. In the future, businesses will establish global networks that incorporate their machinery, warehousing systems and production facilities in the shape of Cyber-Physical Systems (CPS) (Acatech, 2013). Industry 4.0 emphasizes the idea of consistent digitization and linking of all productive units in an economy (Gilchrist, 2014). The Industrie 4.0 working group (Kagermann, 2013) developed recommendations to focus research efforts on three strategic topics:

- horizontal integration through value networks,
- end-to-end digital integration of engineering across the entire value chain,
- vertical integration and networked manufacturing systems.

The definition of Industry 4.0 given by (Hermann, Pentek, & Otto, 2015) is as follows: Industrie 4.0 is a collective term for technologies and concepts of value chain organization. Within the modular structured Smart Factories of Industrie 4.0, CPS monitor physical processes, create a virtual copy of the physical world and make decentralized decisions. Over the IoT, CPS communicate and cooperate with one another and humans in real time. Via the IoS, both internal and cross-organizational services are offered and used by participants of the value chain. While at the global level there are various initiatives around the future of manufacturing (e.g. Industrial Internet Consortium, Factory of The Future, Made in China), the Industry 4.0 is the most consistent and as it was developed in Germany, has the biggest influence on the European industrial market. The most repeatable Industrie 4.0 components (Hermann et al., 2015) are as follow: Cyber-Physical Systems (CPS), Internet of Things (IoT), Internet of Services (IoS) and Smart Factory.

1.2 Industry 4.0 – impact on manufacturing technology

The in-depth research into the literature has revealed many different views on key Industry 4.0 technologies, but the overall conclusion is that impact of the Industry 4.0 initiative on manufacturing technology landscape is significant. Accordingly, with the growing differentiation of customer requirements, increasing individualization of products, the level of a manufacturing technology complexity will constantly be growing. The BCG (RuBmann et al., 2015) introduces nine pillars of Industry 4.0 technological advancement: big data and analytics, autonomous robots, simulation, horizontal and vertical integration, the industrial Internet of Things, cybersecurity, the cloud, additive manufacturing, augmented reality. Many of the nine advances in technology that form the foundation for Industry 4.0 are already used in manufacturing, but with Industry 4.0, they will transform production: isolated, optimized cells will come together as a fully integrated, automated, and optimized production flow, leading to greater efficiencies and changing traditional production relationships among suppliers, producers, and customers—as well as between a human and machine. The new technological characteristics of the industrial landscape will be as follows (Gilchrist, 2014): cyber-physical systems and marketplace, smart robots and machines, big data, new quality of connectivity, energy efficiency and decentralization, virtual industrialization, factory 4.0 (fully connected way of making things with key manufacturing technologies such as: intelligent sensors, 3D printing/additive manufacturing, advanced materials, advanced manufacturing systems (CPS, full interconnected automation), robots, autonomous vehicles, cloud computing and big data). The complement view (Sujeet Chand; Jim Davis et al., 2016) distinguishes a number of disruptive technologies that will enable digitization of the manufacturing sector: data, computational power, and connectivity – big data/open data, Internet of Things/M2M, cloud technology. analytics and intelligence – digitization and automation of knowledge work, advanced analytics, human- machine interaction – touch interfaces and next level of GUIs, virtual and augmented reality, digital-to-physical conversion – additive manufacturing, advanced robotics (e.g. human-robot collaboration), energy storage and harvesting.

The research into the advanced manufacturing technology trends (de Weck & Reed, 2014) identified seven categories: nano-engineering of materials and surfaces, additive and precision manufacturing, robotics and adaptive automation, next generation electronics, continuous manufacturing of pharmaceuticals and biomanufacturing, design and management of distributed supply chains, green sustainable manufacturing. The core manufacturing technologies relevant to the manufacturing companies have been identified and should be taken into consideration in the maturity model development process.

Industrie 4.0 and its core technologies such as secure plug & work-solutions for reconfigurations of machines, augmented reality-based assistance devices for workers, cyber-physical systems with inexpensive sensors to automatically collect data in value streams, machines and components, as well as machine learning and big data algorithms can be seen for e.g. car makers as enablers of transition towards flexible automation or scalable model-mix-factories (Wee, Kelly, Cattel, & Breunig, 2015).

1.3 Industry 4.0 business and technological challenges

Industry 4.0 brings numerous different challenges, e.g. business, strategic, technological, people-related ones. Based on international research (Geissbauer et al., 2016), companies expect Industry 4.0 to favorably impact their revenue, costs and efficiency. The additional revenue will come from: digitizing products and services within the existing portfolio, New digital products, services and solutions, offering big data and analytics as a service, personalized products and mass customization, capturing high-margin business through improved customer insight from data analytics, increasing market shares of core products. The lower cost and greater efficiency are expected to be generated by the following factors: real-time inline quality control based on big data analytics; modular, flexible and customer-tailored production concepts; real-time control of process and product variance; augmented reality and optimization by data analytics; predictive maintenance on key assets using predictive algorithms to optimize repair and maintenance schedules and improve asset uptime; vertical integration from sensors through MES to real-time production planning for better machine utilization and faster throughput times; horizontal integration, as well as track-and-trace of products for better inventory performance and reduced logistics; digitization and automation of processes for a smarter use of human resources and higher operations speed; system-based, real-time end-to-end planning and horizontal collaboration using cloud-based planning platforms for execution optimization, increased scale from increased market share of core products.

In parallel to business improvement opportunities, certain business and technological challenges are connected. The greatest challenges connected with implementing Industry 4.0 will be (Kagermann, 2013) standardization, process/work organization, product (means technology) availability, new business models, security know-how protection (e.g. cybersecurity), lack of specialist staff, research, training and CPD (continuing professional development). The research (Bauer et al., 2016) shows that six of ten manufacturing companies face significant management barriers when working on Industry 4.0 implementation. The main barriers are connected with the level of progress in Industry 4.0 implementation process and are as follows:

- Top five general barriers: difficulty in coordinating actions, lack of courage to push through radical transformation, lack of necessary talent, concerns about cybersecurity while integrating IT-OT systems, lack of a clear business case that justifies investments in the underlying IT architecture,
- More advanced manufacturers' barriers: concerns about data ownership, uncertainty about in- vs. outsourcing and lack of knowledge about providers, challenges of integrating data from disparate sources in order to enable Industry 4.0 applications.

While there are a lot of Industry 4.0 business challenges, the area of technological challenges should be addressed accordingly, because mass customization puts great strains on product developers and system designers (Fasth-Berglund & Stahre, 2013). A look at the development of computer science (CS), information and communication technologies (ICT), and manufacturing science and technology (MSC) reveals their parallel development (Monostori et al., 2016). Therefore, convergence of the related manufacturing technologies dedicated to virtual world and physical world is being expected. Should it come, the ability to build future-proof architecture for manufacturing technology in the factory (or factory network) will become the critical skill.

The key priority (technology and standardization related) areas are as follows (Kagermann, 2013): standardization and open standards for a reference architecture, managing complex systems (e.g. planning models and explanatory models), delivering a comprehensive broadband infrastructure for industry, safety and security (e.g. cybersecurity).

The other technological challenges are connected with the fact that Industry 4.0 related manufacturing technologies are at a different stage of maturity. Based on Gartner Hype Cycle for Emerging Technologies research (Gartner, 2015), most of Industry 4.0 related technologies are in the "peak of inflated expectations" and "through of disillusionment" phases (but the situation has been changing dynamically). Only a few, e.g. 3D printing and virtual reality, are in the "slope of enlightenment" phase. It was principally the reliability and technical working order of machines that enabled the second industrial revolution (2.0) and third industrial revolution (3.0) (i.e., the rise of manufacturing automation technology projects (Nowacki, 1953).

In addition, the business justification of the Industry 4.0 investment should be developed properly. Lessons learned from preparation to the third industrial revolution (production automation) show (Schulz, 1962) that every manufacturing technology development phase, as well as every technological innovation has technological and economic aspects. Therefore, it is important to formulate the following two questions:

1. Which manufacturing processes CAN be automated?
2. Which manufacturing processes SHOULD be automated?

1.4 Manufacturing technology maturity in the Polish manufacturing sector

Despite the fact that Industry 4.0 initiative brings many opportunities to the Polish industrial market, many threats can be identified. The main threat list includes a lack of skilled workforce and access to investment capital (Owerczuk, 2016). Moreover, the long-term historical competitive advantage of the Polish economy – low-cost workforce – is beginning to run out. Poland has significant digitization gap ("Cyfrowa Polska," 2016) in the manufacturing sector (45% in advanced manufacturing, 78% in simple manufacturing). Research into the level of automation in the Polish manufacturing sector (Hajkus & Gracel, 2015) showed that only 15% of factories are fully automated and 76% partially automated (which could mean that just one machine or even most of them are automated). Based on the industrial IT market research, almost 60% of manufacturing companies in Poland gather shop-floor data manually, while 36% are prepared for automated data collection (Hajkus & Gracel, 2015). An industrial robotics market research (IFR 2016) revealed that (in 2015) Poland had 28 robots per 10,000 employees. For the sake of comparison, Germany's robot density level is estimated at 292 per 10,000 employees. Moreover, only 32% of manufacturing and engineering companies have development programs for engineers in place (Gracel, Stoch, Bieganska & Rząca, 2017).

An overview of the Polish manufacturing sector shows significant challenges in numerous dimensions. Polish enterprises (especially SMEs) are not aware of the forthcoming technological changes and, moreover, they do not understand the risk of overtaking production orders by their international partners (customers) due to the latter's higher efficiency gained from Industry 4.0 investments (Goetz & Gracel, 2017). This facts should be taken into consideration while developing Manufacturing Technology Maturity Model related to Industry 4.0. The research shows that Polish managers still cope with **Industry 3.0** challenges.

1.5 Impact on decision making and implementation challenges

The impact of Industry 4.0 on the manufacturing sector is of strategic proportion. (April, 2013) identified various areas of potential, e.g. meeting individual customer requirements, flexibility, responding to demographic change in the workplace, resource productivity and efficiency, a high-wage economy that is still competitive, work-life-balance, creating value opportunities through new services and **optimized decisiontaking.** Industrie 4.0 (April, 2013) provides end-to-end transparency in real time, allowing early verification of design decisions in the sphere of engineering, as well as a more flexible response to disruption and global optimization across all of a company's sites in the sphere of production. The decision making process is supported from different perspectives, e.g., in terms of human resources or algorithms. In the automotive manufacturing (Peters, Chun, & Lanza, 2016): Industry 4.0 and autonomous driving (which could be called Mobility 4.0 in this analogy) use quite similar technologies such as various types of (optical) sensors, data fusion systems and decentralized **decision making algorithms**.

Successful implementation of an Industry 4.0 strategy requires the involvement of top and mid-level management. There are three fundamentally different sources of an individual manager's poor to negative approach to implementing a strategy (Guth & Macmillian, 1986): perceived inability to execute strategy, low perceived probability that strategy will work, perception that outcomes will not help to achieve individual goals. Thus, appropriate tools supporting common understanding of the level of maturity of, e.g., Industry 4.0 manufacturing technologies, etc. across all levels of organization, should be developed.

2. INDUSTRY 4.0 AND AN OVERVIEW OF MANUFACTURING TECHNOLOGY MATURITY MODELS

2.1 Maturity Models introduction

Maturity is defined (Paulk, Curtis, Chrissis, & Weber, 1993) as a specific process for explicitly defining, managing, measurement and control of the evolutionary growth of an entity. Maturity is related to the evolutionary progress in demonstrating a particular capacity or the pursuit of a certain goal, from an initial state to the desirable final state. (Kohlegger, Maier, & Thalmann, 2009) defined maturity models as tools used to evaluate the maturity capabilities of certain elements and select the appropriate actions to bring the elements to a higher level of maturity. A maturity model (Proen?a & Borbinha, 2016) is a technique that has been proved to be valuable in measuring different aspects of a process or an organization. It represents a path towards increasingly organized and systematic way of doing business in the manufacturing industry.

2.2 Maturity model assessment design methodologies

The review of literature on maturity models design and development methodologies has helped to identify some references to leading topics. The most general approach to developing maturity models for assessment of Business Processes Maturity Model (BPMM) and Knowledge Management Capability Assessment (KMCA), described by (De Bruin, Freeze, Kaulkarni, & Rosemann, 2005), proposes six model development phases: scope, design, populate, test, deploy, maintain. IT management related Maturity Model development procedure (Becker, Knackstedt, & PoppelbuB, 2009) comprises the following steps: problem definition, comparison of existing maturity models, determination of development strategy, iterative maturity model development, conception of transfer and evaluation. Third design approach developed by (Mettler, 2010) recommends the performance of the six following steps: identify need and specify problem domain, define scope of model application and use, identify operationalization measures, implement deployment and evaluation method, apply model, evaluate model structure and deployment method, synthesis of design and continuous learning.

For the purposes of the development of the ManuTech Maturity Model (MTMM), Becker’s methodology, as the best recognized, up-close to the manufacturing technology domain and practical, has been chosen.

2.3 Industry 4.0 and Manufacturing Technology related Maturity models overview

The Industry 4.0 context analysis performed in Section 1 have proved relevance of the research subject and its relevance for the manufacturing technology domain. To develop the concept of a new maturity model (based on the Becker’s procedure), second step of the process “Comparison of existing maturity models” should be performed. During the thorough review of the literature (in English), more than 2,000 references to maturity models have been identified, and over 30 have been taken into consideration for further analysis. The review of Industry 4.0 related maturity models revealed 10 key reference models (presented in the Table 1). For the purposes of ManuTech Maturity Model (ManuTech MM, MTMM) development, they have been evaluated. The most comprehensive approach to the subject has been embedded into the following three models: IMPULS - Industrie 4.0 Readiness (Karl Lichtblau & Stich, 2014), Industry 4.0 / Digital Operations Self-Assessment (pwc, 2016) and Industry 4.0 readiness and maturity model (Schumacher, Erol, & Sihn, 2016). The research into the "Industry 3.0" maturity models helped to identify models of an automation maturity, such as: 19 levels of manufacturing process automation (Nof, 2009) or Level of Automation (LoA) assessment methodology (Fasth-Berglund & Stahre, 2013).

Table 1: Listing of the maturity models related to Industry 4.0

	MODEL NAME	DELIVERED BY	DOMAIN	TARGET GROUP	YEAR
1	CPS Maturity Model (Hellinger, Seeger, & Tönskötter, 2012)	RWTH Aachen University	Cyber-Physical Systems	CTO, CIO	2012
2	The Digital Maturity Model 4.0 (Gill & Boskirk, 2016)	Forrester	Strategic Digitalization, Business Focus	CDO (Chief Digital Officer)	2016
3	The Connected Enterprise Maturity Model (Rockwellautomation, 2014)	Rockwell Automation	Industry (OT/IT Networks)	CIO, CTO	2014
4	Digital Compass maps Industry 4.0 (Wee et al., 2015)	McKinsey & Company	Industry 4.0, Digitalization	CEO, CIO	2015
5	IMPULS – Industrie 4.0 Readiness (Karl Lichtblau & Stich, 2014)	VDMA, RWTH Aachen	Industry 4.0 Strategic	CEO	2015
6	Industry 4.0 / Digital Operations Self-Assessment (pwc, 2016)	PwC	Industry 4.0, Strategic	CEO, CIO	2016
7	Industry 4.0 readiness and maturity model (Schumacher et al., 2016)	Fraunhofer Austria	Industry 4.0 manufacturing strategy	CEO, CIO, R&D Director	2016
8	The Digital Maturity Model 4.0 (Gill & Boskirk, 2016)	Forrester	Strategic Digitalization, Business Focus	CDO (Chief Digital Officer)	2016
9	Supply Chain Visibility Maturity (Gartner, 2017)	Gartner	Supply Chain	Supply Chain Director, CIO	2016
10	3DP (3D Printing) Maturity Model (Mueller & Karevska, 2016)	EY	Additive Manufacturing	CEO, CTO	2016

No model focused on assessment of manufacturing technology maturity level has been found during the review of the literature. After the review and (six) workshops with managers of Polish manufacturing companies, the gaps in current models have been identified discussed below.

2.4 Identified issues/gaps in current models

While Industry 4.0 is an innovative initiative, there are some areas still undefined, unprecise and uncertain. According to (Mettler, 2011) while building a maturity assessment model for a highly innovative phenomenon, justificatory knowledge to base upon is weak or missing and principles of form and function are unclear as no dominant design prevails. Furthermore, the cases necessary to derive the maturity levels and recommendations from may be missing as well.

During the gap analysis of existing maturity models, the following difficulties have been identified:

- While the literature review has identified Industry 4.0 related maturity models, the gap analysis shows areas in the models that should be extended, e.g., deep technology insights. Moreover, a maturity model dedicated to Manufacturing Technology has not been found during the review.
- Industry 4.0 Maturity Models are mainly focused on strategic level of companies. Operational and technology levels have been overlooked.
- Current models present superficial approach to manufacturing technologies, which are to be the core components of manufacturing companies aiming to develop manufacturing processes ready for mass customization.
- Current models do not take into consideration a current level of the "technologization" (e.g. level of production automation or robotics density) of the industrial market/companies. It creates a risk to the applicability in less advanced economies (like Poland's).
- Current models do not relate to the manufacturing technologies from Industry 3.0 that are still and will be the core part of manufacturing companies. To retrieve the comprehensive maturity assessment results, a combination of core technologies should be used. Industry x.0 approach should be developed.
- Current models focus on assessment of general set of Industry 4.0 related technologies without focusing on specificity of certain industry (e.g. automotive, electronics, food processing, others). In the new model, manufacturing technologies should be grouped into the set of general technologies (Industry 4.0 related) and specific technologies (based on industry best practices).
- Workshop carried out with manufacturing companies in Poland focused on assessing the Industry 4.0 technological readiness" resulted in demotivation of managers. The reason is people do not like to be said to be "totally not aligned" with Industry 4.0 from technological perspective.
- Current models are focused in the majority on top (C-level) managers, which is important to start the Industry 4.0 initiative, but the research showed that the middle management involvement is a crucial aspect in the implementation of the strategy. Accordingly, the new maturity model should pay a special attention to this fact.
- Current maturity models do not refer to the maturity of assessed technologies. This creates the risk of underestimation of CAPEX (Capital Expenditure), TCO (Total Cost of Ownership), applicability and reliability of the technologies. This should be taken into consideration during the new model development.

The assessment of the gaps in existing maturity models can contain some faults, because of a lack of accessibility to the detailed description of the models and intellectual property issues (since consulting companies and research institutes, which provide the service of maturity assessment to organizations, restrict access to their intellectual property (Willaert, Van den Bergh, Willems, & Deschoolmeester, 2007). However, the gap analysis performed forms a basis for the development of a new maturity model in manufacturing technology domain (ManuTech Maturity Model or MTMM).

3. THE CONCEPT OF MANUFACTURING TECHNOLOGY MATURITY MODEL (MANUTECH MATURITY MODEL, MTMM)

3.1 The model assumptions

The comprehensive gap analysis resulted in development of the concept of the new maturity model related to the Manufacturing Technology (with Industry 4.0 correlation). The model will be named “ManuTech Maturity Model” or, shortly, MTMM.

While conducting the model design step “Determination of development strategy” (based on Becker’s methodology), the main assumptions that should be taken into consideration (based on **Section 2.4 Gap analysis**) are as follows:

1. The model should focus on operational level of manufacturing technologies with respect to its strategic alignment.
2. The target group of the model will be middle management of manufacturing companies.
3. The model will be focused on maturity of manufacturing technologies with key supporting items (such as strategic alignment, cybersecurity, knowledge management, people & culture).
4. The mix of core manufacturing technologies will be selected based on the relevance for a specific industry, reference Industry 4.0 technologies and research in advanced manufacturing technology domain.
5. The model should refer to the Industry 4.0 design principles (Hermann et al., 2015): interoperability, virtualization, decentralization, real-time capability, service orientation, modularity.
6. The model will contain a reference to the maturity of technologies included (e.g. based on Technical Readiness Level, TRL or Gartner’s hype-cycle)
7. Special attention will be paid to practicality and applicability of the model.

Following the best practices in development of maturity models, the structure of the model should contain:

- The maturity levels (levels 1 to 5, with 1 - the lowest, and 5 - state-of-the art).
- The dimensions.
- The method of application.
- The method of representations.

3.2 ManuTech Maturity Model (MTMM) concept

Based on the workshops and interviews with experts and middle managers of manufacturing companies (in Poland) **conducted by the authors over the last five years**, the concept of the ManuTech Maturity Model (MTMM) structure has been developed. The concept includes the following **eight dimensions** with **maturity items** assigned to them (details are presented in Table 2): core technologies, people & culture, knowledge management, real-time integration, infrastructure, strategic awareness & alignment, process excellence, cybersecurity. These dimensions have been proposed as the most comprehensive conceptualization of maturity combining theoretical and practical perspective of the topic. The authors propose **two methods of application**: self-assessment or assessment by an external expert and **numerical method of representation** with visual radar charts.

Table 2: ManuTech Maturity Model (MTMM) dimensions and maturity items

	DIMENSION	MATURITY ITEMS (EXAMPLES)
1	**Core technologies**	**General:** Predictive analytics, Advanced automation of machines, Advanced scheduling, etc. **Industry domain specific:** Robotized palletizing, Additive manufacturing (e.g. 3D printing) for components, Process traceability, etc.
2	**People & culture**	Organizational culture openness to innovation, Technology competences of employees, Empowerment of employees, Employee satisfaction level, …
3	**Knowledge management**	Existence of technology knowledge base, Knowledge gathering, sharing systems, …
4	**Real-time integration**	Automated integration of data flow between core business processes (e.g. Production Management – Maintenance Management, R&D – Production Management, etc.) internally and externally (with networked factories, customers, suppliers)
5	**Infrastructure**	High-speed internet access at shop floor, Energy meters with Ethernet connectors, etc.
6	**Strategic awareness & alignment**	Leadership competences development level, Digitization strategy, Business case in place, etc.
7	**Process excellence**	WCM, Lean, SixSigma, TPM, Project Management standards, etc.
8	**Cybersecurity**	Security policy for IT/OT, Cybersecurity vulnerability report, Active protection equipment, etc.

	DIMENSION	MATURITY ITEMS (EXAMPLES)
1	**Core technologies**	**General:** Predictive analytics, Advanced automation of machines, Advanced scheduling, etc. **Industry domain specific:** Robotized palletizing, Additive manufacturing (e.g. 3D printing) for components, Process traceability, etc.
2	**People & culture**	Organizational culture openness to innovation, Technology competences of employees, Empowerment of employees, Employee satisfaction level, …
3	**Knowledge management**	Existence of technology knowledge base, Knowledge gathering, sharing systems, …
4	**Real-time integration**	Automated integration of data flow between core business processes (e.g. Production Management – Maintenance Management, R&D – Production Management, etc.) internally and externally (with networked factories, customers, suppliers)
5	**Infrastructure**	High-speed internet access at shop floor, Energy meters with Ethernet connectors, etc.
6	**Strategic awareness & alignment**	Leadership competences development level, Digitization strategy, Business case in place, etc.
7	**Process excellence**	WCM, Lean, SixSigma, TPM, Project Management standards, etc.
8	**Cybersecurity**	Security policy for IT/OT, Cybersecurity vulnerability report, Active protection equipment, etc.

After processing the Becker's "Iterative maturity model development" step, the model will be transformed into the assessment tool and tested with manufacturing companies in a specific industrial sector. Due to the fact, that numerous existing maturity assessment models have faced applicability problems, the practical verification of the concept will be a crucial step in the ManuTech Maturity Model development.

3.3 Conclusions

To summarize, the paper presents a comprehensive analysis on the subject of Industry 4.0 with definitions, impact on manufacturing technology and the related business/technological challenges. Moreover, it introduces a definition of maturity models and presents comparison of existing Industry 4.0 and manufacturing technology related maturity models. Finally, it presents the concept of the new manufacturing technology related **ManuTech Maturity Model (or MTMM)**.

The main conclusions are as follows:

1. The rationale behind the concept of the ManuTech Maturity Model (MTMM) is strong and based on both in-depth literature examination and direct interviews with the target group (middle managers of manufacturing companies) in Poland.
2. The concept of the models addresses the main gaps identified during comparisons of existing maturity models in Industry 4.0/Manufacturing technology domains.

3. The concept of the ManuTech Maturity Model (MTMM) creates solid foundation for building a practical and long-term tool for assessment of maturity in manufacturing technology domain.
4. The ManuTech Maturity Model (MTMM) should be evaluated and tested in a specific manufacturing company to prove its practicality and applicability.
5. The ManuTech Maturity Model (MTMM) is based on a generic term "Manufacturing Technology" and with this approach it can be applied to various types of industries and multiple technologies. Consequently, on the MTMM model, long lasting and Industry x.0 independent maturity assessment can be based.

4. REFERENCES

[1] April, W. G. (2013). 001.Recommendations for implementing the strategic. *Acatech,* (Acatech), pp. 4–7.
[2] Bauer, H., Baur, C., Mohr, D., Tschiesner, A., Weskamp, T., & Mathis, R. (2016). Industry 4.0 after the initial hype. *McKinsey Digital.*
[3] Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering,* 1(3), pp. 213–222.
[4] Chlebus, E. (2000) Techniki komputerowe CAx w inzynierii produkcji. Wydawnictwo Naukowo-Techniczne WNT.
[5] Cyfrowa Polska. (2016). *McKinsey & Company.*
[6] De Bruin, T., Freeze, R., Kaulkarni, U., & Rosemann, M. (2005). Understanding the Main Phases of Developing a Maturity Assessment Model. In *Australasian Conference on Information Systems (ACIS)* (pp. 8–19).
[7] de Weck, O. L., & Reed, D. (2014). Trends in Advanced Manufacturing Technology Innovation. *Production in the Innovation Economy*, 235.
[8] Fasth-Berglund, Á., & Stahre, J. (2013). Cognitive automation strategy for reconfigurable and sustainable assembly systems. *Assembly Automation, 33,* pp. 294–303.
[9] *Gartner (2015)* Hype Cycle for Emerging Technologies, *Gartner Inc.*
[10] Geissbauer, R., Vedso, J., & Schrauf, S. (2016). Industry 4.0 : Building the digital enterprise. *PwC,* pp. 1–36.
[11] Gilchrist, A. (2014). Industry 4.0. *Roland Berger,* (March).
[12] Gill, M., & Boskirk, S. Van. (2016). The Digital Maturity Model 4 .0. Forrester.
[13] Goetz, M., & Gracel, J. (2017). Przemys! czwartej generacji (Industry 4.0) – wyzwania dla badan w kontekscie międzynarodowym. *Kwar,* 1(51), pp. 217–235.
[14] Gracel, J., Stoch, M., Bieganska, A., & Rząca, A. (2017). *Inzynierowie Przemysiu 4.0. (Nie)gotowi do zmian? 2017* (Vol. 0). ASTOR.
[15] Guth, D. W., & Macmillian, C. . (1986). Strategy Implementation versus Middle Managements Self-interest. Strategic Management Journal, 7(4), pp. 313–327.
[16] Hajkus, J., & Gracel, J. (2015). W jakie technologie inwestują firmy produkcyjne w Polsce? *ASTOR Whitepaper.*
[17] Hellinger, A., Seeger, H., & Tönskötter, L. (2012). agenda CPS – Integrierte Forschungsagenda Cyber-Physical Systems. *Acatech STUDIE*, pp. 1–297.
[18] Hermann, M., Pentek, T., & Otto, B. (2015). Design Principles for Industrie 4.0. Working Paper A Literature Review, (1), 16.
[19] Kagermann, H. (2013). Securing the Future of German Manufaturing Industry. Recommendations for implementing the strategic Initiative INDUSTRIE 4.0. *Acatech, Final Report of the Industrie 4.0 Working Group,* (April), pp. 4–7.
[20] Karl Lichtblau, & Stich, V. (2014). IMPULS. Industrie 4.0-Readiness, pp. 0–76.

[21] Kohlegger, M., Maier, R., & Thalmann, S. (2009). Understanding maturity models results of a structured content analysis. *Proceedings of IKNOW '09 andISEMANTICS '09,* (September), pp. 51–61.

[22] Mettler, T. (2010). Supply Management im Krankenhaus – Konstruktion und Evaluation eines konfigurierbaren Reifegradmodells zur zielgerichteten Gestaltung. Hochschule für Wirtschafts-, Rechtsund Sozialwissenschaften (HSG).

[23] Mettler, T. (2011). Maturity assessment models: a design science research approach. *International Journal of Society Systems Science,* 3(1/2), 81.

[24] Monostori, L., Kadar, B., Bauemhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Ueda, K. (2016). Cyber-physical systems in manufacturing. *{CIRP} Annals – Manufacturing Technology*, 65(2), p. 621–641.

[25] Mueller, A., & Karevska, S. (2016). EY's Global 3D printing Report 2016 - Executive Summary. Ernst & Young GmbH. *Nof, S. (2009).* Handbook of Automation. Computation and Control. *Springer Verlag.*

[26] Nowacki, P. (1953) *Automatyzacjaprocesow wytworczych.* Przegl^d Elektrotechniczny, 1/1953

[27] Owerczuk, M. (2016). Przemysf 4.0. Szansa czy zagrozenie dla rozwoju innowacyjnej gospodarki ? *Boston Consulting Group.*

[28] Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). Capability Maturity Model for Software , Version 1 . 1. *Carnegie Melon University*, (February), 82.

[29] Peters, S., Chun, J.-H., & Lanza, G. (2016). Digitalization of automotive industry - scenarios for future manufacturing. *Manufacturing Review,* 3(1).

[30] Proen^a, D., & Borbinha, J. (2016). Maturity Models for Information Systems - A State of the Art. *Procedia Computer Science,* 100(2), pp. 1042–1049.

[31] Pwc (2016) The Industry 4.0 / Digital Operations Self-Assessment. Retrieved from: https://i40-self-assessment.pwc.de/i40/landing/

[32] Rockwellautomation. (2014). The Connected Enterprise Maturity Model, 12. Retrieved from http://www.rockwellautomation.com/rockwellautomation/innovation/connected-enterprise/maturity-model.page?

[33] Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0. The Future of Productivity and Growth in Manufacturing Industries. *Boston Consulting Group.*

[34] Schulz, Z. (1962) EfektywnoSC ekonomiczna automatyzacji produkcji przemysfowej. Polska Akademia Nauk.Panstwowe Wydawnictwo Naukowe.

[35] Schumacher, A., Erol, S., & Sihn, W. (2016). A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises. *Procedia CIRP, 52,* 161-166.

[36] Sujeet Chand; Jim Davis, Gilchrist, A., Innowacji, D., Digital, M., Wee, D., Kelly, R., ... Lopez, R. (2016). Industry 4.0. *McKinsey & Company,* 5(1), pp. 4–7.

[37] Wee, D., Kelly, R., Cattel, J., & Breunig, M. (2015). Industry 4.0 – how to navigate digitization of the manufacturing sector. *McKinsey & Company.*

[38] Willaert, P., Van den Bergh, J., Willems, J., & Deschoolmeester, D. (2007). The Process-Oriented Organization: Holistic View Developing a Framework for business Process orientation Maturity.Business Process Management, 64 *(September).*

A SPC Strategy for Decision Making in Manufacturing Processes

Alain GIL DEL VAL[1], Bartosz SAWIK[2], Alba María AGUSTÍN[3], Javier FAULÍN[3], Pedro María DIÉGUEZ[1]

[1]*Public University of Navarre, Mechanical, Energy and Materials Engineering Department Campus Arrosadía, 31006, Pamplona, Spain, e-mail: alain.gil@unavarra.es*

[2]*AGH University of Science and Technology, Faculty of Management Mickiewicza 30, Krakow, Poland, e-mail: BSawik@zarz.agh.edu.pl*

[3]*Public University of Navarre, Institute of Smart Cities, Integrated Group of Logistics and Transportation, Pamplona, Spain.*

Abstract: Tapping is an extensively employed manufacturing process by which a multi-teeth tool, known as tap, cuts a mating thread when driven into a hole. When taps are new or slightly worn the process is usually in control and the geometry of the resulting threads on the work piece is correct. But as the tap wear increases the thread geometry deviates progressively from the correct one and eventually the screw threads become unacceptable. The aim of this paper consists on a development of statistical process control strategy for decision making according to data coming from the current signal of the tap spindle for assessing thread quality. It could operate on line and indicates when the tap wear is so critical that, if the process were continued, it would result in unacceptable screw threads. The system would be very cost-effective since the tapping process could be run without any operator intervention.

Keywords: tapping, SPC, decision, quality, PCA.

1. INTRODUCTION

Tapping by cutting is one of the most common operations in manufacturing. It consists of cutting internal threads on the wall of a previously drilled hole by means of a tool called a tap that has cutting edges on its chamfered periphery. Tapping is a difficult operation due to the large number of cutting edges involved and the complicated synchronization necessary between the rotational and the feed movements of the tap, a task particularly difficult at high speeds.

In this paper, we detail the experimental a development of Statistical Process Control (SPC) strategy for decision making aimed at predicting the point in time when the tapped threads become unacceptable as a result of tap wear.

(Chen, 1990) worked on tap modeling, has focused on fault detection and classification for diagnosis purposes. (Liu, 1991) continued this work using a neural network to classify the faults. Both papers relied on intrusive sensors. Further work along this line was conducted by Li et al. (2002), whose system detects the same three types of faults of the previous papers but used less intrusive sensors. Still, the sensors utilized involve an additional cost, and for training purposes the system requires production of unacceptable threads for each of the types of faults the system is to detect.

Although (Li, 2002) used non-intrusive sensors and this reduces cost, we present a more cost effective solution in which the electrical current signal is acquired directly from the I/O module of the CNC machine itself, and hence, no sensors are required. In contrast with earlier research on tap modeling, our emphasis is on process monitoring of the tap tool and on the decision making of the corresponding thread quality, rather than on fault detection and classification of the tapping process with respect to a specific set of process conditions.

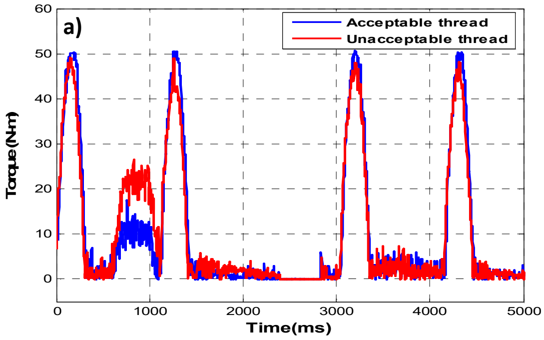
(Lorentz, 1989) was the first to use the principal component analysis in tapping processes. The results were good at improving the cutting and operational conditions.

The aim of this paper is to develop a SPC strategy for decision making when the thread quality is close to be out of tolerances. The paper is organized as follows; first, experimentation, signal and results are showed and demonstrated the evidence that justifies the use of tap torque for monitoring purposes. Second, Principal Component Analysis is used to reduce the number of parameters from monitor signal and, finally, a Generalized Variance (GV) control charts are illustrated the decision making to stop the process before mechanized an unacceptable thread.

2. SET UP, SIGNAL & EXPERIMENTAL RESULTS

Tapping operations are performed on a CNC machining centre using three flutes metric M10x1.5mm High Speed Steel (HSSE) taps Titanium Nitride (TiN) coated, with 4–4.5 threads on a 9º taper. Work piece material is nodular cast iron (GGG50) in 250 by 450mm plates 20 mm thick. A speed of 65m/min and no coolant are selected. To assess thread quality, all tapped threads are inspected by a “go-no-go” gauge.

The electrical current from the I/O module of the CNC machine was selected to make a decision about the thread quality. Therefore, Figure 1.a illustrates the torque sensitivity as acceptable and unacceptable thread profiles are superposed. The effect of wear on the tapped threads quality can be appreciate in Figure 1.b. As a consequence, the go gauge can not go in the threaded hole.



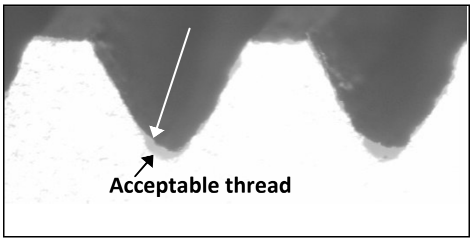


Figure 1: a) Spindle motor torque signals from acceptable and unacceptable threads and b) Lack of Quality: Non defective and defective thread profiles.

6 taps of the same characteristics has been tested and results are shown in Table 1. Second column indicates the number of tapped holes that passed the "go-no-go" gauge test. The process continued until an experienced operator declared that the tap end of life by catastrophic failure was close, and hence tap was changed. Therefore, third column indicates de total number of threads at which. All in all, the difference between columns indicates the number of unacceptable threads.

Table 1: Number of tapped holes corresponding to the 6 HSSE M10x1,5 taps tested.

Tap	Acceptable threads	Total threads
1	192	205
2	108	120
3	36	40
4	60	80
5	81	100
6	109	120

Figure 2 shows the area parameters from current signal for decision making to assess the thread quality.

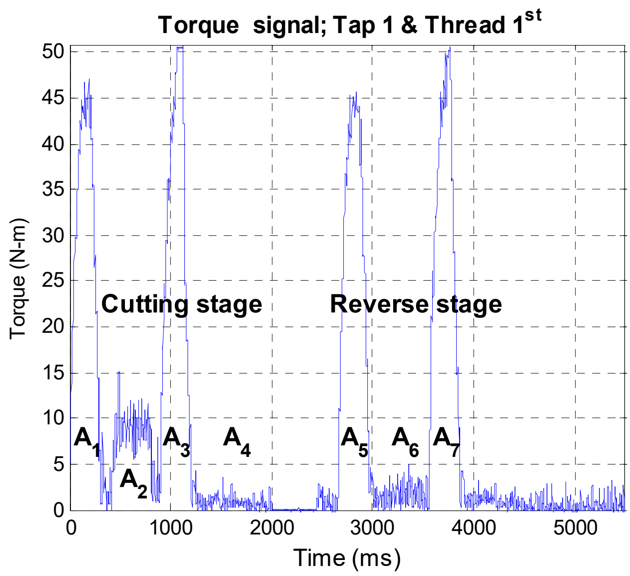


Figure 2: Torque signal from the spindle drive during tapping operation.

A_1, A_3, A_5 and A_7 parameters represent the torque evolution area of the main spindle motor in the acceleration or deceleration periods in cutting and reverse stages. Parameter A_2 is the torque evolution area of the main spindle motor while threads are cut by the tap, no longer be neglected, and therefore the A_3 value decreases. Parameter A_4 is to keep synchronism between rotation and the feeding movements. A_6 is the tap torque evolution area during tap reverse.

With all this parameter, a decision making is complicated and consequently, the parameter number is reduced. For this task, the Principal Component Analysis (PCA) will be applied.

3. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a well-established statistical technique that expresses multivariate data as a set of linear functions of “latent” variables (the principal components, PC’s). (Jackson, 1991) was the first to propose the use of principal components for quality monitoring of multivariate processes.

Let X be an $n \times p$ matrix containing the n observations of the p responses or signals of interest (in the tapping process, these are the A_i torque area values, so p = 7). It is good practice to standardize each column of X such that it has a mean of zero and a variance of 1 (hence, the values in the Z matrix can be negative or positive, contrary to the actual area measured which obviously are non-negative):

$$\mathbf{Z} = \left(\mathbf{X} - \overline{\mathbf{X}}\right)\mathbf{S}, \quad \text{where} \quad \mathbf{S} = diag\left\{1/s_{x_i}\right\} \tag{1}$$

and where s_{x_i} is the standard deviation of the ith column of X. The principal components Y are linear combinations of the (standardized) data:

$$\mathbf{Y} = \mathbf{ZV} \tag{2}$$

where matrix V is found in such a way that the variance-covariance matrix of Y is diagonal (and hence, the PC's are orthogonal). We have that:

$$\operatorname{cov}(\mathbf{Y}) = \mathbf{V}'\operatorname{cov}(\mathbf{Z})\mathbf{V} \tag{3}$$

and therefore V is the matrix that diagonalizes cov(Z). This is well-known by (Jackson, 1991) to be achieved by the $p \times p$ matrix of eigenvectors of cov(Z) (the eigenvectors are orthonormalized, i.e., V′V = I). Thus, the PC's Y point in p orthogonal directions in p-dimensional space where the data varies the most. The dimensionality of the problem is reduced by selecting only a subset of k (k < p) PC's that explain most of the variance in Z. The proportion of variability explained by the ith PC is given by $\lambda_i \big/ \sum_{j=1}^{p} \lambda_j$ where $(\lambda_1, \lambda_2, ..., \lambda_p)$ are the eigenvalues of matrix cov(Z).

the PCA of the six taps tested, it can be concluded that the first two PC's, after simplification are common to all taps, explain between 80% and 90% of the variability and are approximately equal to:

$$Y_1 = 0.25(Z_1 + Z_3 + Z_5 + Z_7) \tag{4}$$

$$Y_2 = 0.5(Z_2 + Z_6) \tag{5}$$

4. STATISTICAL PROCESS CONTROL

SPC is an anticipatory technique which suggests to stop a process before the production of a defective part. Therefore, in the tapping process, the SPC scheme may suggest a decision making to stop the process and replace the tap when the tap would have produced an additional number of acceptable threads until the end of tool life threshold. In decision making methods, this would constitute false positives (FP's). If the process were deemed in control past the end of life threshold of the tap this may not only result in defective threads but the tap may break inside the work piece, which could result in a considerable expense. In decision making analysis terms these would correspond to false negatives (FN's), and would be therefore of more economic importance than false positives (FP's). As shown next, the proposed monitoring scheme provides a good trade-off between these costs, not giving false negatives and giving a reasonably low false positive rate.

Let m be the number of observations (threads) in the training period. There are some studies that indicate how to choose m in multivariate SPC, but they usually recommend a relatively large number of observations (at least 50, Montgomery, 1996). If a too small training period is used, it is well-known (Montgomery, 1996) that the statistical power of the control chart will be low. Based on our desire to obtain a monitoring scheme that works effectively for all tap behaviors observed in our experiments, the following empirical rule was used to determine the end of the learning period in each tap: the end of the training period should occur when the observed A_2 parameter value equals 1.5 times the average A_2 value observed during the five first tapped holes.

The GV chart requires estimation of the instantaneous variance of the PC's, estimated following a suggestion in who use the moving ranges:

$$\mathbf{V_i} = \mathbf{Y_{i+1}} - \mathbf{Y_i}, \quad i = 1,\ 2, \ldots, m-1 \tag{6}$$

and suggested as estimator of the covariance matrix:

$$\mathbf{S}^* = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} \mathbf{V_i} \mathbf{V}_i'. \tag{7}$$

The statistic plotted at time (tapped hole) j on the Generalized Variance (GV) chart is the standard desviation of

$$Z_{i,j} = \frac{Y_{i,j} - \overline{Y_{L,i}}}{\sqrt{S^*_{i,i}}}, \quad i = 1, 2, \ldots, p \tag{8}$$

where $\overline{\mathbf{Y_L}}$ is the average vector of PC's for each tap over the training period. That is, the statistic plotted at time (or tapped hole number) j on the GV chart is then

$$GV_j = \left(\frac{\sum_{i=1}^{p} \left(Z_{i,j} - \overline{Z_j} \right)^2}{p-1} \right)^{1/2} \tag{9}$$

The Upper Control Limit (UCL) of the GV chart is then calculated as

$$UCL = B_4 \overline{GV} \tag{10}$$

where $\overline{GV}$ is average value of the GV statistic for each tap during the learning period, i.e.,

$$\overline{GV} = \frac{1}{m} \sum_{i=1}^{m} GV_i \tag{11}$$

and B4 is a constant that depends on m and corrects for bias (Montgomery, 1996). The process is deemed "out of control" when $GV_j > UCL$.

Once explained the methodology and the statistical techniques used to make a decision of the thread quality.

5. GV CONTROL CHARTS

For brevity only the GV control charts (Figure 3–5), corresponding to the taps that have been discussed so far (numbers 1, 3 and 6), are showed. As it was said before these taps represent reasonable well the behaviour of the whole set. As in Figures 3, 4 and 5, the real end of tool life (end thread quality, ETQ) is indicated by a vertical black line.

The GV control chart of tap 1 indicates that, from tapped hole number 175th onwards, all threads are out of control (non-acceptable). Figure 3 illustrates the moment when the operator might make a decision to stop the process before mechanizing an unacceptable thread. However, the tapping operation continued as the reliability is wanted to see. Experimental results have shown that the first non-acceptable one was actually number 192nd as shown by the black vertical line. Therefore, GV control chart for tap 4 gives an 8% of false positive (FP) (threads that being OK are considered unacceptable). No FNs have happened.

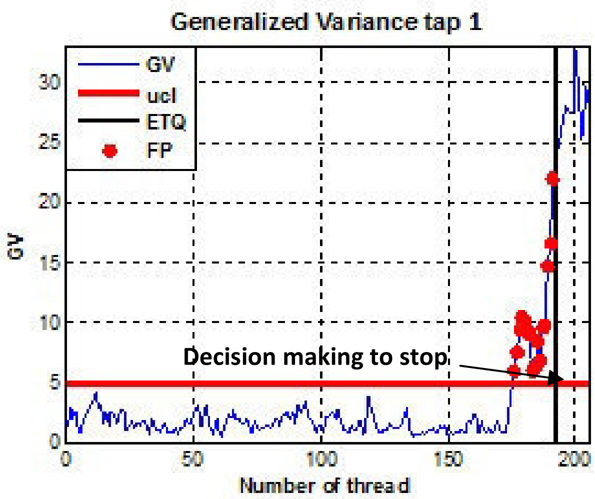


Figure 3: SPC control chart for tap1. First FP is the 175th thread.

Figure 4 shows the tap 3 behavior in a similar way as tap 1 did. The great difference is the tool life as this tap is the shortest tool life. The decision making to stop could be on 33rd thread. It presents a 6% of FP and a 0% FN.

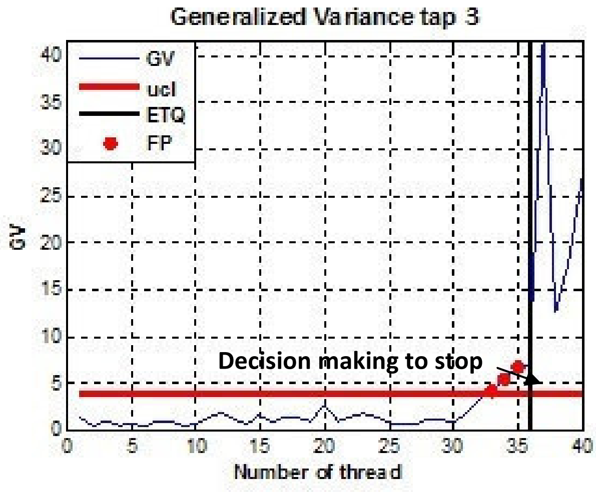


Figure 4: SPC control chart for tap1. First FP is the 33rd thread.

Tap number 6 GV control chart indicates that, again no FNs have occurred. The decision making to stop might be on 76th thread which gives a 30% of FPs as it can be showed in Figure 5. A great tool cost but there are not unacceptable threads which are the greatest industrial costs.

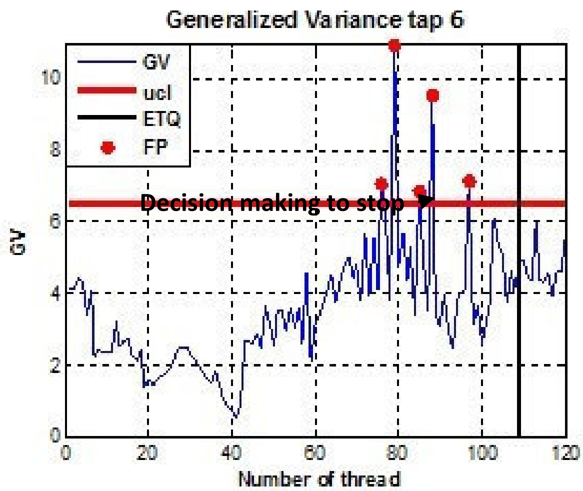


Figure 5: SPC control chart for tap1. First FP is the 78th thread.

The SPS strategy for decision making to stop the process before the lack of thread quality developed in this paper achieves the average FP rate across all the six taps in our experiments was close to 10% but no FN's were observed. Considering that the cost of the tool varies between 10 and 30% of the total manufacturing cost, the cost penalty due to the early process stoppages represents only between 1.4 and 3.2%. This can be considered acceptable for a manufacturing firm. The comparation with the savings of having a monitoring system that allows to work without operator or with only partial operator service.

6. CONCLUSIONS

A SPS strategy for decision making to stop the process for assuring good quality of M10x1.5mm threads machined with TiN Coated HSS taps and using torque signals of the spindle motor has been developed and validated for cast iron GGG50. It does not give any false negative (industrial cost) but it gives an average of 10% of false positives that increases tool costs. However the cost penalty in most cases can be assumed.

The strategy detects when a particular tap goes out of control as a result of tool wear and it has enough generality to be applied to different types of work and tap materials with different cutting speeds, tap diameters and geometries and can replace the operator vigilance task keeping at the same time the threads quality.

7. ACKNOLEDGEMENTS

The authors express their thanks to the Spanish Government for the support given to this research through the IDI-20100674 from CDTI programme, Public University of Navarre for the support this International Conference and also Groupe Tivoly for their experience as manufacturers.

8. REFERENCES

[1] Chen, Y.B., Sha, J.L., Wu, S.M. (1990). Diagnosis of tapping process by information measure and probability voting approach. *Journal of Engineering for Industry*, 112, pp. 319-325.

[2] Liu T. I., Ko E. J., Sha S. L. (1991). Diagnosis of Tapping Operations Using an AI Approach. *Journal Materials Shaping Technology*, 9 (1), pp. 39–46

[3] Li, W., Li D., Ni, J. (2002). Diagnosis of tapping process using spindle motor current. *International Journal of Machine Tools & Manufacture*, 43, pp. 73–79.

[4] Lorentz, G. (1989). Principal component analysis in technology. *Annals of CIRP*, 38 (1), pp. 107–109.

[5] Jackson, J. E. (1991), A user's guide to principal components. John Wiley & Sons, New York

[6] Montgomery, Douglas C. (1996). Introduction to statistical quality control. 3rd edition, John Wiley & Sons, New York

Modelling & Simulation as a Strategic Tool for Decision-Making Process in the Dairy Industry

Cristiani ECCHER[1], John GERAGHTY[2]

[1]*Dublin City University, School of Mechanical and Manufacturing Engineering Glasnevin, Dublin 9, Dublin, Ireland, e-mail: cristiani.eccher2@mail.dcu.ie*

[2]*Dublin City University, School of Mechanical and Manufacturing Engineering Glasnevin, Dublin 9, Dublin, Ireland, e-mail: john.geraghty@dcu.ie*

Abstract: The Dairy Industry faces many challenges compared to other sectors. On the supply side due to the nature of the raw material shelf life, large inventories are not applied; during the manufacturing process the flow of production is continuous which is highly sensitive to any sort of unplanned disruption; and on the demand side, the market dictates the commodity prices. In response to the growth in competition, dairy organizations' strategy must incorporate technology into their daily processes in order to become more efficient, profitable and sustainable. To achieve desired levels of improvement, Modelling and Simulation has been increasing in popularity in the decision-making process. Using a Dairy company as a study case, this paper has highlighted the potential for Modelling and Simulation to be used as a powerful strategic tool.

Keywords: Modelling and Simulation, dairy, decision-making process

1. INTRODUCTION

The Dairy sector is composed of farms, processors, retailers and distributors to reach the final consumers and faces several challenges compared to other sectors mainly because of the short lead time and the high competitiveness. Different to other commodities, dairy products requires specific equipment to be processed and to be stored, which increases considerably the cost of the final product. Moreover, the transportation needs to be efficient in order to improve the delivery time, since milk spoils rapidly without the appropriate refrigeration. As an alternative, many companies are increasing the value-added products mix with high profits by reducing the necessity of refrigeration and increasing the life cycle. Furthermore, the seasonality of this sector is another concern. Operating at low level cost is essential once the dependency of farms on supply raw material and market for the demand price are uncertain. The complexity to elaborate a production plan and follow strictly relies on the products mix processed and the efficiency of their sequencing. From the suppliers to the final customers a number of different levels of decisions are required.

The decision-making process is explored by (Farahani, et al., 2014; Gunasekaran et al., 2004; Ozbayrak, et al.,2007; Slack et al.,2013; Tonanont et al., 2008; Wisner et al.,2008;) and is divided into Strategic, Tactical and Operational. Through this segmentation time horizon and investment required are defined. Moreover, metrics and measures are created and compared according to each level:

1. At the strategic level long-term decisions are made and are normally followed by Business Plan strategies such as location, strategic partnership, new products and lifecycle products, make-buy decisions, new equipment, new plant capacity, competitiveness, international partnership and organisational goals.
2. At the tactical level medium-term decisions are made and the time horizon generally covers one year. The main objective is to support the strategic level previously defined and assure that the operational level will be followed as planned. Decisions regarding planning and scheduling, inventory policies, safety stock levels, transportation route and sequence of orders are generally decided at this level. In addition, measurements and Key Performance Indicators (KPIs) are compared to achieved results.
3. At the operational level short-term decisions are made and the time horizon can be measured daily and/or weekly. Operations previously planned at the Tactical level must be executed and any deviation must be reported and properly recorded to be managed in the future. Lessons learned are also reviewed to balance operations and future improvements. Daily transportation, inventory levels, production efficiency and order entries fulfillment are examples of activities executed at this level.

This paper is organized as follows. Section 2 gives an overview of applications in the dairy sector using simulation tools. Section 3 describes the simulation model built in ExtendSim® to support a dairy company decision-making process. Section 4 presents the statistics analysis used to validate the model. Section 5 then presents the findings and results. Section 6 presents the conclusion and future work.

2. MODELLING AND SIMULATION APPLIED TO THE DAIRY INDUSTRY

In this section, the application of Modelling and Simulation to the Dairy Industry will be presented. The following examples of research based on modelling and simulation have contributed to clarify the research question proposed by this case study: What is the potential of modelling and simulation to support strategic decision-making process? Even though several studies related to dairy and modelling and simulation processes were found, most of the findings have not explored the impact in different decision levels.

A Case Study was conducted by (Abed, 2008) where a discrete-event simulation (DES) model was developed through Arena® package to analyse the effects of each activity in the dairy process such as the mixing of raw material throughout the packing of the final product. Additionally, seven scenarios were produced and compared in order to propose an optimum solution. A simulation and optimisation-based decision support tool was developed (Li et al.,2008) between two factories and the integration of a large quantity of milk suppliers and the uncertainty in demand.

The incorporation of sustainable aspects of SCM is adopted by (Sonesson, 2003) through Life Cycle Assessment (LCA) and five approaches were explored in the flow of material. The model consists of the integration of transportation among farms, dairies, distributors, retails and households and aspects of packing, energy utilization, water and waste was explored. Furthermore, the use of resources such as wood for paper production and pallets; and oil for plastic packaging production was explored. LCA was also explored by (Nutter et al., 2013) to evaluate the global warming potential related to USA milk processing. The model evaluated the GHG emission per kg of packaged milk. However, the results are not considering other products in the dairy industry such as milk powder or casein where the electricity required to process these products is high.

(Shen et al., 2010) proposed a model based on quality-control and moral hazard considering the quality approach of the milk's depots and the processors. Effects such as additional compensation for customers, incoming inspection and independent investigation systems were evaluated in order to increase their quality to the final customers.

A centralized management optimization in (Lei et al., 2010) proposes a simulation method with a General Algebraic Modelling System (GAMS) comparing strategic decisions about centralization and decentralization and its impact on the supply chain network. An optimisation model applying (GAMS) combined with Microsoft Excel® to evaluate the cost of seasonality in Ireland was proposed by (Heinschink et al. 2016). Due to the fact that milk production relies mainly on grass-based and the effect of extra processing capacities required during the peak season is high. A financial analysis suggested changes in supply from a seasonal to a smother patter.

An economic approach was studied in (Guan and Philpott, 2011) describing the effects of contracts arranged several months in anticipation, especially the price-demand curve. A multistage stochastic programming model in a Dairy Company was conducted in this research. A payment system to compensate farmers for delivering quality goods dairy is proposed by (Fuentes et al., (2016) where a spreadsheet decision support tool was designed to calculate the profit based on suppliers' milk quality.

A model to represent the current condition between dairy companies and rural producers was conducted using System Dynamics (SD) by (Cezar et al., 2003). The model proposed a reduction of cost over time through better performance in quality and quantity metrics. SD was also applied by (Reiner et al. 2015) through a model to explore the effects of appropriated pricing strategy for wealth generation and health improvement at the Base of Pyramid (BoP). The research proposed insights to find the best price and decisions regarding distributions of dairy products in Bangladesh.

3. THE DAIRY CASE STUDY

In this case study, the industrial partner is an intermediate milk processor in the Dairy Supply Chain, and the final product is used as a raw material for other companies and processors. Most of their products are negotiated by contracts which are planned in advance. The company produces several dairy products, however, one specific product and the impact of its by-product generated was the main objective of this case study. Since the information about production is confidential, the company is referred to as DSC (Dairy Study Case) and PA for Product A and PB for Product B. The model was designed to simulate the same operation in the DSC processing which means that the milk is supplied hourly during the peak season. Additionally, after the transformation both products are packaged to be delivered.

A discrete-event simulation (DES) model was designed to deal with two branches of products carrying different production flows. The first branch is a continuous process and after reacting with HLC the separation between two products occurs. Still in the first branch, the product goes straight to the decantation process due to separation of solids from the liquid form. As soon as the solid becomes concentrated it is sent to the evaporation process to be transformed into the powder form. This product has a high level value-added and is frequent disrupted by its by-product work-in-process (WIP).

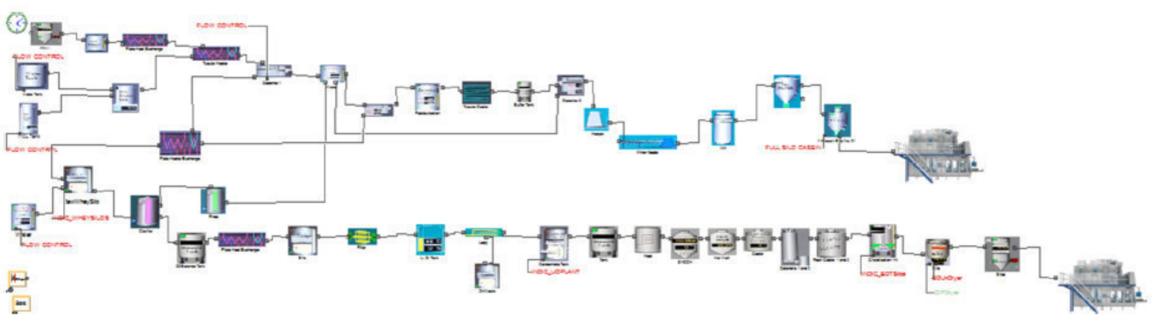


Figure 1: Model developed using ExtendSim®

In the second branch, a batch process is required accumulating a minimum level of product for further transformation. As soon as the by-product achieves a profitable level the flow is sent to the Ultra Filtration Plant (UO) in order to separate solids from liquid form. At this level, part of the product will be transformed into powder and packaged for customers and part of the product becomes waste. Figure 1 depicts the model developed through the ExtendSim® software and the two branches represent this specific process evaluated in the DSC.

An increase in the PA production was proposed due to the fact that the actual supply of PA to the market is currently below the demand generating backlogs which are negotiated among customers and transferred to the next periods.

Some important issues were considered due to direct impact on the product output: (i) the high level of WIP of PB, (ii) unplanned breakdowns, (iii) impact of rate and flow in each process. Despite the fact that PB is a by-product of PA, the models developed had to considered both products and their impact on each other's flow.

An interface was created to enable the input of data from non-technical users and the majority of the input parameters are available for changes. The model is easily customizable and allows managers entering their own rates and capacities. Additionally, a Microsoft Excel® interface was also performed for those users who are more familiar with spreadsheets. After the simulation run the output information is available for analysis. Figure 2 portrays the developed interface.

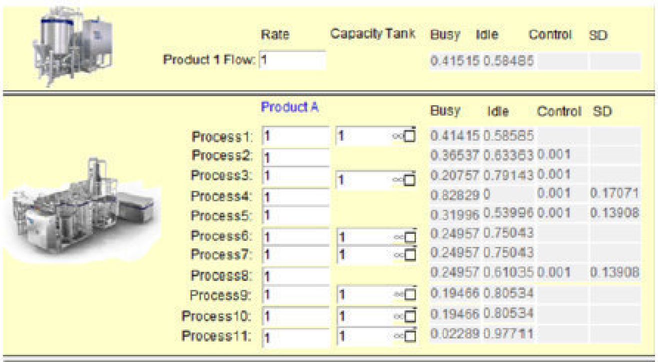


Figure 2: ExtendSim® interface for data input

The structure of the DSC presented was validated by means of: (i) interviewing the operator to understand the PA and PB production in the DSC, (ii) mapping the process, (iii) visiting the production line, (iv) validating the strategy with the strategic board (v) using company profiles, technical reports available from dairy, (vi) researching lasted outcomes literature.

4. MODEL VALIDATION

After the execution, the average of the all replications of PA and PB were calculated and compared to samples from the current records. The simulation package ExtendSim® provides an internal database where all the information can be tracked and stored hourly by each piece of equipment and the product flow.

Through IBM SPSS®, RStudio and Microsoft Excel® the yield produced monthly recorded in the dataset created by ExtendSim® was scrutinized. The comparison and the results produced by the model created will be explored as follows.

4.1 Time series

The real and simulation quantities of PA is demonstrated in a time-series graph in order to visually compare the data produced over an extended period of time. Even though the variability and uncertainties were incorporated in the model, the simulated output produced a small range of disparity compared to the real production. Through Figure 3 the time-series simulation and real production can be observed.

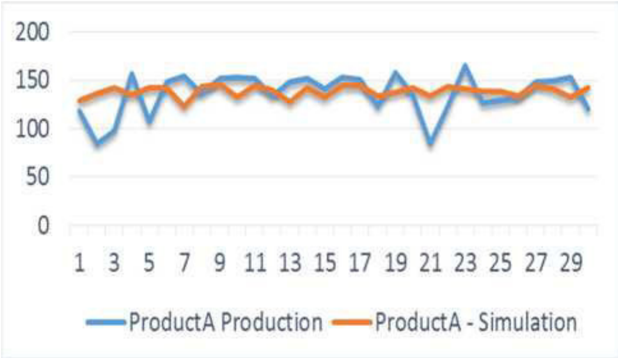


Figure 3: Time-series graph – Comparison between Production and Simulation

4.2 Boxplot Whiskers

The graph presented in Figure 4 depicts the behaviour of PA and PB comparing the production and simulation results through boxplot which are an important data representation of central value. It is possible to observe the records or points which are clustered around the simulated mean in PA and followed by two outliers. The discrepancy between lowest and highest limits are minor compared to PA real production. On the other hand, PB simulated represents similar behaviour compared to PB real production. In order to evaluate significant differences between these two measures a T-test is described in the next topic.

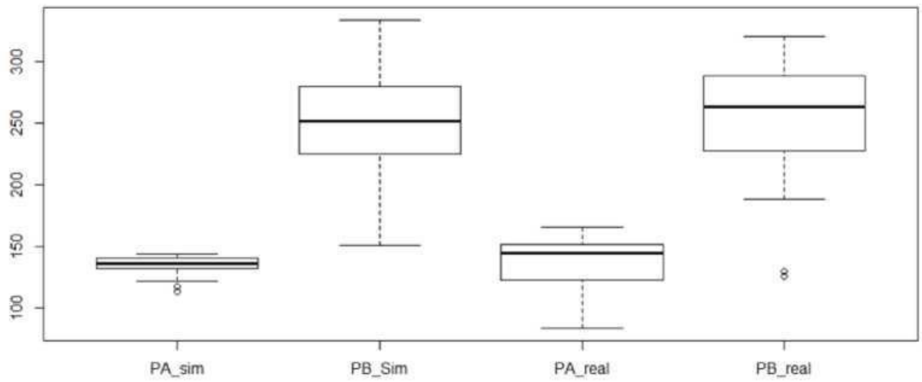


Figure 4: Boxplot Whiskers – Comparison between Production and Simulation

4.3 Independent-samples t-test and ANOVA

The independent-samples t-test was conducted to compare the two unrelated groups of variables (real produced and simulated). For PA, there was not a significant difference in the scores for production (M= 136.11 SD=21.48) and simulated (M=137.78, SD=5.92) conditions; t(30)=-0.441, p =0.683. For PB, there was not a significant difference in the scores for production (M= 252.66 SD=47.68) and simulated (M=247.91, SD=34.88) conditions; t(30)=0.441, p =0.661.

4.4 Nonparametric Tests

The last test was applied to assure that there is no significance differences between simulated and real production. The Whitney U Test was selected automatically by SPSS® according to the data evaluated, it is a nonparametric test that can be used to determine whether two dependent samples were selected. Figure 5 portrays the results performed by both products.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ProductA is the same across categories of Real/Simulated.	Independent-Samples Mann-Whitney U Test	.169	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ProductB is the same across categories of Real/Simulated.	Independent-Samples Mann-Whitney U Test	.701	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 5: Nonparametric Test for PA and PB

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

5.1 Total of Disruptions in hours

The histogram generated by SPSS® is a reasonable measure of frequency and followed by a normal bell shape suggests a good visualisation of the data distribution and concentration around the mean. Considering that the statistical analysis to compare the output of real production and simulation were satisfied a histogram depicts the distribution produced with a positive skewness caused by an excess of WIP in Figure 6 (M=137.78, SD=5.92). The experiments conducted in this section considered the possibility of reducing the impact of these disruptions on DSC flow.

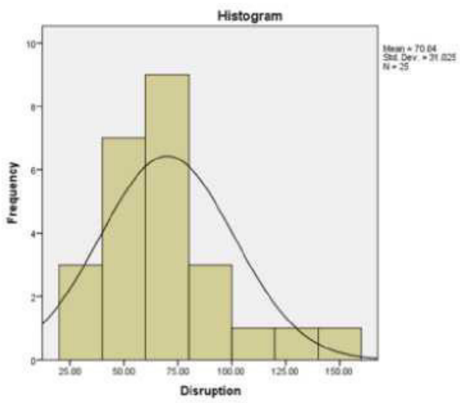


Figure 6: Histogram of disruptions in the current DSC process

In the DSC, as previously explained, PB is a by-product of PA. The increase in the PA flow will lead to approximately 75% more volume. Subsequently, any measured unit of PA will result in approximately 1.75 units measured of PB. However, the value of PA and PB in the market price suggests a different proportion: PB fluctuates around only 0.12 of PA unit value. To deal with this important trade-of, several scenarios were conducted combining reasonable strategies in order to suggest insights into what type of decision could be make and what type of results would be attainable. Not only is the production of PA dependent on the current bottleneck flow, but on farmers' supply, new bottlenecks that may exist and the process capacity of PB.

To visualize an increase in PA production ,the model considered follows the parameters in the DSC: (i) Raw Milk Rate is defined by the flow of milk per hour; (ii) Silo Capacity is one of the bottlenecks identified in the flow and impact the production flow; (iii) Actual Bottleneck is the main restriction in the flow; (iv) Concentrated Silo is a binary value used to demonstrate the production flow in PA and how it is affected by PB; (v) Level of ProdC is the concentrate produced by PB and is the main cause of disruption in PA; (vi) Disruption is the quantity of hours stopped in the PA flow.

The numbers shown in Table 1 are standardized in order to maintain significant results of confidentiality but provide important contributions to the literature. From 144 scenarios created, the top 10 of PA production were selected. The table has been sorted by the highest level of PA and PB obtained as responses through the chances in the factors such as: the Milk Rate, Silo Capacity, Actual Bottleneck, the use of an Alternative Silo. The model responses also demonstrate the level of concentration that would be produced and the quantity of disruption caused.

Table 1: PA and PB production associated with a possible increase in Milk Rate, followed by an increase in the intermediate Silo and Actual Bottleneck capacity, the use a Concentrated Silo and the level of the concentration and total of Disruptions caused in the DSC flow.

Scenario	PA	PB	Milk rate	Silo Capacity	Actual Bottleneck	Concentrated Silo	Level ProdC	Disruption
Scenario 040	2.031824868	1.806941487	0.585540044	1.788854382	1.224744871	1	20493	0
Scenario 101	1.954312652	1.944431354	2.342160175	0	2.449489743	0	0	4
Scenario 140	1.954312652	0.432042813	2.927700219	2.683281573	0	1	200451.77	0
Scenario 084	1.721776001	0.789516468	1.756620131	0.894427191	2.449489743	1	31798.438	0
Scenario 086	1.566751568	-0.10416767	1.756620131	1.788854382	0	1	421313.26	0
Scenario 057	1.385889728	1.188237083	1.171080088	0.894427191	1.224744871	0	0	11
Scenario 005	1.282540106	1.504463778	0	0	2.449489743	0	0	0
Scenario 114	1.282540106	1.078245189	2.342160175	1.788854382	2.449489743	1	51232.5	0
Scenario 042	1.2567027	1.311977964	0.585540044	1.788854382	2.449489743	1	27165.452	0

In Table 2 five scenarios were selected considering the current Bottleneck operating at the same level. These options provide insights into the quantity of PB produced. Even though there is a strong correlation between PA and PB (considering the fact PB is a by-product of PA), increasing the PA production will increase the level of concentrate in the liquid form that is not capable of being processed by DSC.

Table2: PA and PB production associated with a possible increase in Milk Rate, followed by an increase in the intermediate Silo maintaining the Actual Bottleneck at the same level.

Scenario	PA	PB	Milk rate	Silo Capacity	Actual Bottleneck	Concentrated Silo	Level ProdC	Disruption
Scenario 140	1.954312652	0.432042813	2.927700219	2.683281573	0	1	200451.77	0
Scenario 086	1.566751568	-0.10416767	1.756620131	1.788854382	0	1	421313.26	0
Scenario 134	1.050003455	0.170812064	2.927700219	1.788854382	0	1	153453.71	0
Scenario 122	0.998328644	-1.011600795	2.927700219	0	0	1	537567.41	0
Scenario 014	0.920816427	-0.365398418	0	1.788854382	0	1	271939.15	0

5.3 Discussion regarding increasing PA capacity

Some special attention for specific scenarios and its results must be highlighted. In Scenario 040 the maximum level of PA production would increase the current production in two standard deviations, however, it requires a high level of investment due to the fact that not only should the Milk rate be increased but the actual Silo Capacity and also the Actual Bottleneck. Moreover, investment to dealing with the excess of concentrate at the first stage of PB is also required. In this scenario no disruption would be caused.

In Scenario 101 a lower level of investment is required due to the fact that the Milk rate should be increased combined with a considerable elevation in the Actual Bottleneck capacity. No investment to deal with the excess of concentrated is required due to the fact the Bottleneck would be increased and capable to deal with the total flow. Some level of disruptions would occur in this scenario.

Table 3: Strategy adopted and the potential impact

Strategy	Description	Action	Drawback
S1	Increase Input Flow	Establish contracts with new potential suppliers and/or current suppliers in order to guarantee the level of milk desired	High dependency on S2 and S3 or S4
S2	Increase the Silo Capacity	Financial Investment and evaluation of the precise capacity required	High dependency on S3 or S4 High level of financial investmen
S3	Increase the BT Capacity	Financial Investment and evaluation of the precise capacity required	High level of financial investment
S4	Alternative Silo	Establish partnerships with other processors in order to send the Concentrated form to be processed by a 3P Financial Investment and evaluation of the precise capacity required	High level of financial investment

To support the decision-making process in this DSC the four strategies were evaluated in the current process: (i) Increase in the current Input Milk Flow: Not only will the quantity of milk determine the total PA produced but the TS (Total Solids) presented in the milk which is impacted the peak season of production. (ii) Increase in Silo Capacity: The PA is heavily impacted by PB and an extra capacity in PB would reduce the amount of disruption (demonstrated in Tablel) in the flow. (iii) Increase the Bottleneck Capacity: The BT is presented in PB process, however it is an expensive process and a break-even financial analysis is required. (iv) An alternative Silo support for the concentration of PB could be applied in the current operation and an increase in current production with lower financial investment (compared to S3). Table 3 describes the decision and the drawbacks related to each decision.

6. CONCLUSION AND FUTURE WORK

A DES model was developed to represent the real process in this DSC to support the decision-making process increasing a specific product production. Processes such as milk reception, for example, where the separation of the milk occurs daily in order to produce different branches of products was considered. The model was built to operate at the operation level identifying the bottlenecks in order to provide strategic insights to satisfy the customer's demand.

A potential for Modelling and Simulation for decision-making process was purposed. One-layer of the Dairy Supply Chain was considered to test the model and the scenarios created and some changes in the current process were suggested. Through the model developed, every improvement in the current flow, new bottlenecks will undoubtedly appear such as pipes, pumps, and balance tanks, for example, and new financial assessments will be required for the usage of what-if analysis to find reasonable solution in short, medium and long term. This is the natural behavior of any continuous improvement process.

The expansion to a multi-echelon and the relationship among other processors combined with costs and financial analysis would improve the model. Additionally, waste is an important aspect in all the dairy processors and highly restricts the capacity of process flow due to the legal requirement. Investment in a waste plant was not considered in this research due to its complexity and the deviation of the subject proposed. However, it can be supported by modelling and simulation to the making-decision process as future work.

7. ACKNOWLEDGMENTS

This work has been supported by Enterprise Ireland. Grant Agreement Number: TC 2014 0016

8. REFERENCES

[1] Abed, S.Y., (2008), Improving Productivity in Food Processing Industries Using Simulation - A Case Study, *12th WSEAS International Conference on SYSTEMS.* pp. 22–24.

[2] Cezar, F., Scramin, L., Batalha, M.O.,(2003), Assessment method for supply chain benefits, *Journal on Chain and network science.* 3 (2), pp. 135–146.

[3] Farahani, R.Z., Rezapour, S., Drezner, T., Fallah, S., (2014) , Competitive supply chain network design: an overview of classifications, models, solution techniques and applications, *Omega* 45, pp. 92–118.

[4] Fuentes, E., Bogue, J., Gomez, C., Vargas, J., Le Gal, P., (2016), "Supporting small-scale dairy plants in selecting market opportunities and milk payment systems using a spreadsheet model". *Computers and Eletronics in Agriculture.* 122, pp. 191–199.

[5] Guan, Z., Philpott, A.B., (2011), A multistage stochastic programming model for the New Zealand dairy industry, *Int. J. Production Economics*,134, pp. 289–299.

[6] Gunasekaran, A., Patel, C., McGaughey, R.E., (2004). Framework for supply chain performance measurement, *Int. J. Production Economics,* 87, pp. 333–347.

[7] Lei Bei Bei, Yang Chun Jie, Cao Jian, (2006), A centralized Optimization of Dairy Supply Chain Based on Model Predictive Control Strategy. *7th International Conference on Computer-Aided Industrial Design and Conceptual Design.*

[8] Li, W., Zhang, F., Jiang, M. (2008). A simulation-and optimisation-based decision support system for an uncertain supply chain in a dairy firm, *Int. J. Business Information Systems,* 3,(2), pp. 183–200.

[9] Nutter D.W., Kim D., Ulrich R., Thoma G., (2013), “Greenhouse gas emission analysis for USA fluid milk processing plants: Processing, packaging, and distribution”. *International Dairy Journal,* 31, pp. S57–S64.

[10] Ozbayrak, M., Papadopoulou, T.C. and Akgun, M., (2007), "Systems dynamics modelling of a manufacturing supply chain system", *Simulation Modelling Practice and Theory,15* (10), pp. 1338–1355.

[11] Shen Q., Hou Y., Lu X., (2010), Study on the Quality-Control Mechanism of Dairy Supply Chain Based on External Lose sharing model, *International Conference on Future Information Technology and Management Engineering*.

[12] Reiner G., Gold S., Hahn R.,(2015), WealthandhealthattheBaseofthePyramid: Modelling trade-offs and complementarities for fast moving dairy product case, *Int. J. Production Economics,* 170, pp. 413–421.

[13] Slack, N., Brandon-Jones, A., Johnston, R., (2013), Operations Management 7th edition – Pearson Education.

[14] Sonesson, U., Berlin J. (2003), Scenarios as a method to assess environmental impact in future milk supply chains. *Journal of Cleaner Production,* 11 (3), pp. 253–266.

[15] Tonanont, A., Yimsiri, S., Jitpitaklert, W., & Rogers, K.J. (2008). “Performance evaluation in reverse logistics with data envelopment analysis”. *Proceedings of the 2008 Industrial Engineering Research Conference*, pp. 764–769.

[16] Validi, S., Bhattacharyan A., Byrne P.J. (2014). A case analysis of a sustainable food supply chain distribution system—A multi-objective approach. *Int. J. Production Economics*. 152, pp. 71–87.

[17] Wisner, J.D, Tan, K.C, Leong, G.K, (2008), Principles of Supply Chain Management – A balance Approach, Cengage Learning.

Extracting Weights and Ranking Alternatives Based on Fuzzy Pairwise Comparison Matrices Using a Brain Inspired Method for Solving Fuzzy Multi-Criteria Decision Making Problems

Mohamed Naili[1,2], Abdelhak Boubetra[2], Abdelkamel Tari[3], Makhlouf Naili[4]

[1]*Computer science department, Faculty of Exacte Sciences,*
*Univers*ity of Bejaia, 06000, Bejaia, Algeria, e-mail: naimo1987@gmail.com

[2]*Computer science department,*
University of Bordj Bou Arreridj, 34030, Bordj Bou Arreridj, Algeria

[3]*Laboratory of Medical Computing (LIMED), Faculty of Fundamental Sciences,*
University of Bejaia, 06000 Bejaia, Algeria

[4]*Department of Computer Science, Faculty of Exact and Natural Sciences,*
University of Biskra, 07000, Biskra, Algeria

Abstract: Nowadays e-commerce field is full of questions like: "How to improve my website ranking?" or "which is the best website to publish ads in?". this kind of problems fall into what's known as Multi-Criteria Decision Making problems in which a set of criteria has to be carefuly defined and evaluated especialy in case of uncertainty. To tackle this challenge, the authors of this paper present an application of a brain-inspired method, in which criteria' weight are extracted on the basis of pair-wise comparison matrices. Afterwards, the set of criteria are used to rank a set of medical institutions' websites.

Keywords: Artificial intelligence, Fuzzy set theory, Multi-criteria decision making, Pair-wise comparison matrix,Fuzzy AHP, Web intelligence

1. Introduction

Whatever the decision maker is facing a problem of buying, selling or even playing, an underlying process of ranking is ongoing deep inside the brain. Often, the difficulty of dealing with such kind of problems is due to the uncertainty of the decision maker, such uncertainty is coming often from the leak of knowledge or the integrity of the information sources, such problems require more advanced artificial intelligent systems to help the decision maker in his/her decision process. To address the uncertainty, many theories have been proposed such as the *Rough set* theory (Pawlak, 1982), the *Intuitionistic fuzzy set* theory (K. Atanassov, 1986; K. Atanassov, & Gargov, G, 1989), or the most known the *Fuzzy set* theory (L. Zadeh, 1965; Zimmermann, 2010).

In the *Fuzzy set* theory, the membership of an element to a given set is not a binary function like it is in the conventional set theory, but rather than it is a graduated function from 0 (not a member) to 1 (a full member). On the basis of this concept of membership, many methods have been proposed to deal with the problem of fuzzy multi-attribute decision making problems (Kahraman, 2008). To systematically find a solution to such kind of problems, the definition of a consistent, non-redundancy set of attributes, or what we refer to here as criteria, as well as the nature of each one (qualitative or quantitative) is highly required. Besides the definition of the criteria set, the decision maker has to assign also to each one the suitable weight or importance which reflects really its role in the decision making process, then use these attributes to evaluate each alternative (each alternative has (if possible) the accurate scores against each criterion). Using its performances against the criteria set, the global performance of each alternative has to be calculated. There are two very used methods for such problems: the *Fuzzy Technique for Order of Preference by Similarity To Ideal Solution (FTOPSIS)*, and the *Fuzzy Analytic Hierarchy Process* methods *(FAHP),* based on the Fuzzy set theory and the *TOPSIS* (Hwang & Yoon, 1981) and *AHP* (R. W. Saaty, 1987) respectively.

The *FTOPSIS* method is based on the notion of distances between each alternative and the Ideal positive/negative solutions, where each one of these solutions is derived from the best/worst (respectively) scores of alternatives with respect to each criterion. The second family of methods, *FAHP*, is based on the pair-wise comparison matrix of the alternatives for each criterion. The pair-wise comparison matrix has not been used only to calculate the ranking of a set of alternatives, but also it has extensively used to calculate the criteria' weights which a very crucial step in the decision is making process. As we will see in a subsequent section, this method takes into consideration the uncertainty of the decision maker, but what about its impact on him/her? To tackle this problem, we present in this work an application of an artificial intelligence method, inspired from some brain's activities during the cognitive process of decision making in case of uncertainty, it is called the *Brain-inspired method for solving fuzzy multi-criteria decision making* *problems* method *(BIFMCDM)* (Naili, Boubetra, Tari, Bouguezza, & Achroufene, 2015). This method will be used to derive crisp weights from a fuzzy pair-wise comparison matrix.

The remaining of this paper is organized as follows: section two presents some important methods of the Fuzzy AHP methods family used to derive the criteria's weights from a pair-wise comparison matrix. In section three, we discuss the big lines of the *BIFMCDM* method. To study the performance of this method, we discuss in section four the application of the *BIFMCDM* to handle the problem of electronic service quality ranking in the healthcare industry in *Turkey*. We finish this article with a global conclusion in which we highlight some of the important future trends and research opportunities in this field.

2. Related works

In order to make the remaining of this paper understandable, it would be better to present some important concepts in the fuzzy set theory.

2.1 Important concepts

- $\boldsymbol{\alpha}-\mathbf{cuts}$

The membership to a given fuzzy set *F*, could be restricted on the basis of a real number $\boldsymbol{\alpha} \in [\mathbf{0}, \mathbf{1}]$ through an $\boldsymbol{\alpha}-\mathbf{cut}$ or $\boldsymbol{\alpha}-\mathbf{levels}$. The membership function of the derived set $\boldsymbol{F^{\alpha}}$ is then defined as follows,

$$F^{\alpha} = \{x \in U / \mu_F(x) \geq \alpha\} \tag{1}$$

- **Fuzzy numbers**

If the universe of discourse is the real number set ℝ, then the fuzzy set *F* is a *Fuzzy Quantity*. A fuzzy quantity is considered also as a *Fuzzy number* if it is *Normalized* (only one element has the membership equal to 1) and *Convex* with a piecewise continuous membership function (Bloch et al., 2003; Zimmermann, 2010).

- **Triangular fuzzy number**

A fuzzy number is said to be a *Triangular Fuzzy Number* (*TFN*), if its membership function plot has the shape of a triangle, and this is by the definition of three values of its support, let's be *l, m, u*, which would represent the three vertices of the triangle,

- *l and u* stand for the lower and the upper values respectively. For these two values, the membership degree to the fuzzy set represented by the *TFN* is the lowest.
- m represents the modal value, with a membership degree equal to 1.

The membership function $\boldsymbol{\mu_F(x)}$ could be formulated as follows,

$$\mu_F(x) = \begin{cases} 0 \; x \leq l \\ \frac{x-l}{m-l} \; l < x \leq m \\ \frac{u-x}{u-m} \; m < x \leq u \\ 0 \; x > u \end{cases} \tag{2}$$

For a given $\boldsymbol{\alpha - \mathbf{cut}}$, the *TFN* could be represented by an interval as follows,

$$F_\alpha = [(m-l)\alpha + l, u - \alpha(u-m)] \tag{3}$$

- **Linguistic variables**

In order to give more abstraction to the surrounding world, human beings use expressions that approximate the real measures, which are most of time subjective. For example, they express distance by using terms such as far, near or close instead of using accurate measures in meters for example, and these expressions as they present a degree of simplicity, they include also a degree of fuzziness. To take advantage of such approximate reasoning, in the frame of his *Fuzzy set* theory, *Zadeh* has proposed another concept of variables, called *Linguistic variables*. Let's *x* be a linguistic variable, this variable is defined on a universe of discourse *U*. This variable has two important rules, the syntactic and the semantic rules. The syntactic rule *G* is used to associate a term, from a predefined set of terms *T(x)*, to this variable. The semantic rule *M*, is used to define the fuzzy set *M(X)*, for the term *X* (Bloch et al., 2003; L. A. Zadeh, 1975; Zimmermann, 2010).

2.2 Fuzzy Analytic Hierarchy Process

As we have mentioned earlier, the ranking process can't be carried out without defining the criteria set and their weights. Most of the time, these weights couldn't be defined accurately which makes of using approximate expression very helpful for experts. As one of the most used methods for this kind of problems, we find the *Fuzzy AHP* methods family. As a combination between the *Fuzzy set* theory to handle fuzziness in evaluations, and the conventional *AHP* method proposed by *Saaty* (R. W. Saaty, 1987; T. L. Saaty, 1990) these methods give a wide range of possibilities to derive fuzzy and crisp weights from pair-wise comparison matrices. In the next sub-section a brief explanation of the most known methods of this family.

The cornerstone of the AHP method is the decomposition of the problem into a hierarchy of sub-problems. After the determination of goals, criteria and the alternatives sets, the decision maker has to construct the pair-wise comparison matrices of each one of them in order to derive weights, and here is the problem! How could we derive weights from fuzzy pair-wise comparison matrices? In the literature, there are two main classes of the *Fuzzy AHP* methods family proposed to cope with such problem: the fuzzy values (fuzzy numbers) and the crisp values classes. The fuzzy values class includes many methods used to derive fuzzy weights such as *Van Laarhoven, P. J. M.* and *Pedrycz, W* (Boender, de Graan, & Lootsma, 1989; Van Laarhoven & Pedrycz, 1983) method, *Buckley* method (Buckley, 1985) and *Boender, C. G. E., De Grann, J. G.* and *Lootsma, F. A* method (Boender et al., 1989). As shortcoming of this class of methods, the decision maker needs most of the time to combine the method with other techniques such as the *Fuzzy TOPSIS* to find the final ranking, besides the nature of weights which still fuzzy and hard to apprehend by human beings. On the other hand, as we will see in the next subsection, the second class of the *Fuzzy AHP* methods family yields approximate, but not fuzzy value for weights, such results give the decision maker the opportunity to handle the problem with more accurate values rather than fuzzy quantity of information as it is the case with the first class.

- **Crisp Values class**

In order to derive crisp values from fuzzy pair-wise comparison matrices, *Chang* (Chang, 1996) has focused on the comparison between the pair-wise comparison matrix's elements, in such a way the degree of possibility associated to a given statement " the element e_1 (which is a *TFN*) of the pair-wise comparison matrix is greater or equal than e_2" has to be calculated. Another method has been proposed to tackle such problem, it is the *Fuzzy preference programming method (FPP)* of *Mikhailov* (Mikhailov, 2003). In the frame of this proposition, the author has relied on the concept of $\boldsymbol{\alpha}-\mathbf{cut}$ of each element of the pair-wise comparison matrix, where the weights to be derived have to satisfy the following inequality ($\tilde{\leq}$ denotes the statement "*fuzzy less or equal to*"),

$$l_{ij}(\alpha) \tilde{\leq} \frac{w_i}{w_j} \tilde{\leq} u_{ij}(\alpha) \tag{4}$$

Another example of these methods, is the *Fuzzy AHP-based simulation approach to concept evaluation in a NPD environment* of *Ayağ* (Ayağ, 2005). Based on the same concept of $\boldsymbol{\alpha}-\mathbf{cut}$ as the Mikhailov's method, this method aims at the calculation of the biggest fuzzy eigenvalue and its corresponding eigenvector of a new pair-wise comparison matrix derived from the $\boldsymbol{\alpha}-\mathbf{cut}$ of the elements of the initial matrix. On the basis of the new derived matrix, it would be possible to calculate the degree of optimism $\boldsymbol{\mu}$ (Ayağ, 2005) as a linear convex combination as follows,

$$\tilde{a}_{ij}^{\alpha} = \mu a_{iju}^{\alpha} + (1-\mu) a_{ijl}^{\alpha} \tag{5}$$

Using this optimism degree, it would be possible to calculate crisp values for the weight vector associated to the pair-wise comparison matrix.

Relying on the $\boldsymbol{\alpha}-\mathbf{cut}$ concept and other approaches such as the degree of possibility in Chang's method or the optimism degree introduced in *Ayağ* method, uncertainty processing of the pair-wise comparison matrix becomes easier, but do these methods take the impact of the uncertainty on the decision maker in the first place? The answer is no. In order to tackle this shortcoming, a new method based on the brain's activity pattern has been proposed, it's called the BIFMCDM. The BIFMCDM (*Brain Inspired method for solving Fuzzy Multi-Criteria Decision Making problems)* has been designed on the basis of a special neural network associated to some emotions experienced by the subjects when they face a decision making problem with uncertainty. In the following section we see the idea behinds this method.

3. Brain-Inspired method for solving Fuzzy Multi-Criteria Decision Making problems (BIFMCDM)

If we can say that our knowledge can make of our decision wiser and stronger, then it is fair enough to say that our emotions can bias our decision, and sometimes define it. For example, when the one of us wants to subscribe to an online service, let’s say a website hosting service, even with a great knowledge about the domain of websites hosting, in face of a new offer the decision maker will experience at the same time two emotions, *Fear* of making a bad deal or *Happiness* if his/her choice will turn out as the best one. Based on such approach, in which “*The more we feel happy than we are afraid of an alternative, the more we tend to take this one*” we have proposed the *BIFMCDM* method(Naili et al., 2015).

When the decision maker faces uncertainty in the alternatives’ evaluation for example, he/she experiences these two emotions, and his/her brain recognizes some patterns of activities in which many regions involve. Based on the evaluation of a set of alternatives, a perfect alternative could be generated from their best performances with respect to the criteria set. On the basis of the *Symbolic distance(s) SD*, to that perfect alternative; each alternative could cause an amount of *Happiness* and *Fear*. The symbolic distances are terms of a linguistic variable “distance” which has as the universe of discourse the real numbers set ℝ.

In **Błąd! Nie można odnaleźć źródła odwołania.**, we present an abstraction of the neural network which represents the calculation of the final performance of an alternative against a given criterion.

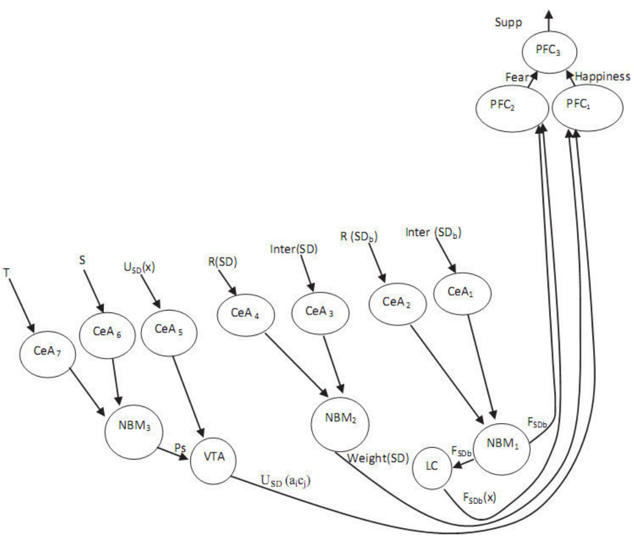


Figure 1: Neural network of calculating the support of being the best alternative against a given criterion (Naili et al., 2015)

Błąd! Nie można odnaleźć źródła odwołania. Figure.1 represents a neural network of three main layers:

Input layer:

- *T*: represents the surface of the triangle created by subtracting the fuzzy performance of a given alternative “i”, which is represented by a *TFN,* from the perfect alternative performance against a given criterion “j”.
- *S*: the intersection between *T* and the support of a given *M (SD*).
- U_{SD} *(x)*: represents the degree of membership of an element of the *T*’s support to a given *M(SD)*. If *T* is a right triangle or *T’s* support contains one element, then the membership to

M(SD) will be doubled. In case of more than one intersection between *T* and *M(SD)*, the membership is equal to the sum of these membership degrees.

- *R (SD)*: represents the importance of *SD*.
- *Inter (SD)*: the *SD*'s support divided by the longest max distance,
- *R* (SD_b): the importance of SD_b , where this last represents a SD which is better than the closest "SD_c" of all the possible *SD*s between the alternative and the perfect alternative, or it is one of the possible *SD*s except SD_c.
- *Inter* (SD_b): the ration of SD_b's support to the longest max distance,

Hidden layer:

- *NBM_2*, calculates the weight of a given SD (Naili et al., 2015),
$$Weight(SD) = R(SD) \times Inter(SD) \tag{6}$$
- *NBM_3*, calculates the probability used to calculate the invalidity $(1 - Ps)$ of a given alternative (Naili et al., 2015),
$$Ps = S/T \tag{7}$$
- *VTA,* calculates the precision of the chosen value $U_{SD}(a_i c_j)$ (Naili et al., 2015),
$$U_{SD}(a_i c_j) = U_{SD}(x) \times Ps \tag{8}$$

Knowing that $U_{SD}(x)$ is the membership degree of *x* to the fuzzy set *M(SD)*.

- *NBM_1*, calculates the expected uncertainty related to fear F_{SD_b} (Naili et al., 2015),
$$F_{SD_b} = Weight(SD_b) = R(SD_b) \times Inter(SD_b) \tag{9}$$
- *LC*, calculates the unexpected uncertainty related to fear $F_{SD_b}(x)$ (Naili et al., 2015),
$$F_{SD_b}(x) = P(x \in SD_b / F_{SD_b}) \tag{10}$$
- *PFC_1*, calculates the happiness (Naili et al., 2015),
$$Happiness_{a_i c_j} = \sum_{k=1}^{K} Weight(SD_k) \times U_{SD_k}(a_i c_j) \tag{11}$$
- *PFC_2*, calculates the fear (Naili et al., 2015),
$$Fear_{a_i c_j} = \sum_{b=1}^{B} F_{SD_b} \tag{12}$$
- *Output layer*. On the level of this layer we find the *PFC_3* neuron; this last calculates the *support* which represents the difference between the fear and the happiness as follows (Naili et al., 2015),
$$Support_{a_i c_j} = Happiness_{a_i c_j} - Fear_{a_i c_j} \tag{13}$$

3.1 Criteria's weights

As mentioned in (Naili et al., 2015), in case of pair-wise comparison matrix, the weights of the criteria set could be calculated by assuming that the columns represent the criteria and the lines stand for the alternatives, and the perfect alternative is the best performance (the *u* or *l* of the TFN if we have a benefit or cost context respectively) found in the matrix. After calculating the support's value for each criterion, the criteria's weights are calculated as follows (Naili et al., 2015):

$S(c_j)$= Support of the criterion *"c_j"*

$$d(c_j) = S_{max}(c) - S(c_j) \tag{14}$$

$$R = \sum_{j=1}^{J} d(c_j) \tag{15}$$

$$weight(c_j) = \left(R - d(c_j)\right)/R \tag{16}$$

Where $S_{max}(c)$ is the best support found with respect to all criteria.

3.2 Ranking

In order to find the final ranking of the set of alternatives, three relations have been defined: *preference, strict preference* and *indifference* as follows:

if "a" and "b" are two alternatives then,

- Preference($\geq$ or Q): a is preferred than b
- Strict preference ($>$ *or* P): a $\geq$ b and not (b $\geq$ a)
- Indifference($\sim$ or I): a $\geq$ b and b $\geq$ a

With all the criteria,

Lets $c_j \in$ Criteria set $(j = 1..J)$, the final preference relation is defined as follows:

$$As = \sum_{j=1}^{J} Support\ (ac_j) \times Weight(c_j) \tag{17}$$

$$Bs = \sum_{j=1}^{J} Support\ (bc_j) \times Weight(c_j) \tag{18}$$

- a P b$\Leftrightarrow$ As $> B$s
- a Q b$\Leftrightarrow$ (As = Bs) AND δ is true

where:

c_j : is the "jth" criterion of the criterion set (ordered (top down) by weight)

$$\delta \text{ is true if: } \emptyset \text{ is true for } c_j, \text{ if } \emptyset \text{ is false AND } \left(U_{CSD_r}(a) = U_{CSD_s}(b)\right) \text{ for all } CSD_{r,s}, \text{ check with } c_{j+1}.$$

- a I b$\Leftrightarrow$ is false AND $\left(U_{CSD_r}(a) = U_{CSD_s}(b)\right)$ for all $CSD_{r,s}$ with all the criteria.

In the end of this section, we can say that unlike the other known methods in the field of pair-wise comparison, the *BIFMCDM* method gives more importance to how the fuzziness of the information could affect the decision maker, which gives more accuracy to the final decision. In order to assess the performance of this method, in the next section we present its application to solve a problem of ranking in the domain of electronic service quality in the healthcare industry, where all information from criteria' weights to alternatives' evaluations are given in the form of *TFN*s.

4. Case study

In the following case study, we aim at studying how could this method applied to such kind of problems, and of course discuss its performances in comparison with *Ayağ's* Fuzzy AHP method (Ayağ, 2005) proposed to deal with the same problem.

- Electronic service quality in the healthcare industry

Any electronic business has to guarantee a secure, accessible, easy to use and even personalized service to ensure the minimum satisfaction of these days' users. As a very profitable electronic business, the electronic service in the healthcare industry has also takes a lot of attention of professional and academics studies, we discuss in this paper the *Büyüközkan* and Çifçi works (Büyüközkan & Çifçi, 2012). In their works, the authors have studied the quality of 13 websites of different medical institutions in *Turkey*. They have based their assessment on a set of criteria and sub-criteria to evaluate in general the service's content, security and interaction with the user. Table 1 summarizes the set of criteria and sub-criteria used in their study.

Table 1: E-sq evaluation criteria for hospital web sites (Büyüközkan & Çifçi, 2012).

Criteria	Sub-criteria
Tangibles (C_1)	Usability (C_{11}), Animation (C_{12}), Design (C_{13}), Functionally (C_{14})
Responsiveness (C_2)	Customer service (C_{21}), Technical performance (C_{22}), Interactivity (C_{23})
Reliability (C_3)	Specialization (C_{31}), Standardization (C_{32}), Reputation (C_{33}), Accuracy of service (C_{34})
Information quality (C_4)	Information richness (C_{41}), Information accuracy (C_{42}), Information up-date (C[illegible])
Assurance (C_5)	Compensation (C_{51}), Trust (C_{52}), Security/privacy (C_{53})
Empathy (C_6)	Customer care (C_{61}), Links (C_{62}), Customization (C_{63})

To evaluate each of these criteria and the alternatives set, the authors have defined a set of *TFN*s to express the value of each criterion and alternative as well (see Table 2).

Table 2: Definition and membership function(MF) of fuzzy scale (FN) (Büyüközkan & Çifçi, 2012).

FN	Definition	MF
$\tilde{9}$	Extremely more importance (EMI)	(8, 9,10)
$\tilde{7}$	Very strong importance (VSI)	(6, 7, 8)
$\tilde{5}$	Strong importance (SI)	(4, 5, 6)
$\tilde{3}$	Moderate importance (MI)	(2, 3, 4)
$\tilde{1}$	Equal importance (El)	(1, 1, 2)

Also the authors have established Table 3 to present the pair-wise comparison matrix of the main criteria with respect to the goal.

Following similar pairwise comparison matrices, the authors have compared the sub-criteria with respect to tangibles, responsiveness, reliability, information quality, assurance and empathy (BÜYÜKÖZKAN & ÇIFÇI, 2012). In their works (BÜYÜKÖZKAN & ÇIFÇI, 2012), the authors have adopted the *Fuzzy AHP* method of Ayağ (Ayağ, 2005) to calculate the criteria's and the sub-criteria's weights with $\alpha = \mu = 0.5$. As results they found that the specialization, interactivity and the accuracy of service are the most important factor in the hospital web sites quality assessment, whereas Animation stands for the weakest criterion . As we can notice in Table 5, Dünya Göz (H_6), Acibadem (H_2), International (H_3) have the three best websites, whereas JFK (H_7) comes in the end of the ranking.

Table 3: Pairwise comparison matrix of criteria (Büyüközkan & Çifçi, 2012).

	C1	C2	C3	C4	C5	C6
C1	1	1 /MI	1/SI	1/MI	EI	EI
C2	MI	1	1/MI	MI	MI	MI
C3	SI	MI	1	MI	SI	SI
C4	MI	1/MI	1/MI	1	EI	MI
C5	1/EI	1/MI	1/SI	1/EI	1	MI
C6	1/EI	1/MI	1/SI	1/MI	1/MI	1

After this overview of results found using the combination of *Fuzzy AHP* and *Fuzzy TOPSIS* methods, we present in the next part of this section the results of processing the same problem using the *BIFMCDM* method. The corresponding relative weights of each sub-criterion are presented in Table 4.

Table 4: All the criteria's and sub-criteria's weights

Criterion	Criterion Relative weight	Sub-criterion	Sub-criterion's Relative weight	Sub-criterion's weights
C_1	0.7751			
		C_{11}	1	0.7751
		C_{12}	0.39235	0.30413
		C_{13}	0.93986	0.7253
		C_{14}	0.66777	0.51762
C_2	0.8671			
		C_{21}	0.4476	0.3881
		C_{22}	0.5523	0.4789
		C_{23}	1	0.8671
C_3	1			
		C_{31}	1	1
		C_{32}	0.6020	0.6020
		C_{33}	0.4454	0.4454
		C_{34}	0.9524	0.9524
C_4	0.8213			
		C_{41}	0.3497	0.2872
		C_{42}	1	0.82138736
		C_{43}	0.6502	0.5340
C_5	0.7866			
		C_{51}	0.3497	0.2751
		C_{52}	1	0.7866
		C_{53}	0.6502	0.5115
C_6	0.7496			
		C_{61}	1	0.7496
		C_{62}	0.5298	0.3971
		C_{63}	0.4701	0.3524

From 0Table 4, we can see that using the *BIFMCDM* method has led us to realize that Specialization, Accuracy of service and Interactivity are the most important criteria whereas the

Compensation of customer in case of problems (C_{51}) has turned out as the weakest criterion. To present the medical institutions ranking found using the *BIFMCDM* method, 0Table 5 gathers the supports and the ranking of each one of them.

Table 5: The Institutions' support, PI and ranking.

	BIFMCDM		Fuzzy AHP TOPSIS	
Institution	Support	Ranking	P.I	Ranking
H_1	-66.53	8	0.0259	5
H_2	-38.62	1	0.0367	2
H_3	-47.56	3	0.0325	3
H_4	-77.40	10	0.0172	10
H_5	-63.51	5	0.0275	4
H_6	-38.93	2	0.039	1
H_7	-84.11	13	0.012	13
H_8	-63.55	6	0.0252	6
H_9	-79.44	11	0.0157	11
H_{10}	-82.34	12	0.0128	12
H_{11}	-59.36	4	0.0241	7
H_{12}	-72.92	9	0.0219	9
H_{13}	-64.85	7	0.0233	8

On the basis of the results shown in 0Table 5, by using the *BIFMCDM* method, the websites have a very different ranking in comparison with the one found using the *fuzzy AHP* method (except with *JFK (H_7)* which represents the worst websites using both methods). To explain these differences we have to look back at the substrates of each method. In the *Fuzzy AHP* method adopted in this problem, the authors have relied on the $\alpha - cut$ and the index of optimism μ. The $\alpha - cut$, as the authors have already mentioned (BÜYÜKÖZKAN & ÇIFÇI, 2012), incorporates the decision maker's confidence related to his/her preferences.

As for the index of optimism μ, this last reflects the degree of satisfaction of the evaluation of the pair-wise comparison. These two parameters as it is illustrated in their definitions, reflect the confidence and the satisfaction of the decision maker, and they have been considered both as 50% for all the evaluations in the pair-wise comparison matrices, such assumption which doesn't efficiently reflect the impact of fuzziness of these evaluations. For instance, the *TFN* (1,1,2) which becomes [1,2] and the *TFN* (2,3,4) which becomes [2,4] are both triangular fuzzy numbers with supports equal to 1 and 2 respectively, but the intervals [1,2] and [2,4] don't reflect the confidence the decision maker had when he/she stated that the first *TFN* may take the value 1 more than any other values between 1 and 2, and his/her uncertainty, when he/she stated that (2,3,4) (the ratio of importance between two given criteria for example responsiveness to information quality) is with a support equal to 2, such kind of confidence and uncertainty are well represented in the *BIFMCDM* method. In this method, the confidence of the decision maker represented by the right *TFN* (1,1,2) has been taken into consideration through doubling the membership to the corresponding fuzzy sets represented by the symbolic distances Close2 (0.16666667), Near1 (0.33333333), Near2 (0.33333333) and Far1(0.16666667) which reflect better the intention behind his/her fuzzy evaluation.

On the basis of the same approach, we can explain all the other differences for the criteria, sub-criteria weights. Besides the impact on of the differences between the criteria and sub-criteria weights yielded from both methods, the alternatives' importance index has been affected also, with the same manner by the uncertainty of the decision maker represented in his/her fuzzy evaluations using TFNs, such fuzziness which has been processed better using the *BIFMCDM*.

5. Conclusion

Based on some brain activities noticed during uncertainty processing, the *BIFMCDM* method takes advantage of the decision maker' fear and happiness emotions to help in getting more accurate results which reflect the real intention of the decision maker hidden behind his/her uncertainty. After applying this method to deal with a problem of ranking of a set of websites, which is a very important issue in the *Web intelligence* field, we have came to a conclusion that this method has derived more accurate crisp weights from pair-wise comparison matrices of criteria and sub-criteria than the *Fuzzy AHP* method. As for alternatives' ranking, we conclude that the *BIFMCDM* method reflects better the fuzzy evaluation of each alternative against the criteria set than the *Fuzzy TOPSIS* technique does. The ranking problem and the decision making process in general represents a very serious challenge in the *Web intelligence* field, as future works, we aim at studying how Regret can affect our future decisions, from a neurological point of view, and how can we apply such brain mechanism to invent more intelligent systems.

6. Acknowledgements

This research did not receive any grant or support from funding agencies or organizations in the public. governmental. commercial. or not-for-profit sectors.

7. References

[1] Atanassov, K. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 20*, 87–96.

[2] Atanassov, K., & Gargov, G. (1989). Interval valued intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 31*, 343–349.

[3] Ayağ, Z. (2005). A fuzzy AHP-based simulation approach to concept evaluation in a NPD environment. *IIE Transactions, 37*(9), 827-842. doi: 10.1080/07408170590969852

[4] Bloch, I., Maitre, H., Collin, B., Ealet, F., Garbay, C., Le Cadre, J., . . . Rombaut, M. (2003). *Fusion d'information en traitement du signal et des images*. Paris.

[5] Boender, C. G. E., de Graan, J. G., & Lootsma, F. A. (1989). Multi-criteria decision analysis with fuzzy pairwise comparisons. *Fuzzy Sets and Systems, 29*(2), 133-143. doi: 10.1016/0165-0114(89)90187-5

[6] Buckley, J. J. (1985). Fuzzy hierarchical analysis. *Fuzzy Sets and Systems, 17*(3), 233-247. doi: 10.1016/0165-0114(85)90090-9

[7] Büyüközkan, G., & Çifçi, G. (2012). A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry. *Expert Systems with Applications, 39*(3), 2341-2354. doi: 10.1016/j.eswa.2011.08.061

[8] Chang, D.-Y. (1996). Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research, 95*(3), 649-655. doi: 10.1016/0377-2217(95)00300-2

[9] Hwang, C. L., & Yoon, K. (1981). *Multiple Attribute Decision Making Methods and Applications: A State-of-the-Art Survey*. New York: Springer-Verlag.

[10] Kahraman, C. (2008). Multi-criteria decision making methods and fuzzy sets. In C. Kahraman (Ed.), *Fuzzy multi-criteria decision making: Theory and Applications with Recent Developments* (Vol. 16).

[11] Mikhailov, L. (2003). Deriving priorities from fuzzy pairwise comparison judgements. *Fuzzy Sets and Systems, 134*(3), 365-385. doi: 10.1016/s0165-0114(02)00383-4

[12] Naili, M., Boubetra, A., Tari, A., Bouguezza, Y., & Achroufene, A. (2015). Brain-inspired method for solving fuzzy multi-criteria decision making problems (BIFMCDM). *Expert Systems with Applications, 42*(4), 2173-2183. doi: 10.1016/j.eswa.2014.07.047

[13] Pawlak, Z. a. (1982). Rough sets. *International Journal of Computer & Information Sciences, 11*(5), 341-356. doi: 10.1007/bf01001956

[14] Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling, 9*(3-5), 161-176. doi: 10.1016/0270-0255(87)90473-8

[15] Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research, 48*(1), 9-26. doi: 10.1016/0377-2217(90)90057-i

[16] Van Laarhoven, P. J. M., & Pedrycz, W. (1983). A fuzzy extension of Saaty's priority theory. *Fuzzy Sets and Systems, 11*(1-3), 199-227. doi: 10.1016/s0165-0114(83)80082-7

[17] Zadeh, L. (1965). Fuzzy sets. *Inform. and Control, 8*, 338–353.

[18] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information Sciences, 8*(3), 199-249. doi: 10.1016/0020-0255(75)90036-5

[19] Zimmermann, H. J. (2010). Fuzzy set theory. *WIREs Computational Statistics, 2*. doi: 10.1002/wics.82

Facility Location for New Stone Crusher in Sarcheshmeh Copper Mine with AHP and TOSIS Methods

Ali Bakhshian[1], Nima Mirzaei[2]

[1]*Faculty of Economics and Administrative Sciences, Department of Business Management, Istanbul Aydin University, Florya, Istanbul, Turkey*

[2]*Department Industrial Engineering, Istanbul Aydin University, Florya, Istanbul, Turkey*

Abstract: This study offers an approach for dealing with prediction of the new stone crusher in Sarcheshmeh copper mine. The most important property of In-pit crusher is reducing the extraction costs. So it must be situated in a place that operating costs of conveyor, trucks and the number of trucks be at least. It is possible the various positions be better than others and the factors of site selection choice are not in a direction, because the organized and comprehensive study of possible places of In-pit crusher is necessary. Multi-criteria decision method is used for ranking of waste in-pit crusher site selection in Sarcheshmeh copper mine. The importance and sensitivity of waste rock is less than ores but size distribution and the ability of load and transport are the most important variables The main criteria in four main groups of technical factors, costing factors, operating factors and environmental factors. Affecting parameters on In-pit crushing site selection choice include. TOPSIS and AHP are used in this study as multi-criteria decision making method.In ranking of crusher site selection, the west case by 0.702 score in first rank, the south case by 0.613 as second rank, east case as third by 0.508 score and the east as fourth rank by 0.383 score was suggested.

Key words: multi-criteria decision making (MCDM), TOPSIS, AHP, Sarcheshmeh Copper mine;

1. Introduction:

Sarcheshmeh copper mine is one of the largest Oligo-Miocene Porphyry Cu deposits all around the world and also mineral processing operation has product huge volume of tailings materials in particular pyrite that it located in 65km southwest of Kerman city, with GPS details; southeastern Iran lat.29°56'30"N., long. 55°52'30" E most of the cu deposits are mining by open-pit [1], [2]Sarcheshmeh copper mine have got many types of mineral elements like: Au, Mo, Ni, Zn, Pb and Ni [3]. Granodiorite intrusive porphyry deposit originated on the name of the source porphyry is concentrated. This intrusive body is dependent on a set quartz diorite more important. That showed protrusion around the deposit.

The intrusive phases of mineralization cut by dikes at the same time or later and for most of the dikes around N25W. Mining and related industries are also can effecting on environment so managing of transporting and find the best way to bring them to process line of heavy metals and acid mine drainage generation are very important [4]. The environmental can be define like as public health and safety, social relationships, air and water quality and some various impacting factors from mining activities were estimated for environmental component [5]. The equipment for mining and engineering judgment have a lot of effect for select the new stone crusher in open- pit types of mine that the present work explores with TOPSIS and AHP method[6]. There are a lot of alternatives and criteria's for defining mining land in other hand multi criteria decision making MCDM methods can be useful [7].

2. The importance and need for research:

According to records of crusher and conveyor system inside the pit mines in the world produce the desired results, as well as increasing the depth of pit sarcheshmeh copper mine In the development plan, followed by rising costs and problems with using the current method, Studies in this field necessary for this mine.

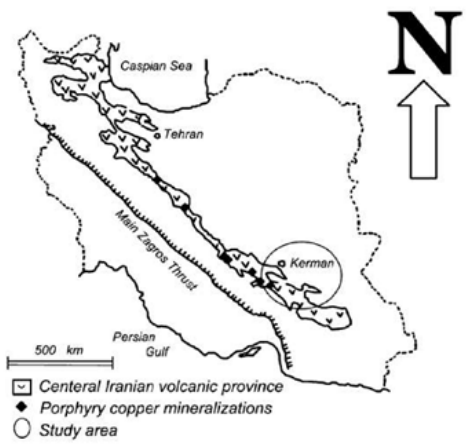


Figure 1: Geographic location of Sarcheshmeh copper complex

Mine facility location selection (MFLS) is one of the general problems that its gone happened to developing in open-pit mining planning and design. At the first step geographical information and potential of that zone should be define by geographical information system (GIS). In second step, find the alternative sites. Third step: determining environmental impact assessment (EIA) [8].

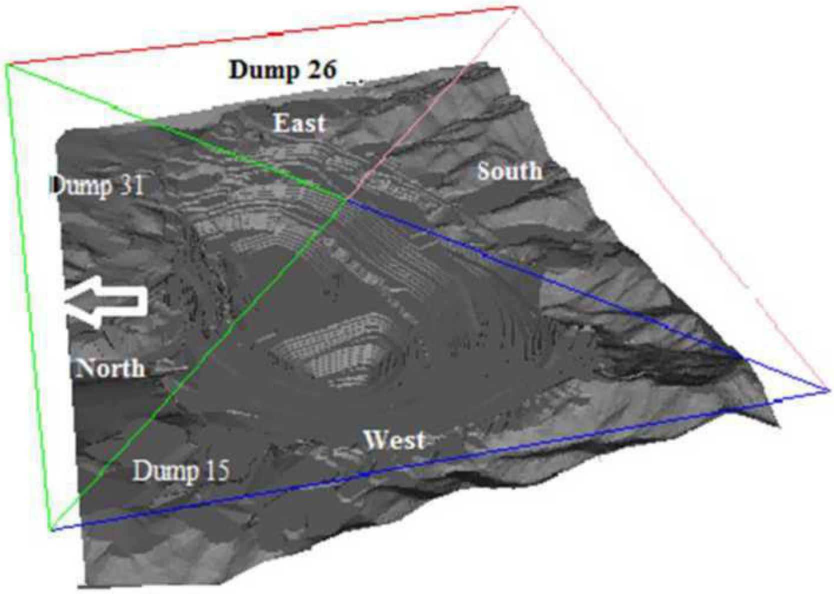


Figure 2: the mining field

Today, the increasing development of science and increase the volume of information and identify new sources, is no secret. In recent decades, environmental, political, economic, social and cultural, in the creation of new places and diversify the places and culture, has left deep effects. In such an environment where users are required to use appropriate methods and tools can no longer focused on traditional tools and methods. Location decision is a major issue for planning new development, specially within the planning of new healthcare infrastructure [9], [10].

MCDA is recognized as the right approach for location or supplier selection when both qualitative and quantitative factors are considered [11], [12]. There are many techniques available for solving MCDA problem. In a study, Multi criteria decision analysis (MCDA) examine and compare two modeling methods to decide about health care infrastructure location decision. Evidential Reasoning (ER) was used to solve the model and then the AHP was used to compare the results [9].

Analytic Hierarchy Process (AHP) was introduced by Saaty in the 1980. Later, the method was adopted in engineering, manufacturing, industry, education, management, and etc. [13].

In general, the AHP method is a popular and effective approach in DMA [14]—[18]. However, there are other methods such as TOPSIS, ELECTRE, and VIKOR which can be found for solving MCDA problem [19],[22]. These method can be combined with fuzzy approach (fuzzy AHP fuzzy TOPSIS) to solve and find solution for MCDM problem [23], [24]

Geographical information system GIS is one of the most comprehensive and most capable systems that serve various disciplines including management science and industrial sites are selected [25].

GIS suitability analysis can be applied for siting facilities such as healthcare, clinic, power plant, transportation stations, bioenergy location, retail site and etc. [21], [26],[29]

The process of selecting the best location is often difficult. The proposed methods can be generally divided into five groups:

multi-criteria decision-making methods, Fuzzy MCDM methods, Geographical Information System (GIS), Multi-Criteria Decision Making with GIS integration, and Select models of the plant [28].

Multi-criteria decision making occurs in different situations, such as facility location selection , transportation and logistics service [24], [30], distribution center [31], supplier selection [32], maintenance strategy [33], [34], landfill for industrial waste and waste disposal site [35], [37], environmental risk management [38], equipment selection [39], dump site selection [40],

There are many types of multi criteria decision making methods that the most common methods its analytical hierarchy process (AHP) [41], techniques for order preference by similarity to an ideal solution(TOPSIS) [42], Simple additive weighting[43], the compromise ranking method (VIKOR) [19], the preference ranking organization method for enrichment evaluation (PROMETHEE) [44].

Some example for Facility locations with different methods ;An assessment of site suitability for marina construction in Istanbul, Turkey, using GIS and AHP multi criteria decision analysis [45].

3. TOPSIS method

This method developed in 1981 by Hwang and Yoon and it also developed by Chen and Hwangin1992.TOPSIS method works with two different side that it called positive and negative side[46][47].

Positive side: in this side we try to make the maximize the benefit and minimize the cost criteria. Negative side: the goal is the maximize the cost criteria and minimize the benefit criteria .

4. The TOPSIS process:

Step l.Need to create evaluation matrix include of m alternatives and n criteria with intersection of each alternative and criteria like x_{ij}. So we have a matrix $(x_{ij})_{mxn}$.

Step 2. Normalize the decision matrix:

$$R=(rij)m \times n \quad (1)$$

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, i = 1,2,...,m, j = 1,2,...,n \quad (2)$$

Step 3. Calculate the weighted normalized decision matrix

$$t_{ij} = r_{ij} w_j , i = 1,2,..., m, j = 1,2,..., n \tag{3}$$

where $w_j = W_j / \sum_{j=1}^{n} W_j , j = 1,2,..., n$ then $\sum_{j=1}^{n} w_j = 1$ and W_j is the original weight given to indicator

$$V_j = j = 1,2,..., n \tag{4}$$

Step 4:

It's time to determine the worst alternative (A_w) and determine the best alternatives (A_b) :

$$A_w=\{(\max(t_{ij}|i=1,2,\ldots,m|j\in J^-),\min(t_{ij}|i=1,2,\ldots,m|j\in J^+)\}\equiv\{t_{wj}|j=1,2,\ldots,n\} \tag{5}$$

$$A_b=\{(\min(t_{ij}|i=1,2,\ldots,m|j\in J^-),\max(t_{ij}|i=1,2,\ldots,m|j\in J^+)\}\equiv\{t_{bj}|j=1,2,\ldots,n\} \tag{6}$$

Step 5:

Calculate the distance between the alternative ***i*** and the worst A_w

$$d_{iw} = \sqrt{\sum_{j=1}^{n} (t_{ij} - t_{wj})^2} , i = 1,2,..., m \tag{7}$$

$$d_{ib} = \sqrt{\sum_{j=1}^{n} (t_{ij} - t_{wj})^2} , i = 1,2,..., m \tag{8}$$

Step 6

Calculate the similarity

$$s_{iw} = d_{iw} /(d_{iw} + d_{ib}), 0 \le 1, i = 1,2,..., m \tag{9}$$

$s_{iw}=1$ in this point if and only if the alternative solution has the best condition

$s_{iw}=0$ in this point if and only if the alternative solution has the worst condition.

Step 7:

Rank the alternatives with pay attention to $s_{iw}(i=1,2,..., m)$

5. AHP

Analytical Hierarchy Process (AHP) is technique for solving problems with complex multi criteria that is developed by Thomas Saaty in the 1970s is one of multi criteria decision making methods.[50]

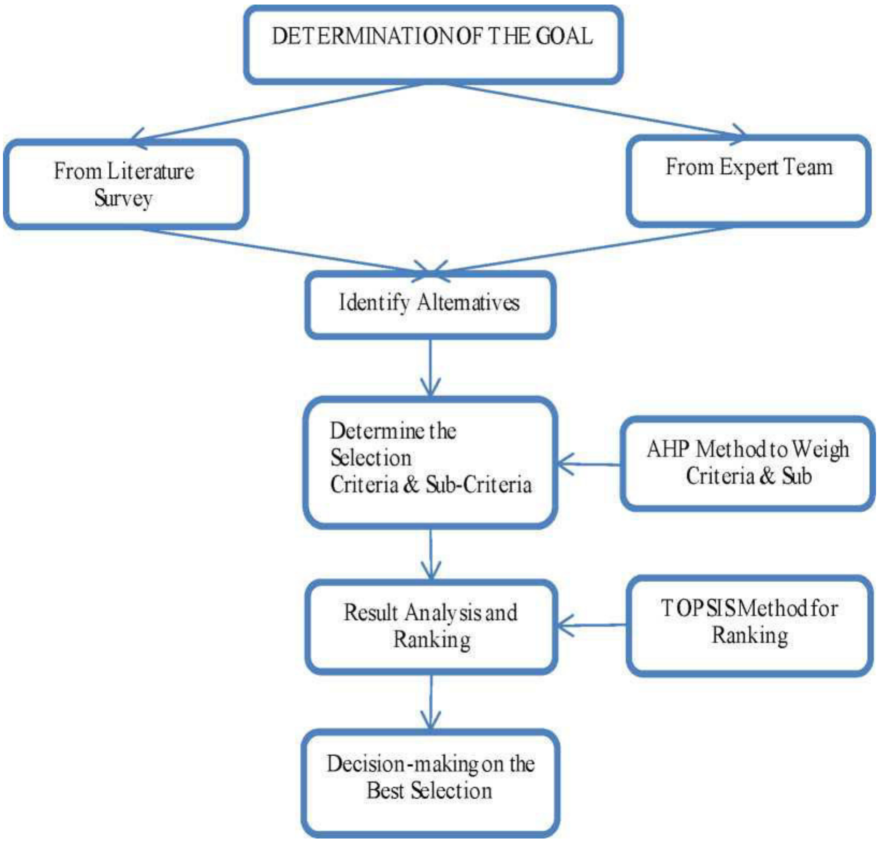


Figure 3: MCDM structure

The application of the AHP to the multi criteria decision making problem involves four major steps:

1. Identify the problem and break down it to level with factors and factors to structure them in hierarchical table
2. Construct pair wise comparison matrix by evaluating factors through scale from 1–9
3. Use normalization method to evaluate the weight table of relevant factors
4. Perform consistency analyses for the problem

AHP is a flexible MCDM method that deals with quantitative and qualitative approach for the problem. The decision maker has an ease to see overall relationship of factors in a hierarchical table. The main steps of AHP process is decomposition of problem into a hierarchical table

He includes main goal, critical and alternatives. In hierarchical part of problem the correlation of alternatives are constructed. In the case when second comparison has higher impact than the first one the pair wise comparison table shows it as 1 divided by 9. Figure 3 illustrate the scheme of AHP.

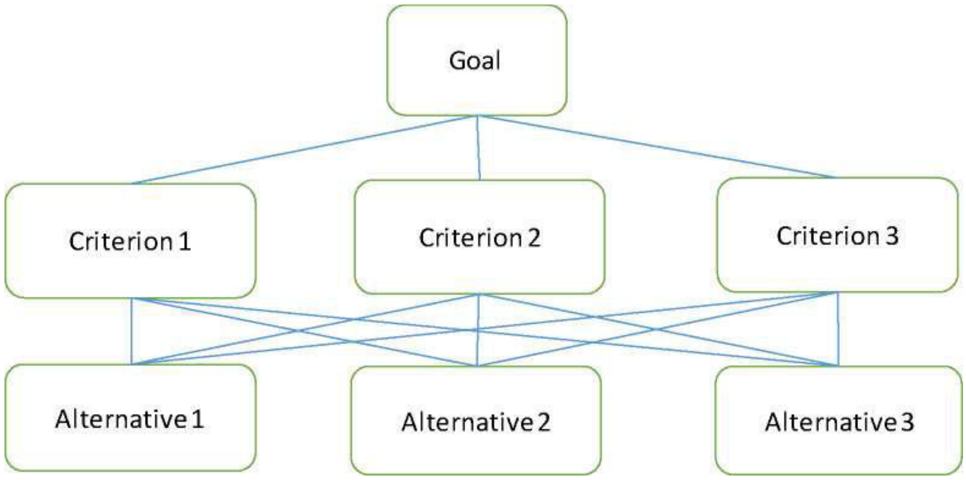


Figure 4: Illustrate the scheme of AHP

Hierarchies are divided into two categories as follows:

1. Structural hierarchy: In a hierarchical structure elements are generally physically associated with each other.
2. The hierarchical task: In a hierarchical task for credit or functional components linked together to form a system the day. Each task hierarchy is made up of a series of levels that exist at the highest level horizon of an element called that goal, but in later levels there may be more elements. The number of these elements can't be high and usually between 5 and 9 elements.

Sometimes measure must also be analyzed so in this part we have new level, this type of hierarchical, call multi-level hierarchy complete.

In this context there are recommendations that will help to build a hierarchy include:

1. To determine the ultimate goal and that of the hierarchy and most important question is what purpose are you pursuing?
2. Determine the purpose of minor
3. Specify the measures that are effective in minor goals.
4. Determine criteria and sub criteria carefully as possible under the numerical criteria for distance or for little like high, medium and low.
5. Factors and sub-factors and the next levels.
6. Determination of options and exits.
7. In some cases the best option for the two hierarchical one another for profit and is useful for a fee. The benefit-cost ratio accounted for the largest proportion of its option is selected. Also in some cases to decide yes - if this method can also be used.
8. The decision of the two modes (yes or no) options can be considered for performance or nonperformance.

Weight in AHP is calculated in two ways A) Calculate the relative weight

To calculate the weight of analytic hierarchy, each level of your element at a higher level than in pair-wise comparison matrix is formed and compared. Assigning a numerical score of either or both of the paired indexes based on Table 3 - 3 takes place.

A paired comparison matrix is shown as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{22} \end{bmatrix} \tag{10}$$

Where in image Preference element image I am of the element image I was. Paired comparison criteria are established in relation to one another as follows:

$$a_{11} = \frac{1}{a_{ji}} \tag{11}$$

Table 1: Allocating points to the importance of standards

Numerical Rating	Compare the relative indices (oral judgment)
9	The absolute importance
7	The importance of a strong
5	Strong importance
3	Poor care
1	Equally important
2, 4, 6 and 8	Preferences between these intervals

Paired comparison matrix $n \times n$ in which n is the number of elements that were compared is. For each pair-wise comparison matrix $n \times n$, elements on the diagonal of the matrix is equal to a need to assess, but other matrix elements should be determined on the basis of paired comparison. Elements are reversed with respect to the original diameter. The number of paired comparisons for each pair-wise comparison matrix $n \times n$ equal to:

$$N_h = \frac{n(n-1)}{2} \tag{12}$$

In general, if the decision-making m Options n Criteria should be n Pair-wise comparison matrix $m \times m$ and a pair-wise comparison matrix $n \times n$ are created. Therefore, comparing the number of hierarchy (theproblem) is with:

$$N_h = \frac{n(n-1)}{2} + n \times \frac{m(m-1)}{2} \tag{13}$$

After determining the matrix of pairwise comparison, the relative weight of elements is calculated. There are different methods to calculate the relative weight based on paired comparison matrix that matter most to the least squares method, method of least squares logarithmic, of eigenvector and measures approximately of them. Among these methods are, especially vector method is more accurate.

B) The method of least squares: the next discussion will be told that if the matrix A consistent, we have:

$$a_{ij} = \frac{W_i}{W_j} \Rightarrow W_i = a_{ij} W_j \tag{14}$$

But in practice it rarely happens that the matrix is consistent. The least squares method is tried values $W_{i,} W_j$ be determined in such a way that the sum of the squares of the differences and is minimized. In other words, the following linear programming problem solved:

$$Min \quad Z = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} W_j - W_i)^2 \tag{15a}$$

$$s.t: \sum_{i=1}^{n} W_i = 1, \quad W_i \geq 0, \quad i = 1,2,...,n \tag{15b}$$

This problem can be solved from the Lagrange multipliers method:

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} W_j - W_i)^2 + 2\lambda (\sum_{i=1}^{n} W_i - 1) \tag{16}$$

We derivative W_i of the above equation to:

$$\sum_{i=1}^{n} (a_{il} W_l - W_i) a_{il} - \sum_{i=1}^{n} (a_{il} W_l - W_i) a_{il} + \lambda = 0 \quad l = 0,2,...,n \tag{17}$$

High non-homogeneous equations, linear equations and unknowns can be obtained. For example, we have, for example: n=2

$$(a_{11}^2 + a_{21}^2 - 2a_{11} + 2)W_1 - (a_{12} + a_{21})W_2 + \lambda = 0 \tag{18a}$$

$$-(a_{12} + a_{21})W_1 + (a_{12}^2 + a_{22}^2 - 2a_{12} + 2)W_2 + \lambda = 0 \tag{18b}$$

$$W_1 + W_2 = 1 \qquad (18b)$$

For n = 3 we have

$$(a_{11}^2 + a_{21}^2 + a_{31}^2 - 2a_{21} + 3)W_1 - (a_{12} + a_{21})W_2 - (a_{13} + a_{31})W_3 + \lambda = 0 \tag{19a}$$

$$-(a_{12} + a_{21})W_1 + (a_{12}^2 + a_{22}^2 + a_{32}^2 - 2a_{22} + 3)W_2 - (a_{23} + a_{32})W_3 + \lambda = 0 \tag{19b}$$

$$-(a_{13}+a_{31})W_1+(a_{23}+a_{32})W_2-(a_{13}^2+a_{23}^2+a_{33}^2-2a_{33}+3)W_3+\lambda=0 \tag{19c}$$

$$W_1+W_2+W_2=1 \tag{19d}$$

These equations can be solved variables W1, W2 and λ calculate:

C) Logarithmic least squares method: As already mentioned, the compatibility or incompatibility of the order states that:

$$a_{ij}=\frac{W_i}{W_j}\Rightarrow a_{ij}\frac{W_i}{W_j}=1 \tag{20}$$

$$a_{ij}\neq\frac{W_i}{W_j}\Rightarrow a_{ij}\frac{W_i}{W_j}\neq 1 \tag{21}$$

In the method of least squares logarithmic attempt to capture the result of multiplying the differences are minimal differences in other words, the geometric mean is minimized. Disputes geometric mean is:

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\frac{W_i}{W_j}\right)^{\frac{1}{n^2}}=Z^{\frac{1}{n^2}} \tag{22}$$

In other words, states have compatibility or incompatibility respectively:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left(Ln\ a_{ij}-Ln\left(\frac{W_i}{W_j}\right)\right)=0 \tag{23}$$

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(Ln\ a_{ij}-Ln\left(\frac{W_i}{W_j}\right)\right)=\frac{1}{n^2}Ln\ Z \tag{24}$$

Since the brackets may be negative in some cases and in some cases is positive, it will be two heads to always be positive. Therefore, this procedure must solve the following linear problem to values W_i achieved:

In general, ordinary least squares method, mean geometric mean log errors and least squares method minimize errors.

$$Min\ Z=\sum_{i=1}^{n}\sum_{j=1}^{n}\left(Ln\ a_{ij}-Ln\left(\frac{W_i}{W_j}\right)\right)^2=0 \tag{25a}$$

$$s.t:\sum_{i=1}^{n}W_i=1,\quad W_i\geq 0,\quad i=1,2,\ldots,n \tag{25b}$$

D) The specific vector method: the method of calculating the relative weight, especially vector method is more accurate. In this way, Are designed to be established that the following equation (Saaty, 1980):

$$AW = \lambda W \tag{26}$$

Where in λ And W respectively paired comparison matrix eigenvalues and eigenvectors is. In the case of larger matrix size is too time consuming to calculate these values. Thus, for calculating λ the determinant of the matrix A- λI placed equal to zero by providing the largest amount λ the resulting values in the following equation W it will be counted.

$$(A - \lambda_{\max} I) \times W = 0 \tag{27}$$

Theorem: For a positive and inverse matrix (such as pair-wise comparison matrix) Special vector can be calculated from the following equation:

$$W = \lim_{k \to \infty} \frac{A^k e}{e^T A^k e} \tag{28}$$

$$\text{Where it is } e^T = [1 \quad 1 \quad \ldots \quad 1] \tag{29}$$

The final weight of each item in a hierarchical process of the product multiplied by the weight of each criterion score achieved the desired option. The total score obtained for each option the relationship below is obtained:

$$A_{AHP(SCORE)} = \sum_{i=1}^{n} a_{ij} W_{ij}, \quad i = 1,2,\ldots,n \tag{30}$$

Where in a_{ij} indicate the relative importance Options i for index C_j And W_j show the importance of j. Also it is necessary to value of alternatives and weights using the following relations are normalized.

$$\sum_{i=1}^{n} a_{ij} = 1, \quad i = 1,2,\ldots,n \tag{31}$$

$$\sum_{i=1}^{n} W_j = 1, \quad i = 1,2,\ldots,n \tag{32}$$

If n criteria as C_1, C_2, …, C_n they have paired comparison matrix is as follows:

$$A = \lfloor a_{ij} \rfloor, \quad i = 1,2,\ldots,n \tag{33}$$

Where in a_{ij} Preferred measure C_i On C_j, The matrix A consistent 40 in the matrix if we say:

$$a_{ik} \times a_{kj} = a_{ij} \quad i, j = 1,2,\ldots,n \tag{34a}$$

If the matrix A is compatible with:

$$a_{ij} = \frac{W_i}{W_j} \tag{34b}$$

Weight matrices compatible through normalization of columns can be achieved.

If an incompatible Pair-wise comparison matrix is, how much of the conflict and how much of it. Inconsistency is better to measure the expression of several important cases should be mentioned:

For each pair-wise comparison matrix (positive and negative), if λ_1, λ_2 ,… and Eigenvalues of matrix pair-wise comparison, the total eigenvalues of the *n* (the matrix) will be:

$$n = \sum_{i=1}^{n} \lambda_i \tag{35}$$

The maximum eigenvalues (λ_{max}) always greater than or equal to (in the case of some λ will be negative).

If a small amount of matrix elements compatibility mode distance, its eigenvalues is also a

small amount of the adjustment will be.

On the other hand, by definition matrix for each square of , we have:

$$A \times W = \lambda W \tag{36}$$

Where in *W* and λ respectively eigenvector and eigenvalues of matrix , respectively. In the case of matrix *A* consistent, a special amount equal to *n* (the maximum eigenvalues) and the rest are zero. So in this case, can write:

$$A \times W = nW \tag{37}$$

In the case of pair-wise comparison matrix *A* is incompatible, image *A* little bit of n distances that can be recited:

$$A \times W = \lambda_{\max} W \tag{38}$$

Since the (λ_{max}) is always greater than or equal to *n*. If distance matrix bit compatibility mode (λ_{max}) of n will be a little distance. So the difference λ_{max} and (λ_{max}-*n*) *A* good measure of the size of the matrix will be the incompatibility. Undoubtedly scale (λ_{max}-*n*) the value of *n* (the matrix) dependent and this dependence can scale to meet the following way define it mismatch index (*I.I.*) they say.

$$I.I. = \frac{\lambda_{\max} - n}{n-1} \tag{39}$$

The inconsistency index for the numbers of random matrices are available, has calculated that it random incompatibility index(*R.I.I.*) Named are the values for the matrix n second on the following equation or table 3. 4 calculated

$$R.I.I. = 1.98\frac{n-2}{n} \tag{40}$$

Table 3-4: random inconsistency index

N	16	17	18	19	20	21	22	23	24	25
RI	1.59	1.6	1.62	1.63	1.63	1.64	1.65	1.65	1.66	1.66
N	26	27	28	29	30	31	32	33	34	35
RI	1.67	1.67	1.67	1.68	1.68	1.68	1.69	1.69	1.69	1.7

$$I.R.. = \frac{I.I.}{R.I.I.} \tag{41}$$

Inconsistencies in the method of calculating the rate of AHP is also of high importance. In general, an acceptable level of incompatibility of a system depends on the decision maker,

But the Saaty number as an acceptable 1.0 provides and believes that if the mismatch is greater than 1.0, it is better judgments appeal.

6. Conclusion

First of all, we collect the data about each part of Sarcheshmeh Copper mine land. We visited all part of that land and try to find the best zone for build a new stone crusher with multi criteria decision making (MCDM) with two different methods that we already explain the methodology. Criteria based on calculation by engineers and discuss with different experts at that mine, they also estimate the weight of the criteria for mineral and dump site. Making the matrix for dump site and mineral site its part of those two different methods that we do already and after calculation we find the place for dump and mineral site. First of all lets talk about TOPSIS and TOPSIS result. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method for multiple criteria choice in a dump machine facility placement problem was assigned.Obtained from previous result the mining area in south were filled for new stone crusher machine, leading it to be excluded from optimal region.The problem was containing 21 factor.Since some of them were containing negative effect the calculation were performed separately.

For finding the place for stone crusher and our dump site we need to pass 7 steps

1. Form a matrix
2. Normalization
3. Build weight matrix
4. Define ideal and non-ideal alternative
5. Calculate similarity for the worst condition 7. Rank the result

The conclusion of calculation is alternative

Result for dump site: A1 (North area of mining)>A3 (South area of mining)>A4 (East are of mining)>A2 (West area of mining)

Alternative 1	**C1***	**0.73340248**	**BEST**
Alternative 2	**C2**	**0.167965932**	**4th**
Alternative 3	**C3**	**0.330342791**	**2nd**
Alternative 4	**C4**	**0.314153619**	**3rd**

Result for mineral (ore) site shows us: A2 (West area of mining)>A3 (South area of mining)>A4 (East area of mining)>A1 (North area of mining)

Alternative 1	**C1**	**0.383909666**
Alternative 2	**C2***	**0.702204626**
Alternative 3	**C3**	**0.613710831**
Alternative 4	**C4**	**0.508297258**

AHP is one of the frequently used multi criteria decision making methods. Through assignment of factors and carefully judged sub factors for the multi decision problem the alternatives the decision maker can achieve optimum alternative. It's a method for solving complex decision making based on the alternatives and multi criteria ,as it names stated. It is also a process for developing a numerical score to rank each decision alternative based on how well each alternative meets the decision maker's criteria.

In this method we need to do some steps:

Phase 1: Develop rating for each decision criteria

Step 1: Defining the decision problem

The weight for criteria for multiple choice system problems must be defined.

Step 2: The Pair-Wise comparison Step 3: Synthesizing the Judgment Step 4: Consistency Analyses for criteria

Phase 2: Develop the rating for each decision alternative for each criterion

Step 1: Defining the Decision Problem

Step 2: The Pair-Wise Comparison

Step 3: Synthesizing the Judgment

Phase 3: Develop rating for overall hierarchy

Step 1: Overall Priority vector

Step 2: Overall Consistency Ratio

You can also follow each step in this study, the resultant weight, provide that A3>A2>A4>A1.

FINAL WEIGHT	Area
0.208905639	A1
0.245189832	A2
0.357476263	A3*
0.209072494	A4
1.020644228	
0.357476263	

We have to pass the same steps for finding the best place for stone crusher

FINAL WEIGHT	Area
0.338368767	A1
0.396314466	A2
0.651098676	A3
0.464157784	A4
1.849939693	
0.651098676	

7. Reference

[1] M. Khorasanipour, M. H. Tangestani, R. Naseh, and H. Hajmohammadi, "Hydrochemistry, mineralogy and chemical fractionation of mine and processing wastes associated with porphyry copper mines: A case study from the Sarcheshmeh mine, SE Iran," *Appl. Geochemistry,* vol. 26, no. 5, pp. 714-730, 2011.

[2] S. Jannesar Malakooti, S. Z. Shafaei Tonkaboni, M. Noaparast, F. Doulati Ardejani, and R. Naseh, "Characterisation of the Sarcheshmeh copper mine tailings, Kerman province, southeast of Iran," *Environ. Earth Sci.,* vol. 71, no. 5, pp. 2267–2291, 2014.

[3] M. Einali, S. Alirezaei, and F. Zaccarini, "Chemistry of magmatic and alteration minerals in the Chahfiruzeh porphyry copper deposit, south Iran: Implications for the evolution of the magmas and physicochemical conditions of the ore fluids," *Turkish J. Earth Sci.,* vol. 23, no. 2, pp. 147–165, 2014.

[4] A. Ebrahimabadi and I. Alavi, "Plant Type Selection for Reclamation of Sarcheshmeh Copper Mine Using Fuzzy-Topsis Approach / Wybor Gatunkow Roslin Do Wykorzystania W Projekcie Rekultywacji Terenow Kopalni Miedzi Sarcheshmeh Z Wykorzystaniem Metod Logiki Rozmytej Topsis," *Arch. Min. Sci.,* vol. 58, no. 3, pp. 953–968, 2013.

[5] M. Monjezi, K. Shahriar, H. Dehghani, and F. Samimi Namin, "Environmental impact assessment of open pit mining in Iran," *Environ. Geol.,* vol. 58, no. 1, pp. 205-216, 2009.

[6] M. J. Rahimdel and M. Karamoozian, "Fuzzy TOPSIS method to primary crusher selection for Golegohar Iron Mine (Iran)," *J. Cent. South Univ.,* vol. 21, no. 11, pp. 43524359, 2014.

[7] A. H. Bangian, M. Ataei, A. Sayadi, and A. Gholinejad, "Optimizing post-mining land use for pit area in open-pit mining using fuzzy decision making method," *Int. J. Environ. Sci. Technol.,* vol. 9, no. 4, pp. 613–628, 2012.

[8] M. Fazeli and M. Osanloo, "Mine Facility Location Selection in Open-Pit Mines Based on a New Multistep-Procedure," in *Mine Planning and Equipment Selection: Proceedings of the 22ndMPES Conference, Dresden, Germany, 14th -- 19th October 2013,* C. Drebenstedt and R. Singhal, Eds. Cham: Springer International Publishing, 2014, pp. 1347-1360.

[9] B. Dehe and D. Bamford, "Development, test and comparison of two Multiple Criteria Decision Analysis (MCDA) models: A case of healthcare infrastructure location," *Expert Syst. Appl.,* vol. 42, no. 19, pp. 6717–6727, 2015.
[10] C. T. Lin and M. C. Tsai, "Location choice for direct foreign investment in new hospitals in China by using ANP and TOPSIS," *Qual. Quant.,* vol. 44, no. 2, pp. 375–390, 2010.
[11] C. Ram, G. Montibeller, and A. Morton, "Extending the use of scenario planning and MCDA for the evaluation of strategic options," *J. Oper. Res. Soc.,* vol. 62, no. 5, pp. 817–829, 2011.
[12] D. Golmohammadi and M. Mellat-Parast, "Developing a grey-based decision-making model for supplier selection," *Int. J. Prod. Econ.,* vol. 137, no. 2, pp. 191–200, 2012.
[13] K. Q. Zhou, A. M. Zain, and L. P. Mo, "A decomposition algorithm of fuzzy Petri net using an index function and incidence matrix," *Expert Syst. Appl.,* vol. 42, no. 8, pp. 3980–3990, 2015.
[14] F. T. Bozbura, A. Beskese, and C. Kahraman, "Prioritization of human capital measurement indicators using fuzzy AHP," *Expert Syst. Appl.,* vol. 32, no. 4, pp. 1100–1112, 2007.
[15] Y. M. Chen and P. Huang, "Bi-negotiation integrated AHP in suppliers selection," *Int. J. Oper. Prod. Manag.,* vol. 27, no. 11, 2007.
[16] H. Y. Kang and A. H. I. Lee, "Priority mix planning for semiconductor fabrication by fuzzy AHP ranking," *Expert Syst. Appl.,* vol. 32, no. 2, pp. 560–570, 2007.
[17] F. Y. Partovi, "An analytical model of process choice in the chemical industry," *Int. J. Prod. Econ.,* vol. 105, no. 1, pp. 213–227, 2007.
[18] S. K. Jakhar and M. K. Barua, "An integrated model of supply chain performance evaluation and decision-making using structural equation modelling and fuzzy AHP," *Prod. Plan. Control,* vol. 25, no. 11, pp. 938–957, 2013.
[19] H. Liao and Z. Xu, "A VIKOR-based method for hesitant fuzzy multi-criteria decision making," *Fuzzy Optim. Decis. Mak.,* vol. 12, no. 4, pp. 373–392, 2013.
[20] P. Wang, Z. Zhu, and Y. Wang, "A novel hybrid MCDM model combining the SAW, TOPSIS and GRA methods based on experimental design," *Inf. Sci. (Ny).,* vol. 345, pp. 27–45, 2016.
[21] S. Beheshtifar and A. Alimoahmmadi, "A multiobjective optimization approach for location-allocation of clinics," *Int. Trans. Oper. Res.,* vol. 22, no. 2, pp. 313–328, 2015.
[22] K. Devi and S. P. Yadav, "A multicriteria intuitionistic fuzzy group decision making for plant location selection with ELECTRE method," *Int. J. Adv. Manuf. Technol.,* vol. 66, no. 9-12, pp. 1219–1229, 2013.
[23] J. M. Sánchez-Lozano, M. S. García-Cascales, and M. T. Lamata, "Evaluation of suitable locations for the installation of solar thermoelectric power plants," *Comput. Ind. Eng.,* vol. 87, pp. 343–355, 2015.
[24] E. Celik, M. Erdogan, and A. T. Gumus, "An extended fuzzy TOPSIS???GRA method based on different separation measures for green logistics service provider selection," *Int.*
[25] *J. Environ. Sci. Technol.,* vol. 13, no. 5, pp. 1377–1392, 2016.
[26] S. R. Graham, C. Carlton, D. Gaede, and B. Jamison, "the Benefits of Using Geographic Information Systems As a Community Assessment Tool.," *Public Health Rep.,* vol. 126, no. April, pp. 298–303, 2011.
[27] J. C. García–Palomares, J. Gutiérrez, and M. Latorre, "Optimizing the location of stations in bike–sharing programs: A GIS approach," *Appl. Geogr.,* vol. 35, no. 1–2, pp. 235–246, 2012.
[28] D. Khan and S. R. Samadder, "A simplified multi-criteria evaluation model for landfill site ranking and selection based on AHP and GIS," *J. Environ. Eng. Landsc. Manag.,* vol. 23, no. 4, pp. 267–278, 2015.
[29] L. Panichelli and E. Gnansounou, "GIS-based approach for defining bioenergy facilities location: A case study in Northern Spain based on marginal delivery costs and resources competition between facilities," *Biomass and Bioenergy,* vol. 32, no. 4, pp. 289–300, 2008.
[30] R. Suárez-Vega, D. R. Santos-Peñate, and P. Dorta-González, "Location models and GIS tools for retail site location," *Appl. Geogr.,* vol. 35, no. 1–2, pp. 12–22, 2012.
[31] J. Zak and S. W^glinski, "The selection of the logistics center location based on MCDM/A methodology," *Transp. Res. Procedía,* vol. 3, no. July, pp. 555–564, 2014.
[32] R. Chakraborty, A. Ray, and P. K. Dan, "Multi criteria decision making methods for location selection of distribution centers," *Int. J. Ind. Eng. Comput.,* vol. 4, no. 4, pp. 491504, 2013.

[33] W. Ho, X. Xu, and P. K. Dey, "Multi-criteria decision making approaches for supplier evaluation and selection: A literature review," *Eur. J. Oper. Res.,* vol. 202, no. 1, pp. 1624, 2010.

[34] T. S. Baines, H. W. Lightfoot, O. Benedettini, and J. M. Kay, "Article information :," *J. Manuf. Technol. Manag.,* vol. 20, no. 5, pp. 547–567, 2009.

[35] B. Kirubakaran and M. Ilangkumaran, "Selection of optimum maintenance strategy based on FAHP integrated with GRA???TOPSIS," *Ann. Oper. Res.,* vol. 245, no. 1–2, pp. 285313, 2016.

[36] H. Mohamed, B. Omar, T. Abdessadek, and A. Tarik, "An application of OLAP/GIS- Fuzzy AHP-TOPSIS methodology for decision making: Location selection for landfill of industrial wastes as a case study," *KSCE J. Civ. Eng.,* vol. 0, no. 0, pp. 1–11, 2016.

[37] U. Yildirim and C. Güler, "Identification of suitable future municipal solid waste disposal sites for the Metropolitan Mersin (SE Turkey) using AHP and GIS techniques," *Environ. Earth Sci.,* vol. 75, no. 2, pp. 1–16, 2016.

[38] G. P. Yal and H. Akg??n, "Landfill site selection utilizing TOPSIS methodology and clay liner geotechnical characterization: A case study for Ankara, Turkey," *Bull. Eng. Geol. Environ.,* vol. 73, no. 2, pp. 369–388, 2014.

[39] Z. Wen, S. Ma, S. Zheng, Y. Zhang, and Y. Liang, "Assessment of environmental risk for red mud storage facility in China: a case study in Shandong Province," *Environ. Sci.*

[40] *Pollut. Res.,* vol. 23, no. 11, pp. 11193–11208, 2016.

[41] M. Dagdeviren, "Decision making in equipment selection: An integrated approach with AHP and PROMETHEE," *J. Intell. Manuf*, vol. 19, no. 4, pp. 397–406, 2008.

[42] M. Yasser, K. Jahangir, and A. Mohmmad, "Earth dam site selection using the analytic hierarchy process (AHP): A case study in the west of Iran," *Arab. J. Geosci.,* vol. 6, no. 9, pp.3417–3426, 2013.

[43] Saaty T.L., *The Analytic Hierarchy Process.* New York: McGraw-Hill, 1980.

[44] *K. Hwang, Ching-Lai, Yoon,* Multiple Attribute Decision Making – Methods and Applications. A State-of-the-Art Survey. *Springer Berlin Heidelberg, 1981.*

[45] *K. R. MacCrimmon,* Decisionmaking among multiple-attribute alternatives: a survey and consolidated approach. *RAND CRP, 1968.*

[46] J. P. Brans and P. Vincke, "Note—A Preference Ranking Organisation Method," *Manage. Sci.,* vol. 31, no. 6, pp. 647–656, 1985.

[47] M. U. Gumusay, G. Koseoglu, and T. Bakirman, "An assessment of site suitability for marina construction in Istanbul, Turkey, using GIS and AHP multicriteria decision analysis," *Environ. Monit. Assess.,* vol. 188, no. 12, 2016.

[48] G. R. Jahanshahloo, F. H. Lotfi, and M. Izadikhah, "Extension of the TOPSIS method for decision-making problems with fuzzy data," *Appl. Math. Comput.,* vol. 181, no. 2, pp. 1544–1551, 2006.

[49] J. Jiang, Y. W. Chen, Y. W. Chen, and K. W. Yang, "TOPSIS with fuzzy belief structure for group belief multiple criteria decision making," *Expert Syst. Appl.,* vol. 38, no. 8, pp. 9400–9406, 2011.

ATENA Methodology and Selected System Requirements for Communication and Knowledge Management Tools

Radosław ZAJĄC[1]

[1]*PW Warsaw University of Technology, Faculty of Management*
Narbutta 85, Warsaw, Poland, e-mail: Radoslaw.zajac@pw.edu.pl

Abstract: Article discusses relations between Atena methodology and system requirements for tools used in given organization for communication and knowledge management. Apart from Application Theme – explicitly dealing about IT systems, each of the rest of principles, themes or processes can be translated into system requirements. Atena methodology and requirements come from existing systems introduced based on previous versions of communication management strategy. Both interrelated areas: theoretical approach and systems preceded with certain set of requirements were refined based parallel business applications and scientific discussions. This is the essence and primary strength of Atena methodology, application and implementation oriented set of best practices extended to some related theoretical models important for creating an d maintaining efficient communication and knowledge management environment.

Keywords: knowledge management, communication management, information systems, requirements, business analysis

1. INTRODUCTION

Turbulent business and social environment forces individuals, teams and organizations to base on schemes and clichés. This implies either general shallowness or specialism. Regarding average managers job seniority at one position of about two years and expectation of quick results, there is not much space left for deliberation. This superficiality becomes result and reason of many other malfunctions. Especially in domain of communication and knowledge management. Newly hired try to achieve quick wins to proof their value, and the best way to do it is to find any failures in existing organization or to be innovative. Lack of understanding of details – both subtle and vital for organization leads to many disservices. Analogously outgoing worker is often cause of loss of tacit knowledge because of no “outboarding” process. Financial sector, or more general – service sector – is lacking because of frequent changes referring either to products or service procedures.

A solution for many of those problems is a general methodology for communication and knowledge management called Atena – presented in this article. This high level tool supports analysis, implementation and maintenance of business communication and knowledge management systems. Whereas this term refers to joint organizational culture, individuals, processes and IT systems consistent business environment.

“Theoretical” models such as PRINCE2 or other IT-related frameworks are goal-oriented and come from practice. Whereas models referring to knowledge management concentrate on theoretical aspects (Walczak, 2005) or specific issues. Interesting exception comes from Tseng (Tseng, 2008), who provides overview of Knowledge Management System implementation from organizational strategy to post-implementation indexes, but only referring to performance management. More comprehensive approach comes from Wiig (Wiig, 1995) – unfortunately despite discussing knowledge creation, there is no connection with communication process and direct reference to system requirements. Although the widespread perspective is impressive there might be some more guidelines for implementation execution.

What is lacking in existing frameworks is a comprehensive approach anchoring knowledge management in business and process environment regarding various domains of the merit. Second missing aspect is embedding knowledge management regarding real lifecycles (from organization to individual). And the last area to be explored and covered with guidelines is transformation knowledge management into IT practice – in the layer of project management and low-level requirements for IT applications.

In this article this third area is discussed – some two-way interrelations between Atena methodology and requirements for applications supporting business system for communication and knowledge management (BSCKM) are pointed out in order. This tight interdependence justifies significance of such methodology as science and real business tool. In theoretical frameworks and real implementation some of those interfaces are respectively missing or not fully managed, and that’s the origin of many negative business consequences such as vanishing knowledge management or knowledge management tools or losing business potential coming from already gathered knowledge, experience and information.

2. ATENA METHODOLOGY AND SYSTEM REQUIREMENTS

Atena methodology (Zając, 2016a, Gołuch-Trojanek, 2017) deals primarily about a few values:

1. Establishing structure, ownership and responsibility for communicate (content, distribution and use) as well as ownership and responsibility for communication and knowledge processes – i.a. by means of principles. Whereas communicate is an atomic object of communication process.
2. Making all users of business system communication and knowledge management prosumers, by transforming them into co-creators of content as well as information and knowledge flows.
3. Harmonization and synchronization with issues connected with: trainings/learning, operational communication and innovations.
4. Establishing hierarchy that makes IT applications tools for creating consistent environment for communication and knowledge management with primacy of prosumers and their needs over given systems and system requirements. In this approach a combination of specialized tools (like CMS, DMS, Quiz creator and separated forum) might be more effective than all-in-one system. It is vital especially if some of those applications had been introduced before and members of the organization were accustomed to them. This approach allows to left aside some of common transformational problems – like learning new tools and changing one's habits and schemes.

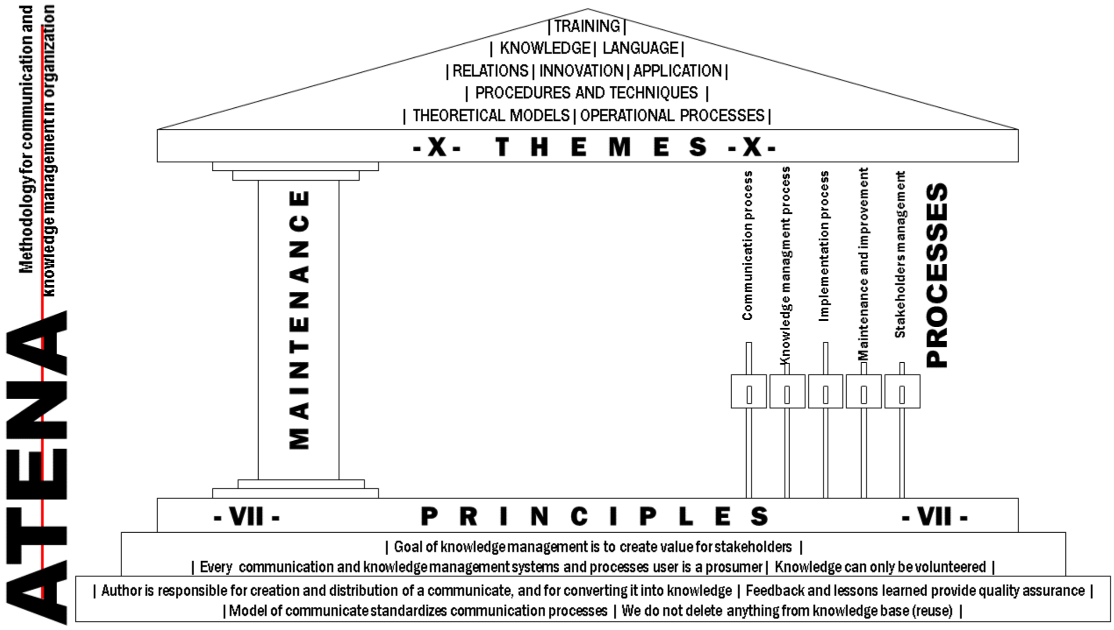


Figure 1: Structure of Atena methodology

Because of size of the article only elements of methodology and its implementation will be discussed. Of course Application Theme is the most obvious area to be investigated while talking about IT systems, but as far as implementation provenience is concerned – each of Atena elements may be translated into functional or non-functional requirements for system supporting communication and knowledge management domain.

As stated before one of fundamentals of Atena is a communicate – meant as logical, formal and IT object. Therefore the following Table one refers directly to formal system requirements.

Communicate should be parametrized template for entering and presenting given set of information. Communicate parameters should contain the following subset. (Req. 1)

Table 1: Communicate structure

Problem	Informal draft of title, easy to understand and deal for Authors and Editors. Might be a duplicate of Title.
Title	Official title of a communicate shown for every user.
Communication Type	Drop down list, i.a. article, note, alert, news, blog article.
Recipients	Drop down list of group of users.
Summary	Dual use field: 1. A few sentence introduction to the communicate and 2. Field shown as text or text and graphic teaser placed in slider/banner in the starting page of a portal for communication and knowledge management or shared with another applications.
Content (with hyperlinks) + guidelines	Content of a communicate using full HTML formatting and might contain hyperlinks. This field should be accompanied with prompts containing guidelines of use of Bloom's taxonomy in creating a communicate. Especially in cognitive layer – transforming communicate into knowledge and enabling analysis, synthesis, assessment and application.
Authors/editors remarks	Field that allows to document discussion about given communicate (especially as a part of editing workflow) – in order to avoid parallel communication among Authors and Editors via mail.
Changes	Dual use field: 1. Highlight most important changes induced by the communicate and 2. In order to force Author to formulate explicit statement about most important facts regarding former organizational state. This field standardizes construction of communicate and makes easier acquiring knowledge and using it in practice by readers.
Question	Optional field enabling connection between communication created as a result of helpdesk or FAQ module questions with given questions.
Expected date of publication	Field that allows to point suggested publication date.
Expiration date	Date by which communicate will be accessible for all authorized users.
Category	Multiple choice field allowing for defining content categories on the basis of merit (like given process or specific product categories) as well as technical (like archive and pending). One communicate may belong to various – non-contradictory categories. Category affects communicate template for editing and display.
Labels (tags, keywords)	Multiple choice field enabling to display one communicate in various places of the portal or even in different IT tools being part of BSCKM. Label/tag logic enable to tailor access to content to individual or group cognitive schemes better than traditional wiki structures. The same mechanism allows to group communicates referring to the same or related topic to ease and accelerate process of acquiring knowledge – and reflect adult learning need of context presentation.
Communicate ID	Automatically created communicate identifier.
Author	Automatically assigned Author and Editors working on given communicate.
Creation date	Automatically assigned date of creation of given communicate.
Last modification date	Automatically assigned dates of subsequent modifications of the given communicate (for the purpose of identification, reporting and editing).
Publishing date	Field for entering publication date for given communicate. With reminder of necessity of time shift enabling users to acquire knowledge and aske potential questions to state the information precisely.
Breadcrumb	Automatically created field enabling reader to identify localization of given article in structure of application supporting BSCKM – as well as switch to other level from the path using one click.
Layout	Parameter reflecting given display template assigned to communicate type.
Change history	Automatically recorded subsequent versions of communicate modifications.
Related communicates	Field enabling to assign and/or display related communicates linked with labels/tags or manual forced connections.

References	Optional field to highlight certain groups of related communicates or documents, i.a. law statements or scientific papers
Comments	Vital field to present user comments – important for engaging prosumers as well as reducing number of similar questions for stating the information precisely.
Ratings	Field for indicating overall rating for the given communicate – it should be parametrized to reuse the same mechanism either for rating communicates or innovative ideas being refined in innovation tool.
Link to forum	Optional link enabling to broaden knowledge by reading forum discussions related to given topic (it can be done for example by using tags)

In a way those parameters and use of social media tools is nothing new – what is innovative and creates value for organization is structuring those tools and objects and introducing overall governance in order to synchronize and harmonize various social media functionalities and standards.

It is required to provide two-way communication making users of BSCKM being able to react for given communicates in order to make them real co-creators of communication and knowledge flow.

It is required to enable authorized users to rate, comment and ask questions about communicates and articles. It is required to implement two-way communication tools such as active FAQ, forum, newsgroup, quizzes, knowledge pills, blogs. (Req. 2)

Because of their complexity, two of Atena methodology components do not refer directly to system requirements – although they base on sets of implemented requirements and IT governance both in implementation and further use in business as usual. These are – Processes and Maintenance – and few requirements derived from them..

2.1 Processes

There are only five high level processes in Atena methodology (Zając, 2017a) – in order to focus on main aspects referring to communication and knowledge management governance: 1. Communication process, 2. Knowledge management process, 3. Knowledge management implementation process, 4. Maintenance and improvement process, 5. Stakeholders management process. Those processes are elaborated by Zając (Zając, 2017a) and because of their high level construction and the fact that they do not directly affect system requirements, they are not the matter of this article.

In Atena methodology process means only a high level one. Middle- or low level processes are called procedures, so they belong to one of Themes. That is why processes it may seem that they have no direct transposition onto system requirements. Despite that, communication process forces implementation of workflow for editing communicates:

It is required to provide workflow for creating, editing, publishing and archiving communicates. (Req. 3)

From process point of view communicate template should address following aspects of communication:

It is required to support planning time shift needed for communicate to be received, analyzed and transformed into both knowledge and action. (part of Req. 1 referring to publication date)

For both – communication and knowledge management processes application support is crucial at least for communicate distribution and retention, as well as for search for information.

Tree structure (information architecture), It is required to preserve all communicates (information) in applications being part of BSCKM (requirement derived directly from Principle about reuse). (Req. 4)

It is required that application should enable various ways to find and access communicates and information – at least by:

1. Tree structure (information architecture),
2. Tags (one communicate shown in various places of the communication and knowledge management portal),
3. Search engine,
4. Desktops for certain user groups. (Req. 5)

It is required to create and customize self-checking quizzes (presenting proper answers with justification – for knowledge pills and stand-alone quizzes). (Req. 6)

Knowledge management process referring to revealing and recording tacit knowledge leads to following:

It is required to enable multiple direction communication and knowledge flows by means of: forum, discussion list, active FAQ and comments. (Req. 7 – detailed version of Req. 2).

As a consequence, the same social media mechanisms support various communication and knowledge management guidelines.

Process of implementation of knowledge management is tightly connected with BSCKM as analog to Agile Manifesto, especially: Individuals and interactions over processes and tools and Customer collaboration over contract negotiation. That implies general requirement: to assure primacy of individuals, interactions and prosumer collaboration over systems and system requirements.

Maintenance and perfecting knowledge management process consist of a few sub-processes. This quality approach implies some of the requirements.

It is required to monitor and report user activities using given KPI's for whole BSCKM – such as number of system and user errors, number of complaints. (Req. 8)

It is required to provide context help and performance support in various product application – with content provided by communication and knowledge management business system (for example as hyperlink or webservice-based transfer). (Req. 9)

2.2 Maintenance

This component reflects importance of governing communication and knowledge management with related areas in order to change whole business environment at the levels of organizational culture, processes, applications and individual attitude. Not less important is planning and monitoring BSCKM after implementation and passing it into operation.

Alike processes, Maintenance component do not directly affect systems requirements.

Maintenance component, being on the top of the Atena methodology structure, has no direct interconnections with IT requirements as a combination of Principles, Themes and Processes. Nevertheless without atomic preceding process and system requirements forming multi-level structure for supporting and establishing communication and knowledge management – this component at least would not work efficiently.

Examples of such requirements are as following.

It is required to authenticate and authorize application users using SSO (Single Sign On) mechanisms whenever it is possible from technical and economical point of view. (Req. 10)

And one more generic requirement about governance: enterprise architects should guard simplicity and balance between number of applications and their value for business and impact on core business value creation.

Other to components – Principles and Themes refer to some requirements that are vital to implementation at the execution level.

2.3 Principles

There are eight axioms called Principles that span background for communication and knowledge management methodology. Those axioms mainly base on communication and knowledge management organizational failures – and are formulated in contradiction to those malfunctions. Implementing them in organizational business reality is trigger whether Atena methodology is introduced as a whole.

Table 2: Chosen Principles in Atena methodology

Principle	Short description	Requirements
Goal of communication and knowledge management is to create value for stakeholders	Communication and knowledge management is a tool and vital part of management. It cannot be devaluated by "corpo" language. Lack of trust for communicates makes knowledge management impossible.	This principle refers to general requirement about communication and knowledge management governance, compliance with IT governance and management as well as compatibility with organizational goal structure – from strategic to operational ones.
Every communication and knowledge management systems and processes user is a prosumer	Users of BSCKM are co-creators of information and knowledge because of their knowledge (also tacit) and experience. Only involving them by promoting to peers and empowering them make them engaged.	***(Req. 2)***
Author is responsible for creation and distribution of a communicate, and for converting it into knowledge	Quality cannot be assured without ownership and responsibility for processes, stages and processes' outputs. That is why responsibility cannot be passed on to communicates users.	This principle is connected to (1) two way communication process and IT tools as well as to (2) editing workflow. Respectively to (1) enable prosumer to assess communicates and quickly ask clarifying questions and (2) to manage the communication process from creating to distribution stage. ***(Req. 2)***
We do not delete anything from knowledge base (reuse)	Organizational memory is short because of dynamic changes within organization and in surrounding business environment. Once created analysis and made decisions are assets with certain cost. They need to be preserved to be reused in order to accelerate further business activities and provide continuous access to lessons learned.	***It is required to leave rights for deleting content or files only for business administrators or part of editing team. (Req. 11)*** ***Search engine and tag mechanism should support finding similar and connected content. (Req. 12)***

2.4 Themes

Themes refer to areas interconnected to or being part of communication and knowledge management. These are issues coming from good practices that must be governed in order to avoid problems in introducing and maintaining BSCKM.

Table 3: Themes in Atena methodology. (Elaborated in details, Zając, 2017b)

Theme	Requirements
Operational Management Theme (OMT)	Operational management Theme affects IT domain in a few areas, at least: mails, attachments. ***It is required to provide DMS (Document Management System) tool to enable exchanging mail attachments to links for knowledge base or document repository. (Req. 13)*** Remark: The same mechanism prevents from meaningless pushing lots of MB's causing LAN overload and prevents prosumers from using various versions of the same document at the same time – what happens when files are attachments.
Training Theme (TT)	Training Theme deals about synchronizing and harmonizing communication and knowledge management domain with knowledge transfer from training channel. ***It is required to interlink LMS (Learning Management System) in process or application layer from the level of applications supporting BSCKM. Application being part of this business system should comply with microlearning and allow prosumers to self-improve their knowledge and check progress (this is to be done by quiz functionalities, like in Coursera.org or Duolingo). (Req. 14)***

Applications Theme (AT)	The purpose of this Theme is to ensure proper level of continuous concentration on IT governance, architecture, processes and applications in order to harmonize evolution of whole communication and knowledge environment. This two-way interdependence is described partially in ITIL and COBIT – in terms of value creation or directly by knowledge as an asset. – in terms of value creation or directly by knowledge as an asset. Generic requirement: All the changes in IT domain should consider impact on information, data and knowledge governance and management.
Procedures and Techniques Theme (PTT)	***It is required to provide workflows for: communication process (esp. editing and distribution), access rights management meant as authorization, authentication and assigning prosumer groups to given communicates. (Req. 15). It is required to provide workflow for revision and updating application being part of BSCKM. (Req. 16). It is required for applications to support communicates flow between internal corporation portal, knowledge base, external website and tools for creating and refining innovations. (Req. 17)***
Innovations Theme (InT)	***It is required for tools for creating and refining innovations to be process- and layout-complying with other communication and knowledge management tools. For example is should be integrated at the level of sign in (SSO – single sign on mechanism) and knowledge exchange (at least links for knowledge base, optimally direct references to communicates from the same domain). (Req. 18)***
Knowledge Theme (KT)	Knowledge Theme is connected among others to sources of knowledge – for example business intelligence, compound data and information about processes, process output [Zając, 2017b], and i.a. to Bloom's taxonomy (Krathwohl, 2002). ***It is required to present the information tailored to group needs (managerial or positions desktops and reports).*** ***It is required for IT elements of BSCKM to semi-automatically or automatically support providing complete and up-to-date set of documents, i.a. law regulation, scientific publications. (Req. 19)***
Theoretical Models Theme (TMT)	This Theme refers partially to the business intelligence aspect from Knowledge Theme (benchmark and news feed). But some of theoretical models can be directly translated into system requirements. ***It is required to provide reminder mechanisms for preserving knowledge (Ebbinghaus, forgetting curve) and quizzes for self-assessment while learning and revising knowledge in planned subsequent intervals. (Req. 20)*** ***It is required to provide triggers and reports for using established KPI's (i.a. from Kirkpatrick model, Grisaffe, 2007)). (Req. 8)***
Language Theme (LT)	***It is required to support use of applications determining FOG Index[1] as well as language corps in order to refine communicates for better conversion into practice. (Req. 21)*** ***It is required to provide search engine and tool for conducting repeating check of communicate and everyday use language compliance. (Req. 22)*** It is also important to implement context approach to communication (Sundgren, 1973).
Relations Theme (RT)	Relations Theme force establishing communication strategy meant as schedule, channels and formats of communicates and reports to provide either continuously or in specified intervals.

[1]FOG Index refers to readability in context of readers education and complexity of language (http://gunning-fog-index.com/)

3. CASE STUDY

Short case study will be presented to provide feel and touch about Atena methodology transformation into system requirements for tools supporting BSCKM.

Let us start from common organizational approach with communication based on e-mail cascade in any given financial corporation. In dynamic business environment with rapidly changing products and processes this means a “Chinese whisper” way of conducting business. In which mid-level managers are responsible not only for their teams performance, but also distribution of the communication, clarification and encouraging team members to put communicates into practice. Clarification questions engage management and professional specialists. Questions often are multiplied because of similar experience staff members. But regarding question formulation differentiation and multiple professionals answering, questions about the same issue may vary. As a consequence one question may engage a few professionals and result in few different answers. In spite of appearances this is a real case because of many product and customer needs combinations. Two of the most difficult issues from communication and knowledge management domain are – profound and fast-paced transformation to be managed and not often realized processes or seldom offered products being on board not to long for being naturally internalized but as long, that they are not most important for managers.

Implementing subsequent transformational changes and IT tools e.g. such as Microsoft Sharepoint or Liferay in order to reengineer customer service and sales processes should help to get rid of most of problems mentioned above. Exemplary three stages are sketched beneath:

FIRST STAGE consist i.a. of assigning e-mail channels for certain kind of communication content such as operational management and notifications. This must be connected with information architecture creation and introducing first version of knowledge base or at least file repository. Those changes as a whole let the organization to distribute links to documents instead of attachments in order to prevent from use of various versions of documents, free space on mail servers and lower synchronous LAN traffic.

SECOND STAGE consist i.a. of implementing information architecture fully in the knowledge base and little by little moving knowledge related product and process communicates flow from e-mail into a knowledge base or other portal solution. This allows to serve communicate in structure defined as in Table 1. Users of BSCKM may become prosumers by rating communicates, asking questions and making comments to clarify communicates and answer others’ questions placed in comments. Encouraged by subtle continuous and coherent organizational culture modification, prosumers will evolutionary form a community, self-evolving and self-advancing.

If clarification questions or discussions are important – original communicate should be replaced by versions regarding the state of knowledge including prosumers professional knowledge and experience. This refinement upgrades quality of single communicate and improves prosumers’ engagement at the same time.

They can also find related content based on tags/labels in order to support their self-development.

Forum mechanism should be introduced to involve most active (usually hater-kind) users to concentrate attention and involvement in one tool or portal.

THIRD STAGE – various informational sections (common errors, most often read) and discussion list as active FAQs providing one question-one answer order should be introduced. Reporting tools such as PIWIK should be used to analyze prosumers journey through portal and provide more efficient portal structure and informational sections. Multimedia content such as knowledge pills consisting of text, short “1”-minute films and quizzes for self-assessment should be widely introduced. Using interfaces or hyperlinks enable to provide context help for other IT tools –

without replication of information. Knowledge base or BSCKM becomes partially-hidden layer connecting various applications (such as product systems or internal and external websites).

Such transformation will free some managers and professional specialists time and result in communicate and knowledge standardization. Of course it requires transformation of corporate culture and individual habits at the same time. Instead of phone calls staff members should be encouraged to use portal – not only when the knowledge is needed but also proactive – in order to get the lessons learned load, to learn from others mistakes.

Regarding first difficult issue – BSCKM will provide processes and IT tools for preserving time and space for efficient conversion information coming from communicates into knowledge and practical use. And in the case of seldom activities will support individuals by easy, instant access to professional community and to needed content.

4. SUMMARY

As shown in the article, elements of Atena methodology come from practice and form consistent structure of generalized guidelines for practitioners and researchers. That is why Atena can be treated as a record of good practices, a framework that allows to organize existing and future specialized tools into a BSCKM to achieve synergy and avoid contradictory impact of various tools and methods. In addition, article provides some specific examples of problems and solutions brought by implementation of BSCKM.

From theoretical point of view Atena methodology may be used to investigate gaps in communication and knowledge management domain – especially when multiple tools are used, they can be assigned to certain Atena components and analyzed in context of their interdependence. From implementation and business point of view, Atena is a methodological framework for managing projects and teams in this domain – assuring complete view on one hand and specific practical hints – on the other.

Next IT related steps refer to creating suite of tools for support communication and knowledge management systems in various organizations – using some of most common systems, like MS Sharepoint, Liferay and Jira/Confluence. Concerning other domains, existing frameworks will be revised and arranged in order to support Atena methodology concept or provide specialized tools such as Tseng’s performance index.

5. REFERENCES

[1] Gołuch-Trojanek, K., Rostek, K., Zając, R. (2017). Wpływ metod tworzenia i wdrażania systemów informatycznych na zarządzanie komunikacją i wiedzą w organizacji, Referat wygłoszony na konferencji: Paradygmaty współczesnego zarządzania organizacjami. Maj, 2017.

[2] Grisaffe, D. B. (2007). Questions about the ultimate question: conceptual considerations in evaluating Reichheld's net promoter score (NPS). Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior, 20, 36.

[3] Kirkpatrick, D.L., Kirkpatrick, J.D. (2006). Evaluating training programs: The four levels. San Francisco, CA: Berrett-Koehler.

[4] Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. Theory into practice, 41(4).

[5] Snowden, D. (2009), Dave Snowden's 7 Principles of Knowledge Management (2009), retrieved from: http://www.gurteen.com/gurteen/gurteen.nsf/id/km-seven-principles.
[6] Sundgren, B. (1973). An Infological Approach to Data Bases. Stockholm: Skriftserie Statistika Centralbyran.
[7] Tseng, Shu-Mei (2008). Knowledge management system performance measure index. Expert Systems with Applications 34.1 (2008): 734-745.
[8] Walczak, Steven (2005). Organizational knowledge management structure. The Learning Organization 12.4 (2005): pp. 330–339.
[9] Wiig, Karl M. (1995). Knowledge management methods. Schema Press. Arlington (TX).
[10] Zając, R. (2016a). The concept of methodology for communication and knowledge management in the organizations. Jagiellonian Journal of Management, 2(Numer 4).
[11] Zając, R. (2016b). Metodyka Atena zarządzania komunikacją i wiedza w organizacji – powiązania. Referat wygłoszony na konferencji: Doktorant – X edycja, Warszawa.
[12] Zając, R. (2017a). Metodyka Atena zarządzania komunikacją i wiedza w organizacji – procesy. Referat wygłoszony na konferencji: Doktorant – X edycja, Kraków.
[13] Zając, R. (2017b). Overview of themes in methodology Atena for communication and knowledge management in organization, paper from Doctoral Consortium, BIS Conference, Poznań, June 2017.

Application of Classical Forecasting Approaches to Predict the Share of Faulty Components

Krzysztof JURCZYK[1], Magdalena KRZYWDA[1], Jerzy MIKULIK[1]

[1]*AGH University of Science and Technology, Faculty of Management*
Mickiewicza 30, Krakow, Poland, e-mail: kjurczyk@zarz.agh.edu.pl

Abstract: The aim of this paper was to compare and select the most appropriate method for forecasting the share of faulty components in the total production. Analysis and calculations were performed based on real data from a manufacturing company. In total three forecasting methods were used and compared, namely: naïve, moving average and exponential smoothing. At the end the generated forecasts were compared based on selected criteria.

Keywords: error analysis, reliability, forecasting

1. INTRODUCTION

There are many various definitions of terms such as forecast and forecasting. Accordingly to A. Zeliaś a "forecast" can be defined as: "a choice – made within given arrangement – of the most likely way how given economic phenomena will evolve in the upcoming future, where the basis for such choice lays in the past phenomena behavior and the current state of the arrangement" [Zeliaś 1997, pp. 15]. Based on M. Cieślak we can say that forecast is a like deducing about the future based on past events [Cieślak 1999, pp. 15]. However, the biggest simplification in terms of defining the concept of forecasting can be found in work written by C. Bozarth and R. B. Handfield

(2008), where forecast is defined as a rough estimation of given variable value. The forecasting process itself is usually realized based on the outline presented in the next paragraph.

During the first stage of the forecasting process one has to formulate the forecasting task which will encompass the forecasted phenomena as well as the forecasts' objectives. The next stage focusses on analyzing the factors which influence the changes in considered phenomena. The third stage aims at gathering statistical data as well as its analysis and processing. In the fourth stage one has to select one of the available forecasting methods. In the last but not least stage the forecasting model is being built upon selected method. In the sixth stage the obtained results are being verified whether they are satisfying or not. If the forecast accuracy is acceptable it can be used in practice. The last stage focusses on assessing the accuracy of the forecasting method [Dittmann, pp. 22–35].

2. FORECASTING METHODS

There is a multitude of forecasting methods which can be applied for various classes and types of time series. To the most fundamental/basic one can account: naïve, moving average and exponential smoothing methods.

Naïve models, accordingly to M. Cieślak [Cieślak 1999, pp. 67–70] are based on the assumption that none significant changes in the level of the forecasted value will be observed (Eq. 1):

$$y_t^* = y_{t-1} \tag{1}$$

where:

y_t^* – forecasted value of variable y estimated for period t,

y_{t-1} – value of variable y in period $t–1$.

In case of time series which exhibit seasonal patterns (seasonality) it is suggested to use method described by Eq. 2:

$$y_t^* = y_{t-r} \tag{2}$$

where:

y_{t-r} – value of variable y in corresponding phase of cycle from previous period,

r – number of cycle phases.

The second group of forecasting models consists of ARIMA and SARIMA models, which has been in a very detailed way described by G. E. P. Box and G. M. Jenkins (1970). These models are based upon auto regression (AR) and moving average (MA).

The auto regressive model in its most general form can be expressed by following equation [Dittmann 2008, pp. 98]:

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + ... + \varphi_p y_{t-p} + \varepsilon_t \tag{3}$$

where:

$y_t, y_{t-1}, y_{t-2}, ..., y_{t-p}$ – value of forecasted variable in period *t*, *t–1*, *t–2*, …, *t–p*,

$\varphi_0, \varphi_1, \varphi_2, ..., \varphi_p$ – model parameters,

ε_t – model error in period *t*,

p – drift.

The formulation of AR models is based on a dependence between forecasted variable and its values from previous periods. The moving average model (MA) can be expressed as follows [Cieślak 1999, pp. 95]:

$$y_t = \vartheta_0 - \vartheta_1 \varepsilon_{t-1} - \vartheta_2 \varepsilon_{t-2} - ... - \vartheta_q \varepsilon_{t-q} + \varepsilon_t \tag{4}$$

where:

y_t – value of forecasted variable in period *t*,

$\vartheta_0, \vartheta_1, \vartheta_2, ..., \vartheta_p$ – model parameters,

$\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$ – model errors in period *t*, *t–1*, *t–2*, …, *t–p*,

q – drift.

In case of MA models it is important to underline that the sum of model parameters $\sum_{i=1}^{q} \vartheta_{t-i}$ does not have to be equal 1 and additionally those parameters does not have to be positive [Cieślak 1999, pp. 95]. By combining AR and MA models one receives an ARMA model [Dittmann et al. 2009, pp. 120]:

$$y_t = \varphi_1 y_{t-1} + ... + \varphi_p y_{t-p} + \varepsilon_t + \vartheta_0 - \vartheta_1 \varepsilon_{t-1} - ... - \vartheta_q \varepsilon_{t-q} \tag{5}$$

Discussed AR, MA and ARMA models assume the stationarity of the forecasted variable time series. In a situation when considered time series does not exhibit stationarity it is recommended to transform it to such by means of differential approach [Dittmann et al. 2009, pp. 239]. The differentiating processes can be expressed by means of Eq. 6 and 7 whereas the Eq. 7 is used for subsequent differences [Dittmann et al. 2009, pp. 120].

$$\nabla y_t = y_t - y_{t-1} \tag{6}$$

$$\nabla^d y_t = \nabla^d y_t - \nabla^{d-1} y_{t-1} \tag{7}$$

As a results of above described procedure it is possible to forecast given time series by means of AR, MA and ARMA models. Models in which case a differentiating approach has been applied are known as integrated processes: ARI (Autoregressive Integrated – integrated process of auto regression), IMA (Integrated Moving Average – integrated process of moving average). Integrated models can be presented in general form as ARIMA (*p*, *q*, *d*) where *p* is the level of auto regression, *d* is the level of differentiating and the *q* is the drift of the moving average [Zeliaś et al. 2003, pp. 239]. The ARIMA (*p*, *q*, *d*) process can be expressed by following formula [Zeliaś et al. 2003, pp. 240]:

$$\Phi_p(B)\nabla^d y_t = \Theta_q(B)\varepsilon_t \tag{8}$$

where *B*, is a backward shift operator which can be calculated based on Eq. 9:

$$B\varepsilon_t = \varepsilon_{t-1} \tag{9}$$

When sesonality in given time series is observed the ARIMA model has to be transformed into SARIMA (Seasonal Autoregressive Integrated Moving Average) model. The general form of SARIMA model is as follows:

$$\Phi_p(B^S)\nabla_S^d y_t = \Theta_q(B^S)\varepsilon_t \tag{10}$$

As shown by A. Lichota [Lichota 2006, s. 57] in this model we assume that the value of the forecasted variable in period *t* is dependent on its past values as well as the errors of the past forecasts and those values in periods *t–s*. In work presented by A. Lichota we can also find a detailed description when such method can be applied [Lichota 2006, pp. 57–58]. In general the autoregressive model should be used when the number of partial correlation coefficients is very small. In opposite case a moving average model should be applied. The ARMA model is applied when the values of auto regression coefficients as well as the partial auto regression coefficients exponentially decrease. The integrated processes can be used when the considered time series exhibits non-stationarity.

The next group of statistical forecasting methods is based on smoothing the time series based on weighted moving average. Those methods has been in detail described by works of R. G. Brown (1959), C. C. Holt (1957) and R. R. Winter (1960).

The Brown model is a form of simple smoothing which can be described by a following equation:

$$y_t^{'} = \alpha y_t + (1-\alpha)y_{t-1}^{'} \tag{11}$$

where:

$y_t^{'}$, $y_{t-1}^{'}$ – exponentially smoothed mean trend values in periods *t* and *t–1*,

y_t – the last observed value of forecasted variable,

α – smoothing constant $\langle 0;1 \rangle$

In case of this model it is important to set the initial value of the forecasted variable in period *t*. In general it is assumed to be equal to that observed in reality:

$$y_1^{'} = y_1 \tag{12}$$

It is worth to highlight that the longer the time series the smaller is the impact of the initial value on the forecast. This model can be applied to time series which are characterized by stationarity, with an almost constant level of forecasted variable as well as some random fluctuations [Sobczyk 2008, pp. 122].

Holt's double exponential smoothing, as noted by M. Sobczyk can be applied when the time series consists of a component in form of a linear trend and a relatively small seasonal fluctuations [Sobczyk 2008, pp. 127]. This model can be used to build short term forecasts for one or several periods ahead. By increasing the forecasts horizon one should expect greater forecasting errors. Holt's models are described by following equations:

$$F_t = \alpha y_t + (1-\alpha)(F_{t-1} + S_{t-1}) \tag{13}$$

$$S_t = \beta(F_t - F_{t-1}) + (1-\beta)S_{t-1} \tag{14}$$

$$y_t^{'} = F_{t-1} + S_{t-1} \tag{15}$$

where:

F_t , F_{t-1} – calculated exponentially means for periods *t* and *t–1*,

S_t , S_{t-1} – mean trend changes calculated based on periods *t* and *t–1*,

α, β – smoothing constants $\langle 0; 1 \rangle$

y_t – the last observed value of forecasted variable.

This model uses two smoothing constants α and β. Their values are selected experimentally in such way which minimizes the mean squared error of ex-post forecasts [Sobczyk 2008, pp. 128]. In case of this model it is important to estimate its initial values F_1 and S_1 . Usually it is assumed that F_1 will be equal to y_1 whereas S_1 is assumed to be equal 0 or as difference between $y_2 - y_1$.

The Winter's model is used when in the time series apart from trend and random fluctuations a seasonal changes can be observed [Sobczyk 2008, pp. 130]. In general we distinguish two versions of Winter's models. An additive and multiplicative one. The additive model can be described by following equations:

$$F_t = \alpha(y_t - C_{t-r}) + (1-\alpha)(F_{t-1} + S_{t-1}) \tag{16}$$

$$S_t = \beta(F_t - F_{t-1}) + (1-\beta)S_{t-1} \tag{17}$$

$$C_t = \gamma(y_t - F_t) + (1-\gamma)C_{t-r} \tag{18}$$

$$y_t^{'} = F_n + S_n(t-n) + C_{t-r} \tag{19}$$

Whereas the multiplicative model is described as follows:

$$F_t = \alpha(y_t / C_{t-r}) + (1-\alpha)(F_{t-1} + S_{t-1}) \tag{20}$$

$$S_t = \beta(F_t - F_{t-1}) + (1-\beta)S_{t-1} \tag{21}$$

$$C_t = \gamma(y_t / F_t) + (1-\gamma)C_{t-r} \tag{22}$$

$$y_t^{'} = [F_n + S_n(t-n)]C_{t-r} \tag{23}$$

where:

F_t , F_{t-1} – calculated exponentially means for periods *t* and *t–1*,

S_t , S_{t-1} – mean trend changes calculated based on periods *t* and *t–1*,

C_t – seasonality index,

α, β, γ – smoothing constants $\langle 0;1 \rangle$,

y_t – the last observed value of forecasted variable.

This model uses three smoothing constants α , β and γ . Their values – similarly as in Holt's model case – are estimated experimentally in order to minimize the mean squared error of ex-post forecasts. The α parameters has impact on the assessment of smoothed variable value in period *t,* β shows the trend direction for period *t,* γ reflects the seasonal fluctuations. As earlier it is important to estimate the initial values F_1 , S_1 and C_1 , C_2 , …, C_r . Usually F_1 is set as y_1 and S_1 as 0 or $y_2 - y_1$. Values of C_1 , C_2 , …, C_r are set in such a way that their sum is equal to 1.

3. FORECASTING THE SHARE OF FAULTY PRODUCTION

Described in the previous section forecasting methods have been used to estimate the share of faulty production in total production in given manufacturing company. Figure 1 visualizes date used in this study.

To create forecasting models we have used a data set covering 66 working days. As can be seen on Figure 1 the share of faulty production exhibits a decreasing trend. The highest share of faulty production has been observed during fifth day of our analysis whereas the lowest during the 57th day. However, the majority of observations oscillates around the mean. This observation is confirmed by presented on Figure 1 histogram and the lattice of coincidence. The values of basic statistical parameters are given in Table 1.

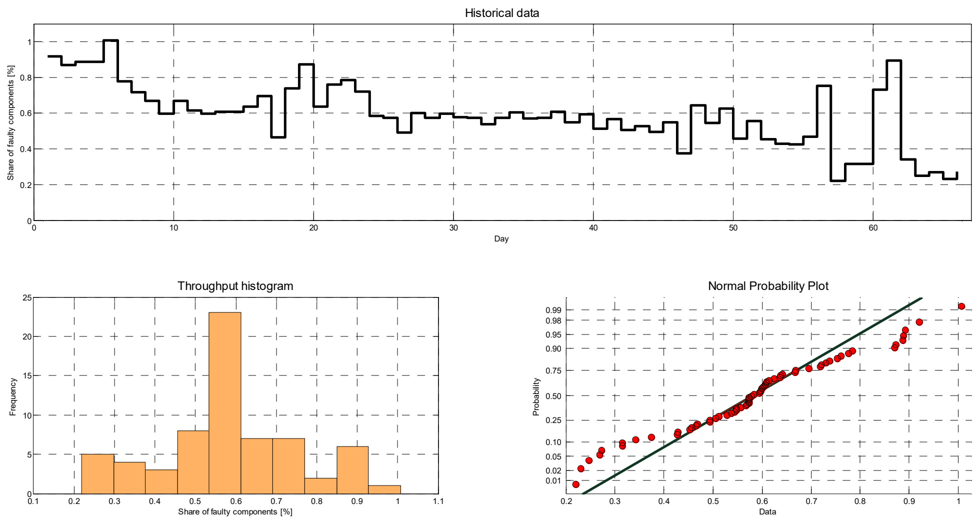


Figure 1: Data set used to create forecasting models

Table 1: Values of basic statistical parameters

Parameter	Value
Mean	0.005853
Standard deviation	0.001737
Min	0.002195
Max	0.010067
Range	0.007872
Median	0.005811
Angular coefficient	-0.000063
Final constant weight	0.007967

Figure 2 visualizes the values of ex-post forecasts made accordingly to the methods presented in this paper.

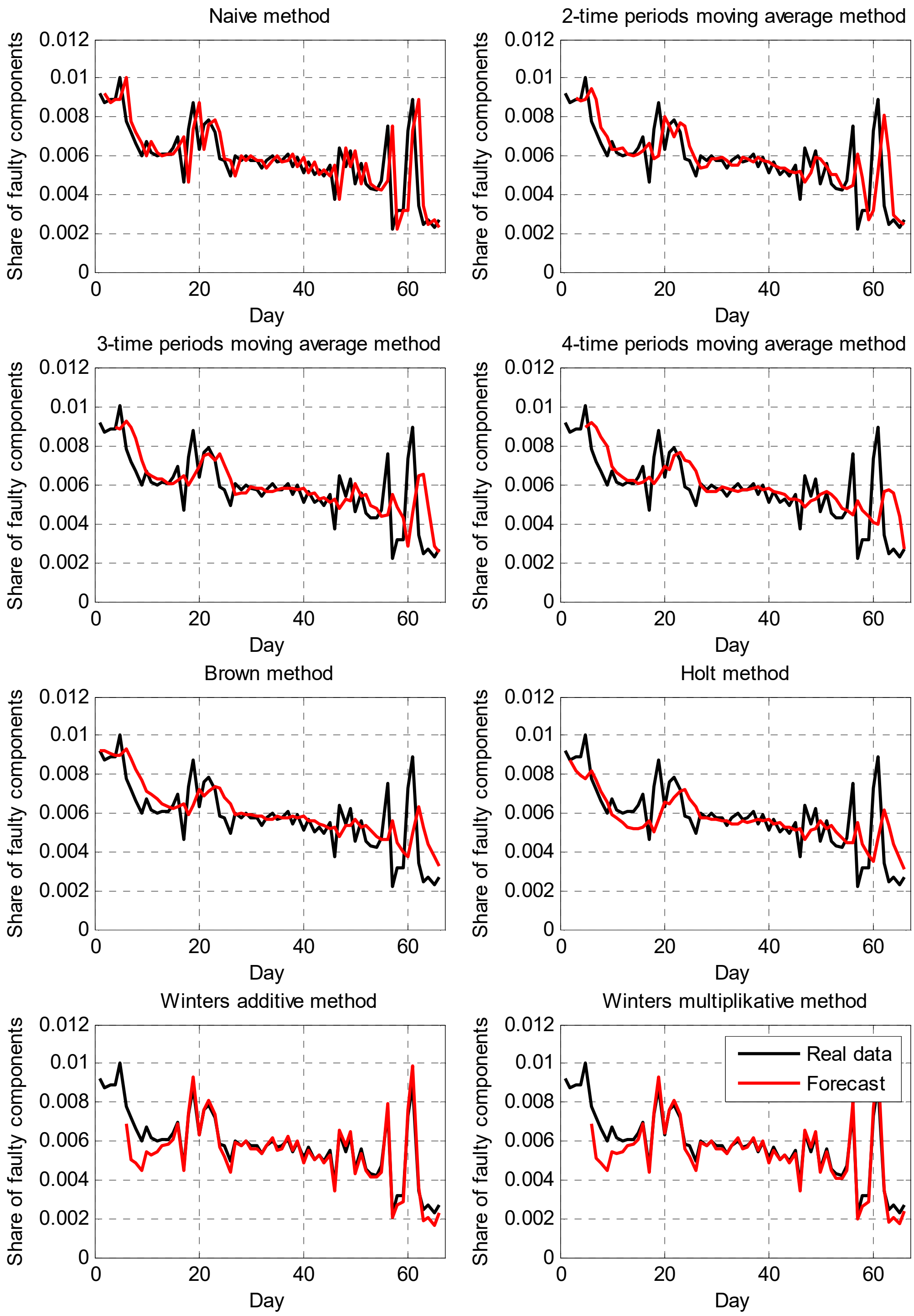


Figure 2: Forecasts of faulty production share in total production

The obtained results have been assessed based on the criteria presented in Table 2.

Table 2: Basic criteria for assessing ex-post errors

Error measure	Formula
Mean error	$ME = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{'}\right)$
Mean absolut error	$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{'}\right|$
Root mean squared error	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{'}\right)^2}$
Mean percentage error	$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i - y_i^{'}}{y_i}$
Mean absolute percentage error	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i^{'}}{y_i}\right|$

Source: own elaboration based on [Krzyżaniak 2008, s. 53].

In Table 3 the calculated values of criteria presented in Table 2 have been collected.

Table 3: Values of various forecasting error types for each one of used forecasting method

Forecasting method	Error measure				
	ME	MAE	RMSE	MPE	MAPE
Naive methods	-0.000100	0.000958	0.001490	-6.75%	19.99%
Moving average, two periods	-0.000147	0.000940	0.001463	-8.65%	19.87%
Moving average, three periods	-0.000200	0.000989	0.001475	-10.45%	21.33%
Moving average, four periods	-0.000254	0.001004	0.001442	-11.78%	21.71%
Brown’s method	-0.000272	0.000939	0.001327	-11.13%	19.98%
Holt’s method	0.000133	0.000971	0.001377	-4.78%	19.92%
Winter’s additive model	0.000216	0.000353	0.000551	4.88%	6.84%
Winter’s multiplicative model	0.000213	0.000343	0.000542	4.90%	6.68%

In case of exponential smoothing method the constants have been set in such a way that minimizes the value of RMSE criterion. The values of constants are given in Table 4.

Table 4: Values of smoothing constants created for Brown, Holt and Winters forecasting models

Forecasting method	Smoothing parameters		
	α	β	γ
Brown’s method	0.3398	–	–
Holt’s method	0.3198	0.1111	–
Winter’s additive model	0.9054	0.2924	0.0010
Winter’s multiplicative model	0.9104	0.2948	0.0010

The lowers values of forecasting errors have been obtained from Winter's models. The analysis of errors distribution justify the selection of Winter's models as the most suitable tool (Fig. 3).

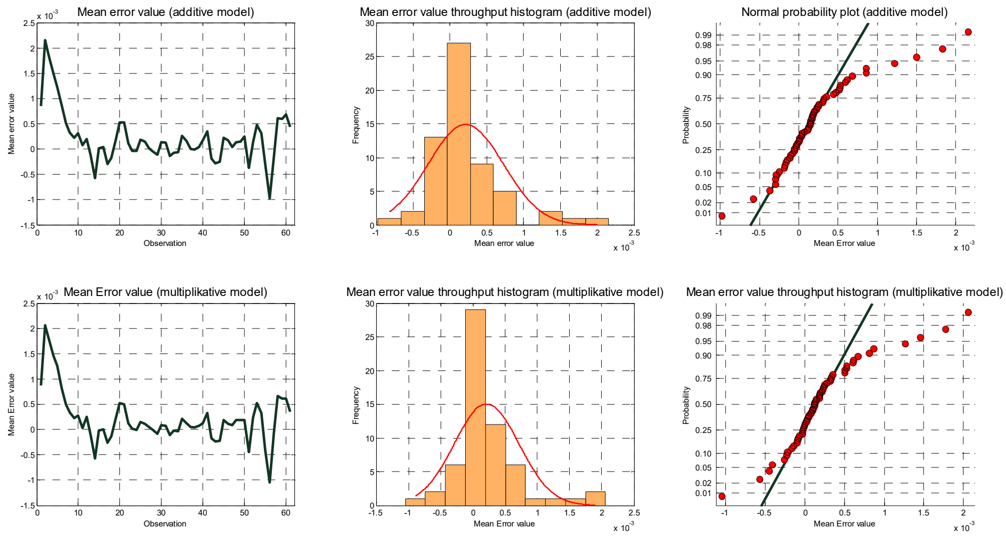


Figure 3: Analysis of errors resulting from using Winter's forecasting models

4. CONCLUSIONS

Forecasting and analyzing various time series is a continuous and extremely important topic both from a purely scientific as well as commercial (business) point of view. The aim of this paper was to analyze the suitability of classical forecasting methods to predict the share of faulty production in total production in given manufacturing company. Methods such as naïve, moving average and exponential smoothing have been used. The best results were obtained for forecasting models based on Winter's approach. This paper revealed some important future research directions which will be undertaken by the authors. Those include the use of seasonality indices and their modification to forecast the share of faulty production.

5. REFERENCES

[1] Box G. E. P., Jenkins G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
[2] Bozarth C., Handfield R. B. (2008). *Introduction to operations and supply chain management*. Upper Saddle River, NJ: Pearson Prentice Hall.
[3] Brown R. G. (1959). *Statistical Forecasting for Inventory Control*. McGraw-Hill, New York.
[4] Cieślak M. [Ed.] (1999). *Prognozowanie gospodarcze. Metody i zastosowania*. PWN, Warszawa, Poland.

[5] Dittmann P. (2008). *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie.* Wolters Kluwer business, Kraków, Poland.
[6] Dittmann P., Dittmann I., Szabela–Pasierbińska E., Szpulak A. (2009): *Prognozowanie w zarządzaniu przedsiębiorstwem*. Wolters Kluwer business, Kraków, Poland.
[7] Holt C. C. (1957). *Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages*. ONR Research Memorandum No. 52.
[8] Krzyżaniak S. (2008). *Podstawy zarządzania zapasami w przykładach*. ILiM, Poznań, Poland.
[9] Lichota A. (2006). *Prognozowanie krótkoterminowe na lokalnym rynku energii elektrycznej*. PhD dissertation, AGH-UST University of Science and Technology, Faculty of Management, Kraków, Poland.
[10] Sobczyk M. (2008): *Prognozowanie*. PLACET, Warszawa, Poland.
[11] Winters P. R. (1960). *General Exponential forecasting; A Computer Program for the IBM 360*. ONR Research Memorandum No. 71.
[12] Zeliaś A., Pawełek B., Wanat S. (2003). *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*. PWN, Warszawa, Poland.
[13] Zeliaś A. (1997). *Teoria prognozy*. PWE, Warszawa, Poland.

A Multi-Criteria Optimization Approach for Modelling Negotiation Process

Andrzej ŁODZIŃSKI[1]

[1] *Warsaw University of Life Sciences, ul. Nowoursynowska 159, 02-776 Warszawa, Poland*
Email: andrzej_lodzinski@sggw.pl

Abstract: The paper presents a multi-criteria optimization approach for modelling negotiation process. The negotiation process is modelled as a special multi-criteria task. The method of finding solutions is the interactive selection process of some proposals. The parties shall submit their proposals to the subjects of the negotiations. These proposals are parameters of the multi-criteria optimization tasks. Selection of solutions is accomplished by solving the optimization task with parameters that define the aspirations of each party involved in the negotiations. Finally, evaluation of the solutions obtained by the parties is done.

Keywords: negotiation process, multi-criteria optimization, nondominated solution, equitably efficient decision, scalarizing the problem, achievement function, set of negotiation, method of solution selection.

1. Introduction

The paper presents a multi-criteria optimization approach for modelling negotiation process. The negotiation process is modelled as a special multi-criteria task. The solution of this task has properties that the parties consider to be justified in negotiations. The solution obtained in such a way is acceptable by the parties. Negotiations serve to agree the decisions when different interests of participants occur. Negotiations are carried out to reach a more favorable result than that which can be achieved without negotiation. Negotiating parties could benefit, coming to the agreement with each other, in comparison to the situation when they act separately. Well arranged agreement is better for the parties than no agreement at all, and some agreements are more favorable for both parties than others. In the complex negotiations, the parties not only want to reach an agreement, but they are looking for the optimal agreement – i.e. the agreement that would be the best for both parties.

Negotiations are characterized by a lack of clear solutions and the necessity of taking into account the preferences of parties. The process of negotiations can be modeled using the game theory. The solution is then the Nash cooperative solution or Raiffy-Kalai-Smorodinsky solution [8], [12], [16], 17], [18].

The work is devoted to apply a multi-criteria optimization for decision selection in the negotiation process. The process of negotiations is modeled as a special multi-criteria optimization tasks, whose solution is equitably efficient decision. The method of decision selection is based on an interactive selection of some proposals of solutions, i.e. the algorithm requires the reaction of parties during this process. The parties submit their proposals for the subjects of negotiations; these proposals are parameters of the multi-criteria optimization task; this way the task is solved. Then, the parties evaluate the solution: they accept it or reject it. In the second case, the parties shall submit new proposals - the new values of parameters and the problem is solved again for these new parameters. The process of selection of solution is not a one-time process, but an iterative process of learning by parties about the negotiated problem.

2. Modeling of the negotiation process

The negotiation process is modeled as an interactive decision-making process. Each party presents its proposals of solutions. The negotiation process is then the process of seeking a common decision, which reconciles the interests of both parties. The parties are trying to find a common compromise solution. Decisions require voluntary consent of both parties and are taken together, not unilaterally. Both parties have to accept those decisions.

During the negotiation process, there are many different purposes, which are implemented using the same set of feasible solutions. The negotiation process is modeled by introducing a decision variable that describes the solution as well as two evaluation functions evaluating the solution from the point of view of each party. During the negotiations, each proposal is evaluated by either party by its evaluation function. Such a function is a measure of satisfaction of a party with a given solution. It evaluates the degree of realization of each subject of negotiations by each party. Higher value of the function means higher satisfaction of a party, so each function is maximized. The basis for evaluation and solution selection are two functions of evaluation - the criteria for both parties.

We assume the following terms:

party 1 and party 2 - parties in negotiations,

n – the number of subjects for negotiations,

$x \in X_0$ – a solution - a decision, the parties of which are to agree, belonging to the set of feasible decisions $X_0 \subset R^n$, $x = (x_1, x_2, ..., x_n)$ - each coordinate $x_i, i = 1,...,n$ defines $i-th$ subject of negotiations,

$f1: X_0 \to R$ – an evaluation function of decision x by party 1,

$f2: X_0 \to R$ – an evaluation function of decision x by party 2.

The problem of a decision selection has the multi-criteria character. Each party wants to maximize its evaluation function, but it must take into account the existence of the other party. The selection of solution is done by using both evaluation functions.

The negotiation process is considered as a task of multi-criteria optimization with the function of purpose $f = (f1, f2)$:

$$\max_x \{(f1(x), f2(x)) : \quad x \in X_0\} \tag{1}$$

where: $x \in X$ – a vector of decision variables,

$f = (f1, f2)$ – the vector function which maps the decision space X into evaluation space Y_0,

X_0 – the set of feasible decisions.

Task (1) is to find such feasible decision $\hat{x} \in X_0$ for which two evaluation functions $f1$ and $f2$ take the best values.

Task (1) is considered in the evaluation space, i.e., the following task is considered:

$$\max_x \{y = (y1, y2) : \quad y \in Y_0\} \tag{2}$$

where: $x \in X$ – a vector of decision variables,

$y = (y1, y2)$ – the vector quality indicator; individual coordinates $y_i = f_i(x), \quad i = 1, 2$ represent single scalar criteria, first coordinate is evaluation criteria of a solution by party 1, the second coordinate is evaluation criteria of a solution by party 2,

$Y_0 = (f1, f2)(X_0)$ – the set of achievements vectors of evaluation.

The set of achievements results Y_0 is given in the implicit form - through a set of feasible decisions X_0 and mapping of a model $f = (f1, f2)$. To determine the value y, the simulation of the model is needed: $y = (f1, f2)(x) \ for \quad x \in X_0$.

The purpose of task (1) is to help in the selection of such a decision, which takes into account the best interests of both parties [7], [10], [15], [20].

3. equitably efficient decision

The solution in the negotiation process should satisfy certain properties that the parties accept as reasonable. The solution should be:

- nondominated solution (Pareto-optimal solution) – i.e. such that you can not improve the solution for one party without worsening the solution for the other party,
- symmetric solution – i.e. that it should not depend on the way the parties are numbered, no one is more important, the parties are treated in the same way in the sense that the solution does not depend on the names of the party or other factors specific to a given party,
- equalizing solution – that is, a vector that has less variation of coordinates of evaluation is preferred in comparison to the vector with the same sum of coordinates, but with a greater diversity of coordinates,
- the solution should take into account the strength of the parties in the negotiations.

A decision, which satisfies the first three conditions is an equitably efficient decision. This is efficient decision which satisfies additional conditions – anonymity and the axiom of equalizing solution.

Nondominated solutions (Pareto – optimal solutions) are defined as follows:

$$\hat{Y}_0 = \{\hat{y} \in Y_0 : (\hat{y} + \widetilde{D}) \cap Q_0 = \varnothing\} \tag{3}$$

where: $\widetilde{D} = D \setminus \{0\}$ – a positive cone without the top. As a positive cone, it can be adopted $\widetilde{D} = R_+^m$ [7], [21].

In the decision space, the appropriate feasible decisions are specified. The decision $\hat{x} \in X_0$ is called an efficient decision, if the corresponding vector of evaluation $\hat{y} = f(\hat{x})$ is a nondominated vector.

In the multi-criteria problem (1), which is used to select a decision in the negotiation process, the relation of preferences should satisfy additional properties: anonymity property and property of equalizing solution.

This preference relation is called an anonymous relation if, for every assessments $y = (y_1, y_2, ..., y_m) \in R^m$ and for any permutation P of the set $\{1,..., m\}$, the following property holds:

$$(y_{P(1)}, y_{P(2)}, ..., y_{P(m)}) \approx (y_1, y_2, ..., y_m) \tag{4}$$

No distinction is made between the results that differ in their arrangement of coordinates. Evaluation vectors having the same coordinates, but in a different manner are identified.

Relation of preferences satisfies the axiom of equalizing transfer if the following condition is satisfied:

for the evaluation vector $y = (y_1, y_2, ..., y_m) \in R^m$:

$$y_{i'} > y_{i''} \Rightarrow y - \varepsilon \cdot e_{i'} + \varepsilon \cdot e_{i''} \succ y \ \text{ for } 0 < y_{i'} - y_{i''} < \varepsilon \tag{5}$$

Equalizing transfer is a slight deterioration of a better coordinate of the evaluation vector and simultaneously improvement of a poorer coordinate, giving the evaluation vector strictly preferred in comparison to the initial evaluation vector. This is a structure of equalizing – the evaluation vector with less diversity of coordinates is preferred in relation to the vector with the same sum of coordinates, but with a greater diversity of coordinates.

Nondominated vector satisfying the anonymity property and the axiom of equalizing transfer is called an equitably nondominated vector. The set of equitably nondominated vectors is denoted by $\hat{Y}_{0e}$. In the decision space, the equitably efficient decisions are specified. The decision $\hat{x} \in X_0$ is called an equitably efficient decision, if the corresponding evaluation vector $\hat{y} = f(\hat{x})$ is an equitably nondominated vector. The set of equitably efficient decisions is denoted by $\hat{X}_{0e}$ [4], [14], [15].

The relation of equitable domination can be expressed as the relation of inequality for cumulative, ordered evaluation vectors. This relation can be determined with the use of transformation $\overline{T} : R^m \to R^m$ that cumulates nonincreasing coordinates of evaluation vector.

The transformation $\overline{T} : R^m \to R^m$ is defined as follows:

$$\overline{T}_i(y) = \sum_{j=1}^{i} T_j(y) \quad \text{for } i = 1, 2, ..., m \tag{6}$$

where: $T(y)$ is the vector with nonincreasing ordered coordinates of the vector y, i.e. $T(y) = (T_1(y), T_2(y), ..., T_m(y))$, where $T_1(y) \le T_2(y) \le ... \le T_m(y)$ and there is a permutation P of the set $\{1, ..., m\}$, such that $T_i(y) = y_{P(i)}$ for $i = 1,.., m$.

The evaluation vector y^1 dominates in equitable way the vector y^2 if the following condition is satisfied:

$$y^1 \succ_e y^2 \Leftrightarrow \overline{T}(y^1) \ge \overline{T}(y^2) \tag{7}$$

The relation of equitable domination $\succ_e$ is a simple vector domination for the evaluation vectors with coordinates which are accumulated values of ordered evaluation vector [4], [14], [15].

Solving the problem of decision selection in the negotiations process consists in determination of an equitably efficient decision that satisfies the preferences of parties.

4. Scalaring the problem

For determination of an equitably efficient solutions of multi-criteria task (1) a specific multi-criteria task is solved. It is the task with the vector function of the cumulative, ordered evaluation vectors, i.e. the following task:

$$\max_y \{(\overline{T}_1(y), \overline{T}_2(y), ..., \overline{T}_m(y)) : \quad y \in Y_0\} \tag{8}$$

where: $y = (y_1, y_2, ..., y_m)$ – an evaluation vector,

$\overline{T}(y) = (\overline{T}_1(y), \overline{T}_2(y), ..., \overline{T}_m(y))$ - a cumulative, ordered evaluation vector,

Y_0 – the set of the achievements evaluation vectors.

An efficient solution of multi-criteria optimization tasks (8) is an equitably efficient solution of the multi-criteria task (1).

To determine the solution of a multi-criteria task (8) the scalaring of this task with the scalarizing function $s : Y \times \Omega \to R^1$ is introduced:

$$\max_x \{s(y, \overline{y}) : x \in X_o\} \tag{9}$$

where: $y = (y_1, y_2, ..., y_m)$ – an evaluation vector,

$\overline{y} = (\overline{y}_1, \overline{y}_2, ..., \overline{y}_m)$ – a control parameters for individual evaluations.

It is the task of single objective optimization with specially created the scalarizing function of two variables – the evaluation vector $y \in Y$ and control parameter $\overline{y} \in \Omega \subset R^m$. It is the function $s : Y \times \Omega \to R^1$. The parameter $\overline{y}$ is available to the parties, it allows them to review the set of equitably efficient solutions.

A optimal solution of the task (9) should be a solution to the multi-criteria task (8). A scalarizing function should satisfy certain properties - property of completeness and property of sufficiency. The property of sufficiency means that for each control parameter $\overline{y}$ the solution of the scalaring task is the equitably efficient solution, i.e. $\hat{y} \in \hat{Y}_{0e}$. The property of completeness means, that by appropriate changes of the parameter $\overline{y}$ it can be achieved any solution $\hat{y} \in \hat{Y}_{0e}$. Such a function completely characterizes the equitably efficient solutions. Inversely, each maximum of such a function is an equitably efficient solution. Each equitably efficient solution can be achieved with some appropriate values of control parameters $\bar{y}$.

Complete and sufficient parameterization of the set of equitably efficient solutions $\hat{Y}_{0e}$ can be achieved, using the reference point method for the task (8). This method makes use of aspiration levels as control parameters. Aspiration levels are such values of evaluation function that satisfy the parties.

The scalarizing function in the reference point method is as follows:

$$s(y,\bar{y}) = \min_{1\le i\le m}(\bar{T}_i(y) - \bar{T}_i(\bar{y})) + \varepsilon \cdot \sum_{i=1}^{m}(\bar{T}_i(y) - \bar{T}_i(\bar{y})) \tag{10}$$

where: $y = (y_1, y_2, ..., y_m)$ – an evaluation vector,

$\bar{T}(y) = (\bar{T}_1(y), \bar{T}_2(y), ..., \bar{T}_m(y))$ – a cumulative, ordered evaluation vector,

$\bar{y} = (\bar{y}_1, \bar{y}_2, ..., \bar{y}_m)$ – a vector of aspiration levels,

$T(\bar{y}) = (T_1(\bar{y}), T_2(\bar{y}), ..., T_m(\bar{y}))$ – a cumulative, ordered vector of aspiration levels,

ε – an arbitrary small, positive adjustment parameter.

Such scalarizing function is called the achievement function. The aim is to find a solution that approaches as close as possible the specific requirements – the aspiration levels.

Maximizing this function with respect to x determines the equitably efficient solution $\bar{y}$ and the equitably efficient decision $\hat{x}$. Note, the equitably efficient decision $\hat{x}$ depends on the aspiration levels $\bar{y}$ [7], [14], [15].

5. Set of negotiation

The aim of the complex negotiations is not only the achievement of an agreement between the parties, even if it is beneficial for both parties, but finding a solution that meets the expectations of parties as much as possible and, if it is not worse than a solution attainable without negotiations.

Before starting the negotiations, parties should consider what is the result they can achieve if negotiations are not successful - the status quo point. This point is the result which can be achieved by each party without negotiation with the other one. If the parties can achieve the result $ys = (y1s, y2s)$ without negotiations - part 1 can achieve the result $y1s$, part 2 – the result $y2s$, then, no one party will agree to the worse result. During negotiations, parties want to improve the solution in relation to this point. The status quo point determines the strength of the parties in the negotiations and, what is their impact on the result.

The set of negotiations is a collection of equitably dominated evaluation values dominating the status quo point.

The set of negotiation is as follows:

$$B(\hat{Y}_{0e}, ys) = \{\hat{y} = (\hat{y}1, \hat{y}2) \in \hat{Y}_{0e} \wedge \hat{y}1 \geq y1s \wedge \hat{y}2 \geq y2s\} \tag{11}$$

where: $\hat{y} = (\hat{y}1, \hat{y}2) \in \hat{Y}_{0e}$ – an equitably nondominated vector,

$ys = (y1s, y2s)$ – the status quo point - the result, which can be achieved by both parties without agreement.

A set of negotiations embraces the points from the set of equitably nondominated results, which give each party at least as much as it can achieve individually (without negotiation).

The parties wish to find such a decision, $\hat{x} \in X_0$, that the corresponding evaluation vector $\hat{y} = (\hat{y}1, \hat{y}2) = (f1(\hat{x}), f2(\hat{x}))$ belongs to the set of negotiations $B(\hat{Y}_{0e}, ys)$ [8], [16], [17].

6. Method of solution selection

The solution to multi-criteria optimization task (8) is the set of equitably efficient decisions. In order to resolve the problem there should be selected one solution that will be evaluated by both parties. Since the solution is a whole set, the parties shall select the solution with the help of an interactive computer system. Such a system allows us a controlled overview of the whole set. Each party attending the negotiation determines its proposed solutions as aspiration levels. These are the values of evaluation of individual negotiation issues, that each party would like to achieve. These values are control parameters of the scalarizing function. For these values the system indicates different equitably efficient solutions for analysis; they correspond to current values of the control parameters. The aim is to find solutions which meet, as close as possible, the specific requirements – the aspiration levels.

1. The method of decision selection is as follows:
2. The initial arrangements.
    a. Iterative algorithm – proposals for further decisions.
    b. The interaction with the system – parties define their proposals for individual subjects of negotiations, as aspirations levels $\bar{y}1$ and $\bar{y}2$.
    c. Calculations – giving another equitably efficient solution $\hat{y} = (\hat{y}1, \hat{y}2)$.
    d. Evaluation of the obtained solutions $\hat{y} = (\hat{y}1, \hat{y}2)$ – the parties may accept the solution or not. In the latter case, the parties shall submit new proposals – they provide new values of their aspiration levels $\bar{y}1$ and $\bar{y}2$ and a new solution is determined (Return to step 2.2).
3. Determination of the decision that meets the requirements of both parties.

A choice is not a single act of optimization, but a dynamic process of searching solutions. That means the parties learn and can change their preferences during the process. Comparing the results of the evaluation $\hat{y}1$ and $\hat{y}2$ to their aspiration points $\bar{y}1$ and $\bar{y}2$, we see that each party has information about what is and what is not achievable, and how far the parties 'proposals $\bar{y}1$ and $\bar{y}2$ are from the possible solutions $\hat{y}1$ and $\hat{y}2$. This allows the parties to do appropriate modifications of their proposals: to provide their new aspiration levels. These levels of aspiration are determined adaptively during the learning process. The process ends when the parties find such a decision, which allows them to achieve results that meet their aspirations or, in a sense, are as close as possible to these aspirations. The method of decision selection is show at Figure 1.

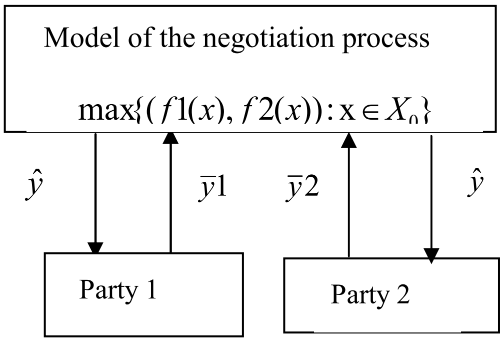


Figure 1: The method of decision selection

This method of decision selection does not impose any rigid scenario on parties and allows them to change their preferences while solving the problem. As we see, parties are learning about the problem during the negotiation. The computer does not replace the parties in selection of solution. It should be witnessed that the entire process of solution selection is controlled by both parties.

7. Example

In order to illustrate the method of multi-criteria optimization for decision selection in a process of negotiations, the following example is presented [13].

The problem related to negotiations is the following:

party 1 and party 2 – parties to negotiations,

$n = 2$ – the number of subjects for negotiations,

$x = (x_1, x_2) \in X_0$ – a solution – a decision to be agreed upon by the parties, belonging to the feasible decisions set, $X_0 \subset R^2$, x_1 – decision concerning the first subject of negotiations, x_2 – decision concerning the second subject of negotiations,

$X_0 = \{(x_1, x_2) \in R^2 : x_1 + 3 \cdot x_2 \le 63,\ 5 \cdot \mathrm{x}_1 + 4 \cdot x_2 \le 117,\ x_1 \ge 0, \mathrm{x}_2 \ge 0\}$ – the feasible decisions set,

$f1 : X_0 \to R^1\ f1(x) = x_1$ – function of evaluation of decision x by party 1,

$f2 : X_0 \to R^1\ f2(x) = x_2$ – function of evaluation of decision x by party 2,

$ys = (ys1, ys2) = (10, 10)$ – the status quo point.

The negotiations process is modelled as a multi-criteria optimization problem with the vector function of the objective $f = (f1, f2)$:

$$\max_x \{(f1(x), f2(x)) : \quad x \in X_0\} \tag{12}$$

where: $x = (x_1 x_2) \in X_0$ – a vector of decision variables,

$y = f(x_1, x_2)$ – the vector function that's maps the decision space X_0 into evaluation space $Y_0 \subseteq R^2$,

X_0 – the feasible decisions set.

In the multi-criteria model (12) individual solutions are evaluated using the evaluation vector $y = (f1, f2)$ where $f1$ is an evaluation function of decision x by party 1 and $f2$ is an evaluation function of decision x by party 2.

As a first step of the multi-criteria analysis, a single-criterion optimization of evaluation function of each party is done. The result is a matrix of implementation goals, containing the values of all criteria of each party, received during solving two single-criterion problems. This matrix allows for the estimation of the range of changes of each evaluation function on the feasible set, as well as provides some information about the conflictual nature of evaluation function. The matrix of the implementation of goals generates an utopia vector representing the best value of each separate criterion.

Table 1: The matrix of the implementation of goals with the utopia vector

Optimization criterion	Solution $\hat{y}_1$ $\hat{y}_2$
Function $f1$ Function $f2$	23,4 0 -0 21
Utopia vector $(y1u, y2u)$	23,4 21

Analysis of Table 1 shows that the negotiation force of both parties are similar.

For each iteration, the Price of Fairness POF each party is calculated [1]. It is the quotient of the difference between the utopia value of a solution and the value from the solution of the multi-criteria task, in relation to the utopia value.

$$POF = \frac{yiu - \hat{y}i}{yiu}, i = 1,2 \tag{13}$$

where: yiu - the utopia value of a party i, $i = 1, 2$

$\hat{y}i$ – the value from the solution of the multi-criteria tasks of a party i, $i = 1, 2$

The value of the POF's is a number between 0 and 1. POF's value closer to zero are preferred by the parties, as the solution is closer to an utopia solution. The more the values of the POF's of both parties get closer to each other, the more the solution will be considered better.

The parties control the decision-making regarding the solution giving their propositions in the form of aspiration levels, constituting the desired values of their evaluation functions, and the system gives solutions corresponding to the current values of parameters that can be analysed by both parties. The multi-criteria analysis is presented in Table 2.

Table 2: Interactive analysis of seeking a solution

Iteration	Party 1 Party 2 $\hat{y}_1$ $\hat{y}_2$
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	23,4 21 9 18 0,61 0,14
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	18 16 9 18 0,61 0,14
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	16 14 10,6 16 0,54 0,23
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	15 13 11,4 15 0,51 0,28
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	14 12 12,2 14 0,47 0,33
Aspiration levels $\bar{y}$ Solution $\hat{y}$ POF	13 12 13 13 0,44 0,38

At the beginning of the analysis, the parties determine their preferences as the aspiration point equal to vector utopia. The solution is clearly more favourable for party 2 and is not acceptable by party 1 as it is worse from its point of reference. The cost of the POF's fairness for party 1 amounts to 0.61, while for party 2 to 0.14. To improve the solution, both parties will reduce their requirements within the next iteration. The solution has not changed. Within subsequent iterations, the parties keep on reducing their requirements. The solutions are each time more favourable for party 1 and less favourable for party 2. The values of the cost of the POF's fairness for party 1 are becoming smaller, while for party 2 they are increasing, and they are getting closer to each other. Within iteration 6, the cost of the POF's fairness for party 1 is 0.44 and for party 2 – 0.38. Further attempts to change the solution do not result in an approximation of the POF's values of the parties. The POF's value of party 1 cannot be improved any more. It is not possible to make the values of the POF for both parties more equal. This is the maximum that party 1 can achieve. For iteration 6, the solution of the negotiation process is the following decision $\hat{x}^6 = (13, 13)$.

The final choice of a specific solution depends on the preferences of the parties. This example shows that the presented method allows the Parties to get to know their decision-making possibilities within interactive analysis and to search for a solution that would be satisfactory for both parties.

8. Conclusion

The paper presents a multi-criteria optimization approach for modelling negotiation process. The negotiation process is modelled as a special multi-criteria task. . It is used to support the decision selection. The model of the negotiation process as a multi-criteria optimization task allows us to create variants of decision and to track their consequences.

The method of interactive analysis, based on the reference point method, is applied for multi-criteria task with a cumulative, ordered evaluation vector.

This method is characterized by:

- the use of information about parties' preferences in the form of aspiration points - values of goal function that are fully satisfactory to them and the optimal option of the scalarizing achievement function in order to organize the interactions with the parties,
- the assumption that the preferences of parties are not completely fixed and they may change during the decision making process.

It allows us to determine solutions, well-tailored to the parties preferences. The numerical example shows that the proper computational task efficiently can be solved by the standard optimization software.

This procedure does not determine the final solution, but supports and teaches the parties about the specific negotiation problem. The final decision is to be taken by the parties involved in the negotiations.

9. References

[1] D. Bertsimas, V. F. Farias, and N. Trichakis, “The price of fairness,” Operations Research, vol. 59, no. 1, pp. 17–31, 2011.

[2] U. Chevaleyre,. J. Endriss, M. Lang, and N. Maudet “A Short Introduction to Computational Social Choice”. Proc. SOFSEM-2007, Springer-Verlag, 2007.
[3] R. Fisher, Ury W., Patron B. “Getting to YES. Negotiating agreement without giving in”. (in polish) PWE. Warsaw, 2002.
[4] M. Kostreva., W. Ogryczak. and A., P., Wierzbicki “Equitable Aggregation and Multiple Criteria Analysis”. European Journal of Operational Research, vol. 122., 2004.
[5] L. Kruś “On some Procedures Supporting Multi-criteria Cooperative Decisions”. Foundations of Computing and Decision 33, (3), 2008.
[6] W. Kubiak “Proportional Optimization and Fairness”, Springer, New York, NY, USA, 2009.
[7] Lewandowski A. and A., P. Wierzbicki eds. “Aspiration Based Decision Support Systems”. Lecture Notes in Economics and Mathematical Systems. Vol. 331, Springer-Verlag, Berlin-Heidelberg, 1989.
[8] D. Luce., H. Raiffa. “Games and decisions”. (in polish) PWN, Warsaw, 1996.
[9] H. Luss, Equitable Resource Allocation: Models, Algorithms, and Applications, John Wiley & Sons, Hoboken, NJ, USA, 2012.
[10] Łodziński. „Decision suport system for satisficing decision” (in polish) Zeszyty naukowe AGH, Zagadnienia techniczno – ekonomiczne, Uczelniane Wydawnictwo Naukowo – Dydaktyczne AGH, pp. 159–165 , Kraków 2007.
[11] Łodziński.”Interactive support of multiple-choice decision” (in polish) Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą, str. 147–155, Bydgoszcz 2008.
[12] M. Malawski, A. Wieczorek and H. Sosnowska. “Competition and Cooperation. Game Theory in Economics and the Social Sciences”. (in polish) PWN, Warsaw, 1997.
[13] W. Ogryczak. (2002) “Multicriteria Optimization and Decisions under Risk.” Control and Cybernetics, vol. 31, 2002 No. 4.
[14] W. Ogryczak, A., M. Kostrzewa, A. P. . Wierzbicki and M. Milewski: “Equitable Aggregations and Multiple Criteria Analysis”, European Journal of Operational Research, 158, 2004
[15] W. Ogryczak, A., P. Wierzbicki and M. Milewski: “A Multi-Criteria Approach to Fair and Efficient Bandwidth Allocation”, OMEGA, 36, 2008.
[16] D., Ph .Straffin “Game Theory”. Scolar. (in polish) Warsaw, 2004.
[17] H. Raiffa.; I. Richardson; D. Metcalfe “Negotiation Analysis”. The Belknap Press of Harvard University Press, Cambridge, 2002.
[18] E. Roszkowska E. “The selected models of negotiation” (in polish). Wydawnictwa Uw B, Białystok 2011.
[19] H., P. Young “Equity: In Theory and Practice”. (in polish) Warsaw, 2003.
[20] A., P. Wierzbicki “A mathematical basis for satisficing decision making”. Mathematical Madelling, 3, 1982.
[21] P. Wierzbicki. “Negotiation and mediation in conficts. Plural rationality and interactive decision processes”, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin Heidelberg New York Tokyo., 1984.
[22] R. R. Yager “On ordered weighted averaging aggregation operators in multicriteria decisionmaking” IEEE Transactions on systems, Man, and Cybernetics, 1988.

The Use of Computer Assistance for the Warehouse Location Problem in the Logistics Network of the Production Companies

Mariusz IZDEBSKI [1], Ilona JACYNA-GOŁDA[2], Emilian SZCZEPAŃSKI[3]

[1,3]*Warsaw University of Technology, Faculty of Transport*
Koszykowa 75, 00-662 Warsaw, Poland, mizdeb@wt.pw.edu.pl

[2]*Warsaw University of Technology, Faculty of Production Engineering*
Narbutta 85, 02-524 Warsaw, Poland, e-mail: jacyna.golda@gmail.com

Abstract: This paper presents warehouse location problem in the logistics network of the production companies. In order to solve this problem the computer assistance so called SIMLoc was developed. The construction and the operation of this application was presented. The operation of this application is based on the mathematical model of warehouse location which was described. The limitations and the optimization criterion of the model were determined. The optimization criterion refers to the transportation costs between the facilities of the network. The final location of warehouse facilities was obtained using a genetic algorithm. This paper describes the stages of this algorithm i.e. the crossover and mutation process.

Keywords: genetic algorithm, warehouses location problems, optimization

1. INTRODUCTION

The problem of warehouses location is an issue that is generally known and widely discussed in the context of decision-making problems which occur on the stage of constructing the logistic network in any company (Akinc et al. 1977; Brandeau, 1989; Özcan, 2011). The classical issue of warehouse location is defined in literature as the “capacitated warehouse location problem” CWLP (Sharma et al. 2007). In this class of problems the structure of the transport and logistics network consists of warehouse facilities and customers. The main objective in the general warehouses location problem is to find such a location of warehouse facilities for which the costs arising from forwarding a specific volume of goods to the customers via warehouses are as low as possible. The transport cost depends on the volume of goods which are carried between warehouses and customers and also on the distance between them, and hence it is also necessary to set out volumes of carried goods. Limitations, on the other hand, arise from satisfying the needs of all customers and the capacitive limit for the dispatch of goods from the given warehouse facility.

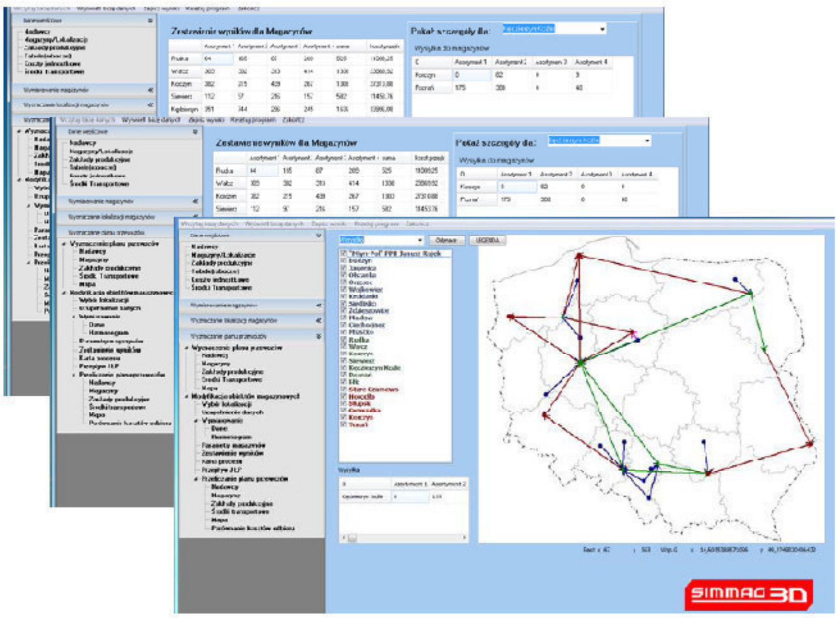


Figure 1: The application SIMLock

The warehouses location problem depends on the complexity of the logistics network. The logistic network of production companies consists of suppliers and warehouses (Jacyna-Gołda 2013; Jacyna-Gołda 2015; Jacyna-Gołda 2015, Jacyna-Gołda 2017, Jacyna et al. 2015**)**. The main aim is to deliver the cargo to the production companies from the suppliers in the directly way or via the warehouses. It depends on the given criterion function e.g. minimal cost associated with providing cargo to the production companies. This cost depends on transportation costs between the network objects and storage costs in the warehouses. The transportation task in the production companies can be defined in the following way: the vehicles leave the production company and goes to the supplier or the suppliers, when the vehicles are loaded they go to the warehouses or the production company. The starting point of the task can be also the warehouse. Transport of the cargo between the warehouses and the production companies is treated also as the task. In order to solve the warehouses location problem the computer assistance so called SIMLoc was developed. This application contains three modules (Jacyna-Gołda et al. 2016): the input module which introduces the characteristics of network objects, the optimization module which designate the optimal location of warehouses, the output module which generates the results. The input data are downloaded from the database. The optimization module is based on the genetic algorithm (Szczepański E. 2014, Jacyna-Gołda 2016; Izdebski 2014, Izdebski 2014). The output data are presented in the graphical and tabular way (Fig. 1).

2. THE MATHEMATICAL MODEL OF THE PROBLEM

The application SIMLock is based on the mathematical model of warehouse location which is presented in the following way. For mathematical formulation of the problem, the following assumptions are made:

- the mathematical model of warehouses location refers to the logistic network which consists of suppliers, warehouses and production companies (recipients),
- the suppliers provide different types of the raw material to the production companies on a given working day,
- the production capacity of suppliers on a given working day meet the needs of production companies.

In accordance with above list of assumptions proposed optimization task takes the following form (Żochowska R. 2014; Karkula 2014, Pyza 2011; Jacyna-Gołda 2017):

V – set of numbers of spot elements of the logistical network: suppliers, warehouses, production companies, $\boldsymbol{V} = \{v\text{: } v=1,2, \dots, v', \dots, V\}$,

DS – set of numbers of suppliers, $\alpha(v)=\{0,1,2\}$- the mapping which assigns the number of elements of the logistic network to object type, $\boldsymbol{DS}=\{v\text{: } \alpha(v)=0 \text{ for } v \in \boldsymbol{V}\}$,

MS – set of numbers of warehouses, $\boldsymbol{MS}=\{v\text{: } \alpha(v)=1 \text{ for } v \in \boldsymbol{V}\}$,

P – set of numbers of production companies (recipients), $\boldsymbol{P}=\{v\text{: } \alpha(v)=2 \text{ for } v \in \boldsymbol{V}\}$,

H – set of number of types of the raw material, $\boldsymbol{H} = \{1, \dots, h, \dots, H\}$,

T – set of numbers of working days, $\boldsymbol{T} = \{1, \dots, t, \dots, T\}$,

D1 – distance matrix in relations: suppliers – warehouses, $\mathbf{D1}=[d1(v,v')\text{: } d1(v,v') \in \boldsymbol{R}^+, v \in \boldsymbol{DS}, v' \in \boldsymbol{MS}]$,

D2 – distance matrix in relations: suppliers - enterprises, $\mathbf{D2}=[d2(v,v')\text{: } d2(v,v') \in \boldsymbol{R}^+, v \in \boldsymbol{DS}, v' \in \boldsymbol{P}]$,

D3 – distance matrix in relations: warehouses - enterprises, $\mathbf{D3}=[d3(v,v')\text{: } d3(v,v') \in \boldsymbol{R}^+, v \in \boldsymbol{MS}, v' \in \boldsymbol{P}]$,

D4 – distance matrix in relations: warehouses - warehouses, **D4**=[$d4(v,v')$: $d4(v,v') \in \boldsymbol{R}^+$, $v \in \boldsymbol{MS}$, $v' \in \boldsymbol{MS}$],

Q1 – volume of deliveries from suppliers on a given working day, **Q1**=[$q1(v,h,t)$: $q1(v,h,t) \in \boldsymbol{R}^+$, $v \in \boldsymbol{DS}$, $h \in \boldsymbol{H}$, $t \in \boldsymbol{T}$],

Q2 – demand volume of enterprises on particular working days, **Q2**=[$q2(v,h,t)$: $q2(v,h,t) \in \boldsymbol{R}^+$, $v \in \boldsymbol{P}$, $h \in \boldsymbol{H}$, $t \in \boldsymbol{T}$],

Q3 – volume of the raw material which is stored in the warehouses on a given working day, **Q3**=[$q3(v,h,t)$: $q3(v,h,t) \in \boldsymbol{R}^+$, $v \in \boldsymbol{MS}$, $h \in \boldsymbol{H}$, $t \in \boldsymbol{T}$],

Q4 – maximal processing capacity for each warehouse, **Q4**=[$q4(v,h,t)$: $q4(v,h,t) \in \boldsymbol{R}^+$, $v \in \boldsymbol{MS}$, $h \in \boldsymbol{H}$, $t \in \boldsymbol{T}$],

UMZ – minimal volume of the raw material, which decides on selection of the warehouses added to the logistic network, **UMZ**=[$\delta(v,h)$: $\delta(v,h) \in \boldsymbol{R}^+$, $v \in \boldsymbol{MS}$, $h \in \boldsymbol{H}$],

POJZ – warehouses capacity **POJZ** =[$pojz(v)$: $pojz(v) \in \boldsymbol{R}^+$, $v \in \boldsymbol{MS}$],

C – transport costs of load unit per distance unit between particular facilities of the logistic network, **C**=[$c(v,v',h)$: $c(v,v',h) \in \boldsymbol{R}^+$, $v,v' \in \boldsymbol{V}\times\boldsymbol{V}$, $h \in \boldsymbol{H}$],

one should find the values of decision variables:

The type of the decision variables determines the volume of cargo which flows between the facilities on a given working day. The first type of variable formulated as matrix **Y11** (relation: suppliers – warehouses), **Y12** (relation: suppliers – production companies), **Y13** (relation: warehouses – production companies), **Y14** (relation: warehouses – warehouses) is interpreted as the volume of raw materials which are transported between network points on a given working day and takes the following form:

$$\mathrm{Y11}=[\, \mathrm{y11(v,v',h,t)}: \mathrm{y11(v,v',h,t)} \in \mathrm{R+}, \mathrm{v} \in \mathrm{DS}, \mathrm{v'} \in \mathrm{MS}, \mathrm{h} \in \mathrm{H}, \mathrm{t} \in \mathrm{T}] \quad (1)$$

$$\mathbf{Y12}=[\, y12(v,v',h,t): y12(v,v',h,t) \in \boldsymbol{R}^+, v \in \boldsymbol{DS}, v' \in \boldsymbol{P}, h \in \boldsymbol{H}, t \in \boldsymbol{T}] \quad (2)$$

$$\mathbf{Y13}=[\, y13(v,v',h,t): y13(v,v',h,t) \in \boldsymbol{R}^+, v \in \boldsymbol{MS}, v' \in \boldsymbol{P}, h \in \boldsymbol{H}, t \in \boldsymbol{T}] \quad (3)$$

$$\mathbf{Y14}=[\, y14(v,v',h,t): y14(v,v',h,t) \in \boldsymbol{R}^+, v \in \boldsymbol{MS}, v' \in \boldsymbol{MS}, h \in \boldsymbol{H}, t \in \boldsymbol{T}] \quad (4)$$

with the constraints:

- the production capacity of suppliers cannot be exceeded - suppliers can provide the raw material directly to the recipients or indirectly by the warehouses:

$$\forall v \in \boldsymbol{DS}, \forall h \in \boldsymbol{H}, \forall t \in \boldsymbol{T} \quad \sum_{v' \in \boldsymbol{MS}} y11(v,v',h,t) + \sum_{v' \in \boldsymbol{P}} y12(v,v',h,t) \le q1(v,h,t) \quad (5)$$

- recipients demands must be met – the raw material can flow to the recipients from suppliers or warehouses:

$$\forall v' \in \boldsymbol{P}, \forall h \in \boldsymbol{H}, \forall t \in \boldsymbol{T} \quad \sum_{v \in \boldsymbol{MS}} y13(v,v',h,t) + \sum_{v \in \boldsymbol{DS}} y12(v,v',h,t) = q2(v',h,\ t) \tag{6}$$

- warehouses capacity cannot be exceeded - the raw material can flow to the warehouses from suppliers or other warehouses:

$$\forall v' \in \boldsymbol{MS}, \forall t \in \boldsymbol{T} \quad \sum_{h \in \boldsymbol{H}} \sum_{v \in \boldsymbol{DS}} y11(v,v',h,t) + \sum_{h \in \boldsymbol{H}} \sum_{v \in \boldsymbol{MS}, v \neq v'} y14(v,v',h,t) + \sum_{h \in \boldsymbol{H}} q3(v',h,t) \leq pojz(v') \tag{7}$$

- the raw material flowing out from the warehouse cannot exceed the sum of the volume of the raw material which is stored in these warehouses and the volume of the raw material which flows into the warehouse:

$\forall v \in \boldsymbol{MS}, \forall h \in \boldsymbol{H}, \forall t \in \boldsymbol{T}$

$$\sum_{v' \in \boldsymbol{P}} y13(v,v',h,t) + \sum_{v' \in \boldsymbol{MS}, v \neq v'} y14(v,v',h,t) \leq q3(v,h,t) + \sum_{v' \in \boldsymbol{DS}} y11(v',v,h,t) + \sum_{v' \in \boldsymbol{MS}, v \neq v'} y14(v',v,h,t) \tag{8}$$

- the maximal processing capacity of warehouses cannot be exceeded:

$$\forall v' \in \boldsymbol{MS}, \forall h \in \boldsymbol{H}, \forall t \in \boldsymbol{T} \quad \sum_{v \in \mathrm{DS}} y11(v,v',h,t) + \sum_{v \in \mathrm{MS}, v \neq v'} y14(v,v',h,t) \leq q4(v,h,t) \tag{9}$$

- the minimal stream of the raw material flowing into the warehouses decides about the choice of those warehouses to the logistic network:

$$\forall v' \in \boldsymbol{MS}, \forall h \in \boldsymbol{H}, \forall t \in \boldsymbol{T} \quad \sum_{v \in \boldsymbol{DS}} y11(v,v',h,t) + \sum_{v \in \boldsymbol{MS}, v \neq v'} y14(v,v',h,t) \geq \delta(v',h) \tag{10}$$

which minimize the following criteria functions:

The criterion of transportation costs between the facilities of the logistic network:

$$\begin{aligned} &\mathrm{F1}(\mathrm{Y11}, \mathrm{Y12}, \mathrm{Y13}, \mathrm{Y14}) = \\ &\sum_{v \in \boldsymbol{DS}} \sum_{v' \in \boldsymbol{MS}} \sum_{t \in \boldsymbol{T}} \sum_{h \in \boldsymbol{H}} y11(v,v',h,t) \cdot d1(v,v') \cdot c(v,v',h) + \sum_{v \in \boldsymbol{DS}} \sum_{v' \in \boldsymbol{P}} \sum_{t \in \boldsymbol{T}} \sum_{h \in \boldsymbol{H}} y12(v,v',h,t) \cdot d2(v,v') \cdot c(v,v',h) + \\ &\sum_{v \in \boldsymbol{DS}} \sum_{v' \in \boldsymbol{P}} \sum_{t \in \boldsymbol{T}} \sum_{h \in \boldsymbol{H}} y13(v,v',h,t) \cdot d3(v,v') \cdot c(v,v',h) + \sum_{v \in \boldsymbol{MS}} \sum_{v' \in \boldsymbol{MS}} \sum_{t \in \boldsymbol{T}} \sum_{h \in \boldsymbol{H}} y14(v,v',h,t) \cdot d4(v,v') \cdot c(v,v',h) \\ &\longrightarrow \min \end{aligned} \tag{11}$$

3. THE OPTIMALIZATION MODULE IN THE APPLICATION

The optimization module is based on the genetic algorithm. The purpose of the genetic algorithm is to determine the amount of cargo which flows between the individual transport network objects for which the criterion function will reach a minimum value. The stages of the genetic algorithm take the following form: determining the structure processed by the algorithm, determining the adaptation function which evaluates the structure according to minimum transportation and reloading costs, the selection, crossover and mutation.

The crossover process and mutation are reiterated a given number of times, until the stop condition has been achieved. A condition for stop in the developed algorithm is the fixed iterations number. In the selection process the roulette method was adopted, while the process of crossover and mutation occurs with a defined likelihood set at the beginning of functioning of the algorithm.

3.1 The structure of the genetic algorithm

The structure of input data was presented as matrix **M(*h*)**, which presents the flow of *h*- th cargo between particular elements of the transport network on a given working day. Lines and columns of this matrix define facilities of the transport network structure. In order to determine the flow of cargo, lines were defined as the starting points from which cargo flows out to the other facilities. Matrix cells are located in the following sequence: suppliers, terminals and recipients. The graphical representation of the matrix structure **M(*h*)** with sample volumes was shown on Fig. 2 (**D**-suppliers, **TZ**-warehouses, **O**-recipients). On the basis of volumes in the matrix **M**(*h*) the criterion function presented in the chapter 2 can be calculated e.g. the decision variable which determines the cargo flow between D1 and TZ1 takes the value 20.

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	10	15
D2	0	0	15	20	0	0	5	5
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

Figure 2: The structure of the genetic algorithm

3.2 The adaptation function

On the basis of the adaptation function the genetic algorithms designate the final solution. The genetic algorithms look for the maximal solution. In order to take into account the mentioned aspect the adaptation function $\mathbf{Fp}_n$ for n-th structure takes the following form:

$$\mathrm{Fp}_n = \mathrm{C} - \mathbf{KPS}_n \tag{12}$$

where:

C – the value much higher than the value of the costs of the cargo flow in the network,

$\mathbf{KPS}_n$ – the cost of the cargo flow in n-this structure, formula (12) for *h*-th cargo.

The tendency of genetic algorithms is to maximize the function of adaptation. Maximization of the function Fp_n consequently is the process of minimization of the function $\mathbf{KPS}_n$, what is the assumed optimization aim.

3.3 The crossover process

To implement the crossover process, two matrices are developed: **DIV**(h) which comprise rounded up average values from both parents, and matrix **REM**(h) containing information whether the rounding up was indeed necessary. Assuming that the value of matrices **M1**(h) and **M2**(h) (parents) in all cells assume determination $m^1{}_{v,v'}$, $m^2{}_{v,v'}$, values of elements of matrices **DIV**(h) and **REM**(h) are calculated from the following dependencies:

$$dim_{v,v'} = \lfloor (m^1{}_{v,v'} + m^2{}_{v,v'}) / 2 \rfloor, \tag{13}$$

$$rem_{v,v'} = (m^1{}_{v,v'} + m^2{}_{v,v'}) / mod\,2, \tag{14}$$

The full description of the crossover operator process was presented in (Michalczewski 1996) and presented in a graphical way to Fig. 3, Fig. 4 and Fig. 5. The values with the matrix **REM** are added to the matrix **DIV**. As a result of this operation two new structures are developed. The applied crossover operator guarantees the correctness of individuals following a completed crossover, without the necessity of using repair algorithms.

M1(h)

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	10	15
D2	0	0	15	20	0	0	5	5
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

M2(h)

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	7	18
D2	0	0	15	20	0	0	8	2
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

Figure 3: Structures in the crossover process

DIV(h)

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	8	16
D2	0	0	15	20	0	0	6	3
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

REM(h)

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	0	0	0	0	1	1
D2	0	0	0	0	0	0	1	1
TZ1	0	0	0	0	0	0	0	0
TZ2	0	0	0	0	0	0	0	0
TZ3	0	0	0	0	0	0	0	0
TZ4	0	0	0	0	0	0	0	0

Figure 4: The matrix DIV and REM

New structure 1

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	9	16
D2	0	0	15	20	0	0	6	4
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

New structure 2

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	8	17
D2	0	0	15	20	0	0	7	3
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

Figure 5: New matrices after crossover process

3.4 The mutation process

To implement the crossover process The operation rule of mutation operator consists of sampling of two figures p and q from the range: $2 \leq p \leq k$ and $2 \leq q \leq n$, which determine the number of lines and columns of a sub-matrix with dimensions $p \times q$ (k - number of lines in the main matrix (processed by the algorithm), n - number of columns in this matrix). The generated matrix is modified in such a way that the total value in columns and lines before and after the modification process is not changed. The detailed mutation process has been outlined in (Michalczewski 1996) and in a graphical way it was presented on Fig. 6.

Before mutation

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	10	15
D2	0	0	15	20	0	0	5	5
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

After mutation

	D1	D2	TZ1	TZ2	TZ3	TZ4	O1	O2
D1	0	0	20	10	0	0	7	18
D2	0	0	15	20	0	0	8	2
TZ1	0	0	0	15	10	10	0	0
TZ2	0	0	0	0	20	25	0	0
TZ3	0	0	0	0	0	10	10	0
TZ4	0	0	0	0	0	0	10	35

Figure 6: Mutation process

4. CONCLUSION

The application SIMLock is the practical optimization tool which can be used in the warehouse location problems. It should be emphasized that genetic algorithms belong to heuristic algorithms. The solution generated by these algorithms for complex decision problems is a sub-optimal solution. However, considering the complexity of the warehouse location problem in the production companies, the solution is accepted from a practical point of view.

5. REFERENCES

[1] Akinc, U., Khumawala, B.M. (1977). An efficient branch and bound algorithm for the capacitated warehouse location problem, Management Science Vol. 23, Iss. 6: 585–594.

[2] Brandeau, M.L., Chiu S.S. (1989). An overview of representative problems in location research, Management Science, Vol. 35, No. 6 (1989): 645–674.

[3] Izdebski, M. (2014). The use of heuristic algorithms to optimize the transport issues on the example of municipal services companies, Archives of Transport, ISSN 0866-9546, vol. 29, iss.1, 2014, pp. 27-36.

[4] Izdebski, M., Jacyna, M. (2014). Use of computer assistance in order to designate the tasks in the municipal services companies, Journal of Kones, Vol.21, pp. 105-112, No.2.

[5] Jacyna M., Wasiak, M. (2015). Multicriteria Decision Support in Designing Transport Systems. In: Mikulski J. (eds) Tools of Transport Telematics. Communications in Computer and Information Science, vol 531. Springer, Cham.

[6] Jacyna-Gołda I. (2013). Chosen aspects of logistics network design method for production service companies. International Journal of Logistics Systems and Management 15(2/3), pp. 219–238.

[7] Jacyna-Gołda I. (2015). Decision-making model for supporting supply chain efficiency evaluation, Archives of Transport, ISSN 0866-9546, vol. 33, Issue 1, 2015, p. 17–31.

[8] Jacyna-Gołda I. (2015). Transport implementation model in supply chains, Proceedings of 19th International Scientific Conference. Transport Means, pp. 652-658.

[9] Jacyna-Gołda I., Izdebski M., Podviezko A. (2016). Assessment of the efficiency of assignment of vehicles to tasks in supply chains: A case-study of a municipul company. Transport, Volume 31, 2016, Issue 4, pp. 1-9.

[10] Jacyna-Gołda I., Izdebski M., Szczepański E., (2016). Lokalizacja obiektów magazynowych w sieci logistycznej – metody i narzędzia. Oficyna Wydawnicza Politechniki Warszawskiej.

[11] Jacyna-Gołda I., Izdebski M., Szczepański, E. (2016). Assessment of the Method Effectiveness for Choosing the Location of Warehouses in the Supply Network, Challenge of Transport Telematics Springer, Vol. 640, pp.84-97.

[12] Jacyna-Gołda, I., Lewczuk, K. (2017). The method of estimating dependability of supply chain elements on the base of technical and organizational redundancy of process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): p.382–392.

[13] Karkula, M. (2014). Selected aspects of simulation modelling of internal transport processes performed at logistics facilities, Archives of Transport, volume 30, issue 2, pp.43-56

[14] Michalczewski Z. (1996). Algorytmy genetyczne+ struktura danych = programy ewolucyjne, Wydawnictwo Naukowo –Techniczne.

[15] Özcan, T., Çelebi, N., Esnaf, Ş., (2011). Comparative analysis of multi-criteria decision making methodologies and implementation of a warehouse location selection problem, Expert Systems with Applications , Volume 38, Issue 8, :9773–9779.

[16] Pyza, D., Golda P. (2011). Transport Cargo Handling Shipments in Air Transport in the Aspect of Supply Chains, 2011 21st International Conference on Systems Engineering, Las Vegas, NV, pp. 442-445.

[17] Sharma, R. R. K., Berry,V. (2007). Developing new formulations and relaxations of single stage capacitated warehouse location problem (SSCWLP): Empirical investigation for assessing relative strengths and computational effort. European Journal of Operational Research, 177(2):803–812.

[18] Szczepański, E., Jacyna-Gołda, I., Murawski, J. (2014). Genetic algorithms based approach for transhipment HUB location in urban areas. Archives of Transport, ISSN 0866-9546, vol. 31, is. 3, p. 73-82.

[19] Żochowska, R. (2014). Selected issues in modelling of traffic flows in congested urban networks, Archives of Transport , volume: 29, issue 1. pp. 77-89.

Risks and Implications for Decision Making Processes Associated with Existing Design Codes or Their Non-Existence

Jan BLACHUT[1], Dariusz SALA[1]

[1]*AGH – University of Science and Technology, Kraków, Poland*

Abstract: Buckling phenomenon is a perplexing and unresolved issue in many safety critical structures, and it has been heavily regulated. The paper highlights the risks to decision making processes due to growing tendencies of eliminating from public domains disastrous events through confidentiality arrangements, erosion of existing human know-how, and falling standards of education. It is illustrated how existing, natural feedback routes to improvements of product design is broken by the imposition of legal, damage recovering, court proceedings. All of this can potentially lead the modern, automated support system to be blindfolded and unaware of harmful consequences when stability loss strikes.

Keywords: decision-making, product design

1. INTRODUCTION

Computer programs are nowadays responsible for a large portion of decision making processes. This applies to many industrial activities. Computer codes facilitate identification of processes, help to analyze and interpret data, as well as estimate consequences of decisions made. These decision support systems meant to cover all phases of decision making process, Refs [1, 2]. The origin of their development is usually associated with Carnegie Institute of Technology and MIT. The latter is credited with the development of first codes aimed at the automated/structured decision making, Ref. [2]. The current systems tend to apply modern techniques of data acquisition and processing. This includes active computer systems (self-learning, and/or adaptive), capable to adjust themselves to the need of a specific decision making requirement and environment. It is worth noting that such systems are capable of dealing with imperfect data like imprecision, lack of consistency, limited certainty, etc. They are also capable of delivering strategic recommendations as well as detailed operational reports for a given industrial entity based on their own data – see Fig. 1. There are also decision support systems which have on-line access to data and to analytical decision making models. This in turn can efficiently support decision processes especially when standard computer programs of lower ranking are either unable to find solution or it would be uneconomical to use them. A subset of computer support systems provides information for top company management personnel. These executive information systems, as they are called, gather the relevant data across all levels of company, sort it out, and prepares it in the format which explicitly can help managers to take efficient decisions on a day to day basis. A further level of executive information system is associated with partially processed data, some conclusions made, as well as partial interpretation and recommendations available. This kind of information system is aimed at top class of managers who make strategic decisions. Another form of decision support system is based on the use of business intelligence, Ref. [3]. This kind of systems uses data from all available sources as well as data about personnel's know-how and experience of all those who are involved in business decisions. The elements of feedback from consumers and partners are also included in this kind of system modus operandi – see Fig. 2.

The inclusion of feedback loop into the decision making has created a superior tool. Users become, for example, aware of past decisions, corrective actions taken and the sensitivities of adjustments. Fresh inputs can be added to the system as new technologies emerge or business environment changes. Bits of information related to corporation individuals, their involvement in company well being, creativity and motivations are also gathered. They constitute personal information about the company, and give managers a better appreciation of corporation workforce, Ref. [4].

A relatively new development in decision making processes is associated with data mining. Recovered data is usually subjected to on-line analytical processing which provides the instigator with multi-faceted analysis of data found.

At face value the above approaches can deal with imprecise, or incomplete data, and estimate the consequences of decisions made. But it is unclear how, if at all, the support systems can deal with some engineering activities where high level of safety is required due to the dangerous phenomenon associated with buckling, Refs [5, 6]. In cases like that design codes have been developed, and in many instances they had become legally binding. Virtually all countries have sets of their own rules, e.g., British Standards in the U.K. (BS), Industrial standards in Germany (DIN), or Polish standards (PN). One specific area where these standards are applied is related to pressure vessels. Despite advances in science, manufacturing, and materials these design codes still rely on phenomenological data obtained from experiments over the past decades, Ref. [5]. Pressure vessels are used across industries on land, under sea, and in the air. Their primary loading is usually internal or external pressure. As far as geometry is concerned they are built from cylindrical shells closed by domed ends. One of perplexing issues, and still not satisfactorily resolved, is large sensitivity of load carrying capacity to the initial geometric imperfections. This is particularly present in axially compressed cylinders and externally pressurised domed closures due to the possibility of buckling. In what follows the paper is to concentrate mainly on domed closures onto cylindrical vessels.

2. Scope

The paper aims to highlight buckling/collapse issues in pressurised domed closures. On one hand, the design codes meant to secure safe use of pressure components but arrival of new business opportunities requires the codes to be upgraded or to be developed. This however is a slow process and safety of vital engineering tasking can only be secured by rigorous application of tried and reliable methodologies. As advances in mechanics, computations and materials challenge the existing design by rule approaches, the paper is to identify some bottle necks which are likely to affect automated decision support processes.

3. Background

One of intense and important human activity, at the frontier of technological developments, is deep-sea exploration. It transpires that at the moment there is a quiet colonisation of oceans' sea bed. Whilst the extraction of oil and gas from sea bed has continued for several decades, it is the availability of rare raw material which drives the deep-sea research (cobalt, copper, manganese, silver, wolfram, gold, etc.). In many instances it is estimated that these materials can be extracted at 30 % of on land costs. There are only few nations which have the required, manned vehicles reaching the depth of 6 km – see Fig. 5 (with Chinese Jiaolong vehicle rated at 7 km just announced). The main technological obstacles are the availability of sufficient buoyancy units. These are shells of hemispherical, torispherical or elliptical shapes. Their strength is greatly affected by sensitivity to small deviations form perfect shape. Hence all experimental data, for metallic heads, has been collated and a lower bound was drawn to them. Next, a safety zone was proposed by lowering the design curve as shown in Fig. 4. For a while designers could rest assured that no failure occurred within the safety buffer zone. But this comfort has been removed by advanced computations backed by testing of industrial and laboratory shells as shown in Fig. 5, Ref. [7]. This implies that additional safety margin has to be considered for a specific product and its application. Industries have looked for new materials from which the buoyancy units could be made. One possibility was for domes being manufactured from composites (CFRP = Carbon Fibre Reinforced Plastics). But the results of these expensive research programmes are not instantly available in public domain – mostly due to their commercial/security sensitivity. Sponsors of research into novel structures frequently remove critical information from public domain. Good example is Ref. [8] where the supplement, containing critical experimental data and conclusions was only available to examiners during PhD viva and then made confidential.

4. Case studies

There are situations where no design code exist and structures have to be built in accordance with the best know-how at a time which sometimes does not secure a safe future use of a product as illustrated next.

4.1 Failure of rear bulkhead

Consider a wide body airplane L1011-385 which has been in service for decades and capable of carrying 400 passengers and 20 air crew. On a flight from Tenerife to Manchester in 1995, when the aircraft was approximately half-way between Tenerife and Faro, the flight crew heard a loud bang. As the cabin pressure started to change - the flight crew released oxygen masks and the aircraft immediately commenced an emergency descent. Apart from the pressurisation failure the aircraft had handled normally. During the transit to Faro, 4000 kg of fuel was dumped to reduce the aircraft landing weight to 160000 kg which then landed safely. As the decompression happened at the rear bulkhead – it was subject of careful investigation at Farnborough by Air Accident Investigation Board, Ref. [9]. The rear pressure bulkhead is a thin shell structure comprising a series of 1 mm thick panels lap-jointed to form a dome-shaped pressure membrane of approximately hemispherical profile. On passenger side the shell has been reinforced by a grid of stringers and rings – as seen in Fig. 6. The purpose of the reinforcing grid was to limit the size of any rupture which might occur. The apparent cause of the cabin depressurisation was the rupture of the skin within one grid cell. It has transpired that during each take-off the skin buckled and after landing it snapped back to normal shape. Large number of these 'in' and 'out' buckling snaps caused fatigue of the material and fracture. Luckily in this case the venting of cabin pressure did not damage other vital elements of the aircraft structure. Catastrophic structural failure which can follow from a pressure bulkhead failure, due to the release of pressurised cabin air into parts of the aft fuselage let to a major accident to a Japan Airlines Boeing 747 in 1985 with death of more than 500 people. The rear bulkhead of B747 had only stringer reinforcements, see Fig. 7, and when rupture happened the large flow of cabin air effectively disabled the aircraft. It has to be said that predicting buckling of a rear bulkhead under operational loading is difficult, and at a time when the above wide body aircrafts were designed, in sixties and early seventies of last century, there were no reliable tools for predicting this mechanism of failure. Equally surprising were crash investigator's recommendations in the case of B747 which merely stated 're-design the rear bulkhead' without giving any reasons behind the statement. Since the occurrence of the above two cases, a major effort was directed towards safety of rear bulkhead. Difficulty of predicting 'buckling dimpling' of a grid reinforced structure is also illustrated in Fig. 8. It shows glider's skin in a buckled condition (dimples) which disappear after another manoeuvre of the glider. The internal grid reinforcement secures, in this instance, the integrity of the wing. But after 'many' 'in' and 'out' buckling snaps the skin can rupture and this can lead to a free fall of the glider.

4.2 Failure of composite dome

CFRP domes have been researched for a possible use in deep-sea activities. As with each novel programme one had to address manufacturing processes and develop theoretical, predictive method for buckling/collapse of these pressure components. Fig. 9 depicts a collapsed 30-ply CFRP head made from draped carbon fabric, Ref. [6]. A question here arises, how the accumulation of the know-how in this research programme can efficiently be used in automated decision making tools ? One approach might be ‘wait and see’. But this equally may lead to the loss of business opportunities once others can take on board these vanguard developments.

4.3 Failure of imperfect dome

As mentioned earlier imperfections in domes seriously affect the magnitude of buckling strength (i.e., they lower it). Consider as an example, CFRP torisphere in which two types of shape imperfections exit. The first one is affine to eigenshape and the second corresponds to a local inward dimple. Computed results are depicted in Fig. 10. It is seen here that the amplitude of imperfection amounting to one wall thickness, t, reduces the load carrying capacity by 75 %. This is a very serious erosion of buckling strength. But shape imperfection are not the only one possible type. There are imperfections in material properties, boundary conditions, load application, or variation in wall thickness. The ongoing research here is unable to resolve this issue unequivocally. Hence accumulated know-how backed by selective experimentation is required for continuous upgrade of design codes. It is worth noting here that NASA – Ref. [11], for example, is still using design rule for imperfect structures which were developed in sixties of last centenary.

4.4 Pictures of failed industrial vessels

Reference [12] provides a number of failed, industrial structures where buckling was the responsible mode. It is seen in Fig. 11 how application-wide, and costly these failure can be.

5. Closure and conclusions

It is clear from the above that engineering activities where buckling is a possibility constitutes a challenging environment for decision making processes. Within the area of existing design codes one can accept solutions with a reasonable confidence although they might at first sight look too conservative. It is essential here that one has to carefully adhere to the recommended envelopes of shape, loading, boundary conditions, materials, etc. Here however it has been recently argued that modern computing tools can be used to replace the existing rules (mostly based on experimentation). The term ‘design by analysis’ has been coined, Ref. [13]. It appears that this route has to be used very carefully, and only by skilled personnel. At the moment it is unclear how future of this strategy will unfold. Saying that, decision processes will have to take notice of this development.

At the frontiers of human activities (deep-sea exploration, missions to Mars, nano-technology) where definitive business answers are fuzzy, it is felt that active monitoring should suffice for the moment.

6. Figures

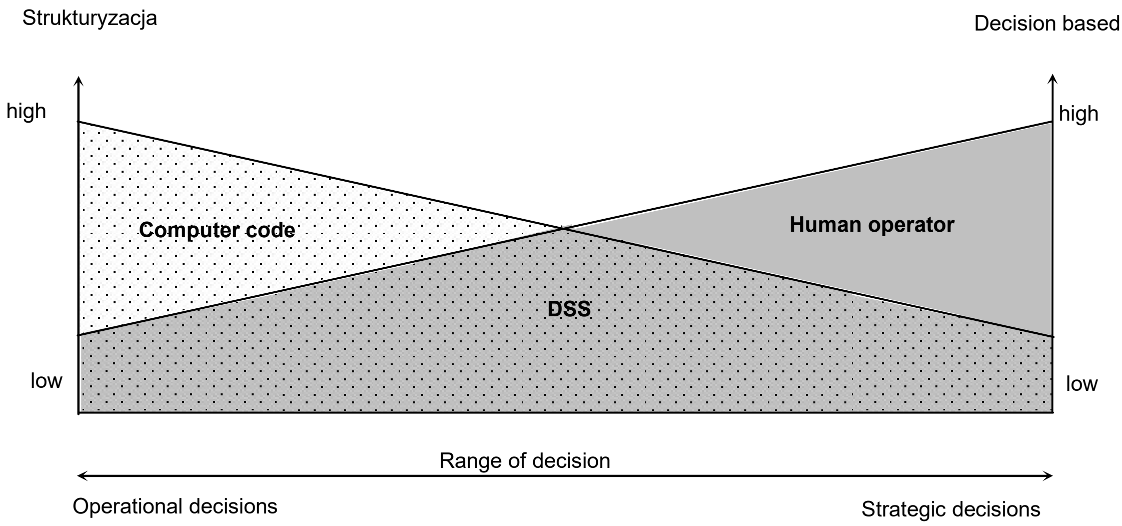


Fig. 1: Scheme of decision making processes and their interaction.

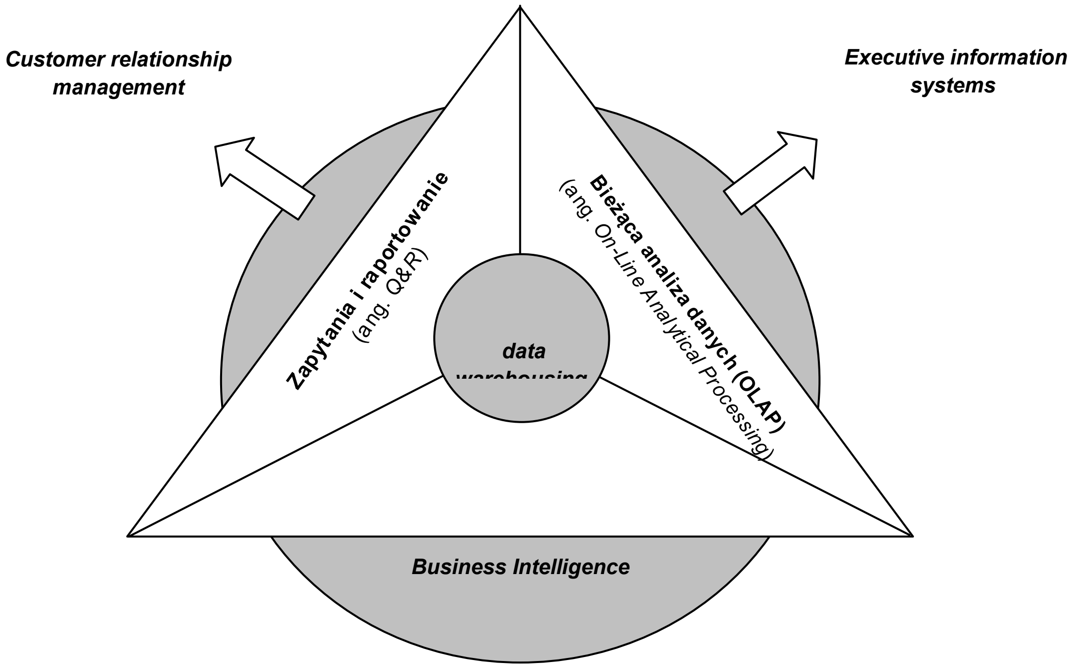


Fig. 2: Illustration of information systems and their relative positioning.

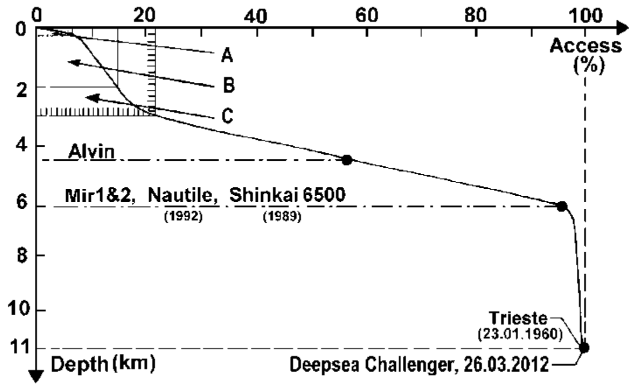


Fig. 3: Access to sea bed versus exploration depth (current and planned). Available vehicles are also denoted. (A ≡ current, North Sea; B ≡ current, Gulf of Mexico, costal waters of Brazil & Nigeria; C ≡ planned, researched).

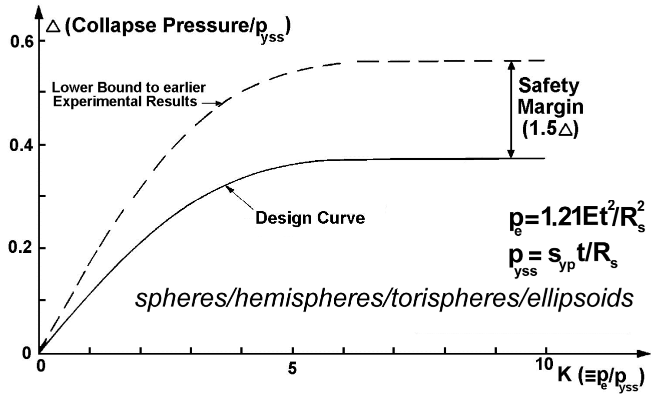


Fig. 4: Lower bound to known experimental data, and design curve with the safety margin.

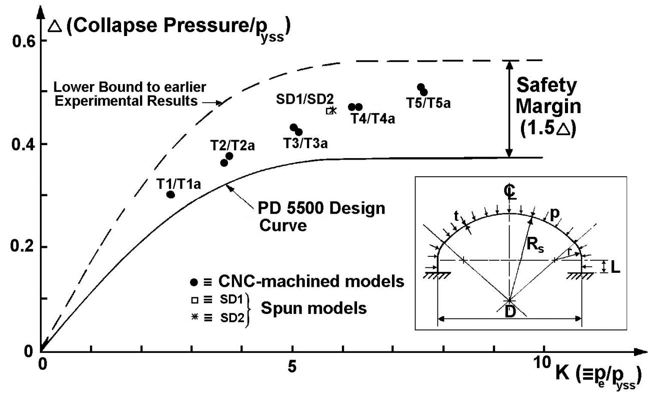


Fig. 5: Illustration how safety margin has been compromised by experimental test data.

Fig. 6: View of rear bulkhead in a wide body plane L1011 being reinforced by stringers and rings on internal, passenger side.

Fig. 7: Fragments of recovered rear bulkhead JAL B747 plane in Tokyo airport museum.

Fig. 8: View of dangerous inward dimples in wings of a glider (with permission of Dr D. Bushnell, Lockheed Martin, Palo Alto, Ref. [10]).

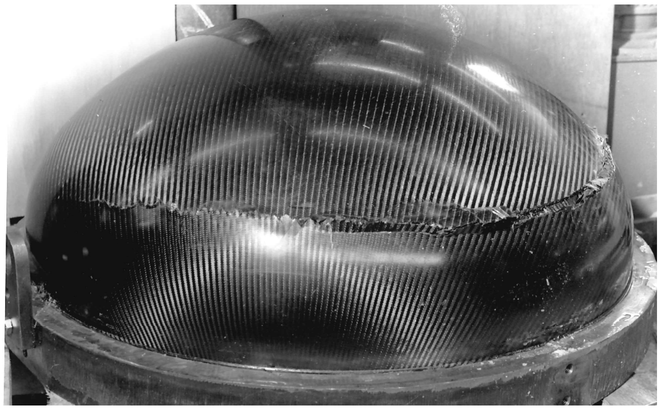


Fig. 9: Photograph of failed 0.8m diameter CFRP dome.

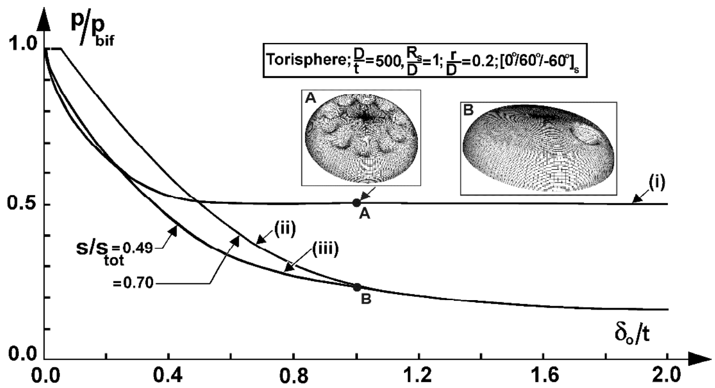


Fig. 10: Sensitivity of buckling pressure, p, to the amplitude of inward shape deviation, δ.

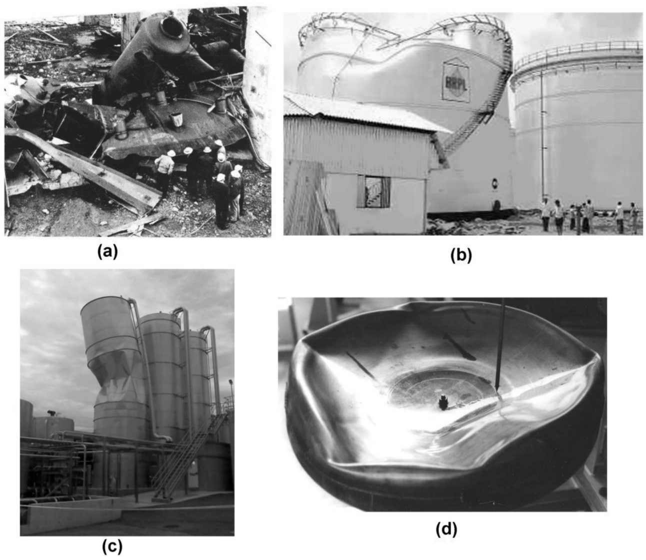


Fig. 11: A sample of structures failed due to insufficient buckling strength.

7. REFERENCES

[1] P.G.W. Keen, M. Scott Morton, *Decision Support Systems: An Organizational Perspective*. Reading, MA: Addison-Wesley, 1978.
[2] D.J. Power, *A Brief History of Decision Support Systems,* DSS Resources, World Wide Web, http://dss.cba.uni.edu/dss/dsshistory.html, 1999.
[3] M. Biere, *Business Intelligence for the Enterprise*, Prentice Hall PTR/IBM Press, June 2003.
[4] Z. Szyjewski, *Zarzadzanie wiedza korporacyjna*, www.placent.pl, Placent, 2002.
[5] J. Singer, J. Arbocz, T. Weller, "Buckling experiments – Experimental Methods in Buckling on Thin-Walled Structures – Volume 2", John Wiley & Sons, NY, 2002.
[6] J. Błachut, "Experimental perspective on the buckling of pressure vessel components", Applied Mechanics Reviews, Transactions of the ASME*,* vol. 66, 2014, 1-51.
[7] J. Błachut, "Buckling of sharp knuckle torispheres under external pressure", Thin- Walled Structures, vol. 30, 1998, 55-77.
[8] J. Didier, "Etude du comportoment au flambage des coques cylindriques multicouches metal/materiau mousse sous chargements combines pression interne/cisaillement/flexion", PhD Thesis, 2014, INSA, Lyon, France.
[9] Air Accident Investigation Branch UK, AAIB, Bulletin No. 2/96, EW/A95/6/1, 1996, 34-39.
[10] D. Bushnell, W.D. Bushnell, www.shellbuckling.com, 2017.
[11] NASA, "Buckling of Thin-Walled Doubly Curved Shells," NASA Space Vehicle Design Criteria, 1969, Report No. NASA SP-8032.
[12] C. Catriona de Paor, 'The effect of random geometric imperfections on the buckling of thin cylindrical shells due to external pressure", PhD Thesis, 2010, University of Cork, Ireland.
[13] J.M. Rotter, "Challenges and their resolution in both philosophy and process to exploit advanced computation in shell structure design", in Proc. of SSTA Conf., Shell Structures – Theory and Applications, 2-9 October 2017, Gdansk, (ed.) W. Pietraszkiewicz, CRC Press Taylor & Francis, London, pp. 1-10.

A MILP Model for Capaciteted Vehicle Routing Problem in Municipal Services with Time Windows and Limited Resources

Antoni KORCYL[1], Roger KSIĄŻEK[1], Katarzyna GDOWSKA[1]

[1]*AGH University of Science and Technology, Faculty of Management, Department of Operations Research, Kraków, Poland*

Abstract: One of responsibilities of municipal administration is to organize and to schedule waste collection services, so that the total costs of service is minimized while ensuring the desirable level of service quality. Once all the clients in the system must be provided with the service, the quality level can be understand as the number of customers served accordingly to their preferences and constaints. Optimization problem of setting waste collection routes is a special case of capacited Vehicle Routing Problem with time windows and backhauls, where following constraints should be included: clients must be served accordingly to their predefined time windows, available vehicles of the fleet can go along defined network segments only, the number and capacity of depots and landfills cannot be exceeded, each vehicle can be assigned to at least one depot (the vehicle starts and ends its route in the depot to which the given garbage truck is assigned). The problem needs to include also time windows defined for the availability of garbage trucks and their drivers. After collecting waste from the last client in the route and before going back to the depot each vehicle must be drained at a landfill. While planning the route of a vehicle, we should choose landfill that is located as close as possible to both the last client in the route and the vehicle's home depot, so that the total distance is reduced. Another specific aspect of this case of VRP is that the system is served by a heteronomous fleet of garbage trucks, where the number of available cars of each type is predefined, as well as the number of hours of their availability. Scarcity of *resources influences significantly the achieved quality level, because network segments have their specific characteristics (e.g. width), so that they can be served by a specific type of vehicles only. In research to which this abstract refers, a new MILP model for capaciteted vehicle routing problem in municipal services with time windows and limited resources was developed and tested for small and medium-sized instances of muli-depot waste collection system with time windows and scarce resources. Developed problem belongs to the class of VRPs, so it is a NP-hard one. Applicability of the model can be enhancd by developing tools using heuristics algorihms based on the newly developed MILP model.*

Keywords: MILP, VRP, capacitated VRP, time windows, municipal services

A MILP Model for Assignment Problem with Time Windows for Loading Parcel Lockers

Roger KSIĄŻEK[1], Katarzyna GDOWSKA[1,2]

[1]*AGH University of Science and Technology, Faculty of Management, Department of Operations Research, Kraków, Poland*

[2]*INESC TEC, Centre for Industrial Engineering and Management, Porto, Portugal*

Abstract: Last mile delivery grows with the increasing interest in remote shopping, but it lso results in heavy congestion of commercial vehicles in the whole city area. In order to reduce the number of unattended home deliveries, some on-line shopper may choose delivery to a parcel locker. Locker-banks are groups of reception box units (lockers), which are similar to collection points although they are not sited at each customers premise but sited in apartment blocks, work places, car parks, railway stations etc. Customers may be notified by message about when their delivery has arrived, the box number and location, and the code to open the box. Locker-banks require the customer to make the final leg of the journey. In research to which this abstract refers a MILP model for Assignment Problem with Time Windows for Loading Parcel Lockers was considered. The deterministic static model finds an optimal solution for the moment when we decide to assign parcels-to-be-shipped to available lockers in a parcel station. If a suitable locker for a given parcel is not available delivery cost on this parcel grow. If a parcel is not delivered on time the company need to compensate inconvenience to the client. Basing on the history of users behavior it is possible to assess probability of the customer's appeareance at the parcel locker and receiving the parcel at the beginning or at the end of the time window duing which the parcel waits in the locker, so that we can or cannot decide on shifting the moment of loading the parcel to the locker-station. The model minimizes the total costs including delivery cost and the compensation costs connected with the absence of the parcel due to shifting it to another locker station or delaying the moment of loading it to the station. The approach to the problem may be based on event-based rolling horizon, where evens are subsequent moments of loading locker-sations.

Constraints that guarantee delivery and matching dimentions of the parcels and lockers must be satisfied. The model optimize a static decision on loading a locker-stations.
This work was supported by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme within project «POCI-01-0145-FEDER-006961», and by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia as part of project "UID/EEA/50014/2013". This work was supported by AGH statutory fund – grant No. 15/11.200.357 and grant No. 15/11.200.356

Keywords: MILP, last-mile delivery, parcel lockers, pop stations, city logistics, assignment problem, time windows

Multiple-Period Interval Synchronization Problem in Public Transportation

Katarzyna GDOWSKA[1]

[1]*AGH University of Science and Technology, Faculty of Management, Department of Operations Research, Kraków, Poland*

Abstract: Multiple-period interval synchronization at long overlapping route segments of an urban public transport network is an important issue in public transport networks with long radial routes. Interval synchronization aims to set departure time of every trip of every line, so that time gaps between arrivals of consecutive trips of different lines at shared stops are equalized. Ride frequency of public transport is adjusted to passenger flows, so it is a common practice to set different ride frequencies to smaller planning periods (morning peak, valley hours, afternoon peak etc.). The duration of these periods does not have to be equal. For each period headways, travel times and number of trips to be executed are specified with respect to travel service needs. Discrepancies between headways adopted in consecutive periods affect significantly departure times and – in consequence – arrival times at bus stops, so that the objective of multiple-period interval synchronization is to smoothen transitions between periods, which means to guarantee that every each line at every bus stop separation time between the last trip of the previous period and the first trip of the next period has to fit some range. Multiple-period interval synchronization in urban public transport can be formulated for a system with fixed or flexible headways. Several approaches to multi-period synchronization of public transport are examined in the substantial literature. Depending on the type of headway setting the range for separation time is different and it can be either included in a model for synchronization problem or be solved as a independent sub-problem. In the paper different approaches to period length setting are surveyed focusing on the number of planning periods, differences in their length, etc. A multiple-criteria mixed-integer linear programming model for the multiple-period interval synchronization problem in public transportation is presented and selected results of computational

experiments conducted for small instances are reported and discussed .
This work was supported by AGH statutory fund – grant No. 15/11.200.357.

Keywords: interval synchronization, public transportation timetabling

Decision-Making Problems in Last-Mile Delivery with Crowdshippment

Katarzyna GDOWSKA[1,2]

[1]*AGH University of Science and Technology, Faculty of Management, Department of Operations Research, Kraków, Poland*

[2]*INESC TEC, Centre for Industrial Engineering and Management, Porto, Portugal*

Abstract: Last mile delivery grows with the increasing interest in remote shopping, but it lso results in heavy congestion of commercial vehicles in the whole city area. Different business models involving crowdsourcing are to be observed in the area of transport and logistics. In last mile delivery there are business models enabling self-employed people to treat completing crowdsourcing tasks as a full time job (e.g. Zipments) as well as concepts focused on sustainable mobility end pollution reducing and for that reasons involving ad-hoc drivers in delivering parcels. Nowadays crowdsourcing becomes so popular and it plays such an important role in transport and logistics that it became absolutely normal to use such terms as "crowd logistics", a "crowd logistics company" or "crowdshipping". The use of mobile crowdsourcing for building various applications and services for smart city is also considered. The paradigm of centrally-coordinated crowdsourcing system is discussed and its applications to smart campus monitoring and last-mile urban logistics (package pickup and delivery). However, developing new business models for real-life intelligent transport systems including customers and multiple stakeholder perspectives pickup and delivery). The crowdsourced delivery problem is a specific manifestation of a general transportation problem which is well-known also as pickup and delivery problem (PDP) that aims to transport goods from origins to destinations at minimum cost – the problem was formulated as dynamic pickup and delivery problem (DPDP); pickup and delivery problem with single depot; pickup and delivery problem with single depot and intermediate warehouses/hubs; pickup and delivery problem with time windows. Dealing with processes involving crowdsourcing needs constant developing methods dedicated for sensing, interfering, predicting and guiding such processes. The specific *features of crowdsourse-based problems are also considered as a opportunity to leverage methods and results from , among others, graph theory, optimization, and machine learning. Another important issue for matching tasks with delivery partners and crowdsourced delivery routing; the issue is how to collect data for the time dependent VRP and proposes a heuristic for that problem. The model is shown to provide a boost in efficiency as it takes into account phenomena of traffic jam and congestion in the city. A VRP heuristic can be used to solve the model, which is a one-transshipment system where a fleet of short-haul trucks are dispatched from a depot and unload packages at terminals of service zones where passenger cars with available mobility are attracted to pick up and deliver packages to demand destinations. The analysis of the results concludes that sharing logistics can be beneficial under certain conditions, which are: start with low-demand-density areas, deliver during off-peak hours (to pay lower wages to the drivers) and increase car loads. The prospects of this business model is grim if the number of drivers available for parcel delivery does not increase so that the wages do not increase along with the quantity of goods to deliver, given that there is no economy of scale for models based on occasional drivers. Another approach considers a model for a last-mile delivery within urban area using trucks, and enabling last-leg delivery(or first-leg pickup) to be croudsoursed. Delivery partners submit their bids to the truck carrier, maximizing their benefit. Truck carrier then selects the bids and coordinates crowdsourced last-leg delivery (first-leg pickup) with its truck operations. The considered truck carrier problem includes selection of bids, determine crowdsource-truck relay points, truck routs and schedule. The problem is a nonlinear integer problem, and solution approach is proposed. Computational experiments performed on test instances generated based on real data.*

This work was supported by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme within project «POCI-01-0145-FEDER-006961», and by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia as part of project "UID/EEA/50014/2013".

Keywords: MILP, last-mile delivery, parcel lockers, pop stations, city logistics, assignment problem, time windows

Jubileuszowe XX Sympozjum
Zastosowania Teorii Systemów
2017

Materiały konferencyjne

Wykorzystanie własności krajobrazu przestrzeni rozwiązań w konstrukcji algorytmów metaheurystycznych

Wojciech BOŻEJKO[1], Andrzej GNATOWSKI[1], Czesław SMUTNICKI[1], Mieczysław WODECKI[2]

[1]*Politechnika Wrocławska*

[2]*Uniwersytet Wrocławski*

Streszczenie: W efektywnych algorytmach optymalizacyjnych (dokładnych i przybliżonych) korzysta sie zazwyczaj z wielu, często specyficznych dla danego problemu własności. Omawiane problemy nie maja klasycznych własności analitycznych, takich jak: wypukłość, spójność, różniczkowalność i w większości maja bardzo dużo ekstremów. Ponadto moc przestrzeni (względem rozmiaru danych) rośnie bardzo szybko („eksplozja wykładnicza”), co wyklucza stosowanie metod przeliczeniowych. Istnieje wiec konieczności intensywnego badania krajobrazu przestrzeni (fitness landscape) opisującego zależności pomiędzy wartościami funkcji celu a odległościami rozwiązań w przestrzeni. W pracy proponujemy wykorzystanie własności krajobrazu konkretnej instancji problemu (na podstawie wstępnego próbkowania, a także z danych pozyskanych w trakcie działania adaptacyjnego algorytmu metaheurystycznego) w celu efektywniejszego niż w dotychczasowych metodach rozwiązywania trudnych problemów optymalizacji dyskretnej.

Słowa kluczowe: optymalizacja dyskretna, analiza przestrzeni rozwiązań, metaheuristyki.

1. Wprowadzenie

Dla wielu problemów optymalizacji dyskretnej (np. TSP, QAP, flow shop, job shop, itd.) rozwiązania mogą być reprezentowane przez permutacje.

Dla n-elementowego zbioru $I = \{1, 2,..., n\}$, tj. elementów ponumerowanych kolejnymi liczbami od 1 do n, przez Ω oznaczmy zbiór wszystkich permutacji elementów z I. Rozpatrujemy problem (w skrócie OP):

$$\min\{F(\alpha) : \alpha \in \Omega\},$$

gdzie funkcja celu $F : \Omega \rightarrow R^{+}$. Jeżeli problem OP jest NP-trudny, wówczas do jego efektywnego rozwiązania (szczególnie przykładów o praktycznych rozmiarach) stosuje sie algorytmy przybliżone, które nie gwarantuje wyznaczenia rozwiązania optymalnego. Obecnie do najlepszych należą algorytmy oparte na metodzie lokalnego poszukiwania, polegającej na iteracyjnym przeszukiwaniu pewnych podzbiorów zbioru rozwiązań — otoczeń *Otoczenie* jest to odwzorowanie:

$$N : \Omega \rightarrow 2^{\Omega},$$

przyporządkowujące każdemu elementowi $a \in \Omega$ pewien podzbiór $N(a)$ zbioru rozwiązań dopuszczalnych Ω, $N(\alpha) \subseteq \Omega$.

Liczba elementów otoczenia, sposób ich wyznaczania oraz przeglądania ma decydujący wpływ na efektywność (czas obliczeń i wartość funkcji celu wyznaczonego rozwiązania) algorytmu opartego na metodzie lokalnego poszukiwania. Klasyczne otoczenia, stosowane od kilkudziesięciu lat zawierają wielomianowa liczbę elementów i są generowane przez przekształcenia zwane powszechnie *ruchami,* tj. „niewielkimi" zmianami pewnych elementów permutacji polegającymi na zamianie pozycjami dwóch elementów (swap) lub przestawieniu elementu na inna pozycje (insert) w permutacji.

2. Przegląd literatury

Jak pokazał Wolpert and Macready w no free lunch theorem [19] aby uzyskiwać lepsze rezultaty konieczne jest wykorzystywanie własności specyficznych dla problemu. I tak, jeden z najlepszych klasycznych algorytmów TSAB [15] zaprojektowany do rozwiązywania problemu gniazdowego szeregowania żądań (Job Shop Scheduling Problem, JSSP) korzysta z tzw. własności blokowych, pozwalających na ograniczenie rozmiaru przeglądanego otoczenia tylko do „obiecujących" rozwiązań Niestety identyfikacja przez badacza przydatnych cech problemu nie zawsze jest możliwa. Co więcej, część z nich może być właściwa tylko dla konkretnych instancji, jako ze efektywność algorytmów najczęściej uzależniona jest struktury przestrzeni rozwiązań. Głównym narzędziem służącym do jej badania jest analiza krajobrazu przestrzeni rozwiązań (Fitness Landscape Analysis, FLA) [5], w kontekście której zdefiniowano szereg miar określających cechy krajobrazu funkcji celu, jak np. chropowatość (ruggedness, [18]), neutralność (neutrality, [16]), istotna w kontekście algorytmów genetycznych epistazę (epistasis, [14]) czy rozkład funkcji celu ({it fitness distribution, [17]). Lista nie jest zamknięta, nadal wprowadzane są nowe definicje miar, jak choćby miary w dziedzinie częstotliwości dla dynamicznych problemów optymalizacji [12]. Bardziej szczegółowy przegląd miar krajobrazu można znaleźć w [6,12]. Jedna z własności problemów optymalizacji dyskretnej, w tym JSSP, identyfikowalna za pomocą FLA jest własność Wielkiej Doliny (Big Valley). Mówi ona o tym, ze istnieje istotna pozytywna korelacja wartości funkcji celu rozwiązania i jego odległości od optimum globalnego.

Oprócz prób znajdywania własności problemów, w literaturze FLA stosowane było miedzy innymi do: szacowania trudności instancji [13,17], wyboru odpowiedniego algorytmu przeszukiwania przestrzeni [8], czy algorytmu sterującego parametrami [2]. Zgodnie z taksonomią z [4], sterowanie parametrami można podzielić na parametr tuning, gdzie dobór parametrów odbywa sie przed uruchomieniem właściwego przebiegu algorytmu; oraz parametr control, gdzie proces przebiega on-line. Pośród rozwiązań opierających sie o sterowanie parametrami, wymienić można ParamlLS [7] lub iterated racing (irace) [11] w których to na podstawie przebiegów algorytmu na instancjach ze zbioru uczącego, dobierane sa nastawy. W pracy [2] użyto z kolei parametr tuning w połączeniu z FLA jako elementu algorytmu memetycznego. W kolejnych iteracjach obliczone sa cztery miary: evolvability measure, two neutrality measures and a fitness distribution measure. Na ich podstawie, za pomocą kilku strategii i algorytmu uczenia online dynamic weighted majority [9] dokonywany jest wybór operatora ewolucyjnego.

Interesujące spojrzenie na dobór parametrów algorytmu, łączące cechy uczenia on-line i off-line, przedstawiono w pracy [10]. FLA użyte jest tam do podziału instancji uczących problemu na grupy. Dla każdej z nich, do znalezienia możliwie najlepszej konfiguracji, używany jest algorytm irace. Następnie przed rozwiązywaniem konkretnej instancji problemu, ponownie wykonywana jest analiza FLA i na tej podstawie wybierana jest odpowiednia konfiguracja.

3. Grafy oraz generatory otoczeń

Niech $M(\alpha)$ będzie zbiorem ruchów generujących otoczenie pewnej permutacji a $\alpha \in \Omega$. Dowolny ruch $m \in M(\alpha)$ przekształca $\alpha \in \Omega$ w nowa permutacje $\beta = m(\alpha) \in N(\alpha)$. Niech $\gamma = \alpha^{-1}\beta$, gdzie α^{-1} jest permutacją odwrotną do permutacji α. Wówczas ruch $m \in M(\alpha)$ generujący z α permutacje β, może być utożsamiany z permutacja γ.

$$m(\alpha)=\beta=\alpha\gamma, \text{ gdzie } \gamma \in \Phi_n$$

Podobnie, zbiór ruchów $M(\alpha)$ można utożsamiać ze zbiorem permutacji

$$G(\alpha)=\{\beta \in \Omega : \exists m \in M(\alpha)(m(\alpha)=\alpha\beta)\}$$

Zbiór permutacji $G(\alpha) \subseteq \Omega$ odpowiadający zbiorowi ruchów $M(\alpha)$ jest nazywany generatorem otoczenia permutacji α. Wówczas

$$N(\alpha)=\{\alpha\beta : \beta \in G(\alpha)\}$$

Jeżeli $\forall\alpha,\beta(\alpha \neq \beta)$ zachodzi *G(α)* = *G(β)*, to zbiór permutacji *G* = *G(α)* będzie nazywany **generatorem otoczeni** *N*. Jedna z metod definiowania otoczenia jest określenie jego generatora (podzbioru zbioru wszystkich permutacji O).

Dla dowolnego otoczenia *N*, definiujemy graf otoczenia *G(N)* = (V, *A)*. Zbiór jego wierzchołków $V = \Omega$, a zbiór łuków

$$A=\{(\alpha,\beta): \beta\in N(\alpha),\ \alpha\in\Omega\}$$

Ponieważ każdy element generatora *G* można utożsamiać z pewnym ruchem, stad elementy otoczenia są generowane przez wykonanie na n pojedynczych ruchów. Tak wiec elementy z *N* są odległe o jeden (ruch) od *n*. Z kolei, jeżeli $\delta\in N(\pi)$, to elementy otoczenia permutacji δ (tj. z $N(\delta)$) są odległe od π o co najwyżej dwa ruchy. Proces ten można kontynuować. Mając generator, można wiec określić „odległość" (liczbę ruchów) pomiędzy elementami przestrzeni. Nie wszystkie przedstawione w tym rozdziale ruchy i złożenia ruchów są bezpośrednio miarami określonymi na zbiorze permutacji. Do analizy odległości (pomiędzy para dowolnych permutacji $\pi,\ \sigma\in\Phi_n$) najczęściej stosowane są następujące miary (Diaconis [3]):

1. Miara tau Kendalla (ang. Kendall's tau):

   $I(\pi, \alpha)$ — minimalna liczba zamian par przyległych, przeprowadzając permutacje π^{-1} w σ^{-1}.
2. Miara Cayleya (ang. Cayley's distance):

   $T(\pi, \sigma)$ — minimalna liczba transpozycji, jakie należy wykonać przeprowadzając permutacje *n* w permutacje σ.
3. Miara Ulama (ang. Ulam's distance):

   $L(\pi, \sigma)$ - długość najdłuższego rosnącego podciągu w iloczynie per- mutacji $\sigma \bullet \pi^{-1}$.

Wymienione miary maja, bezpośredni związek z typowymi ruchami definiowanymi na permutacjach i stosowanymi obecnie w generowaniu otoczeń. Miara tau Kendalla jest związana z wykonywaniem ruchów typu zamień na sąsiednich elementach w permutacji. Tak generowane otoczenie ma n-1 elementów, a algorytm obliczania wartości tej miary ma złożoność obliczeniową $O(n^2)$. Miara Cayleya odpowiada liczbie ruchów typu zamień przeprowadzających dowolny element przestrzeni w dowolny inny. Otoczenie generowane przez te ruchy ma $n(n-1)/2$ elementów, a złożoność obliczeniowa wyznaczania tej miary wynosi $O(n)$. Z kolei miara Ulama ma ścisły związek z ruchami typu wstaw.

Niech *G(N)* = (V, *A)* bedzie grafem pewnego otoczenia *N*, a *d* miara odległości pomiędzy jego wierzchołkami (waga, krawędzi $(i,j) \in A$ jest odległość pomiędzy wierzchołkami *i* i *j*). Ciąg wierzchołków $(v_s, v_{s+1},..., v_{t-1}, v_t)$, $t > s$, $v_i \in V$, $i = s, s + 1,...,t$ jest drogą, z wierzchołka v_s do v_t, jezeli $(v_i, v_i+1) \in A$ dla $i = s, s + 1, \ldots, t - 1$. Długość drogi jest równa sumie wag krawędzi. Graf *G(N)* jest spójny, jeśli dla każdej pary różnych wierzchołków $u \neq v$, $u,v \in V$ istnieje droga z *u* do *v*.

Odległość $d(u, v)$ z wierzchołka *u* do wierzchołka *v* w *G* jest równa długości najkrótszej drogi z *u* do *v*. Dla wierzchołka $v \in A$, niech

$$d_v = \max\ \{d(v, u) : u \in A\}$$

będzie maksymalna odległością z wierzchołka v w grafie *G*. Wówczas

$$diam(G) = min\{d_v : v \in \mathrm{A}\}$$

nazywamy średnicą grafu, a wierzchołek $v°$ dla którego d_{v^o} = *diam(G)* — wierzchołkiem centralnym.

4. Empiryczne badanie przestrzeni rozwiązań

Rozpatrujemy pewien przykładowy problem OP, w którym G jest generatorem otoczeń, $G = (V, A)$ grafem otoczeń, a d pewną miarą na zbiorze rozwiązań dopuszczalnych (permutacji) Ω. Poniżej opisujemy statystyczna metodę badania przestrzeni rozwiązań, której celem jest ustalenie położenia minimów lokalnych i ostatecznie ograniczenie obszaru poszukiwań optymalnego rozwiązania problemu. Na bazie przeprowadzonych eksperymentów obliczeniowych przyjęliśmy następujące założenia związane z hipoteza istnienia Wielkiej Doliny:

1. minima lokalne są skoncentrowane w pewnym obszarze przestrzeni rozwiązań,
2. zależność pomiędzy odległością od rozwiązania optymalnego, a liczba minimów ma rozkład normalny.

Proces ustalania obszaru poszukiwań można przedstawić następująco.

Próbkowanie przestrzeni rozwiązań. Wyznaczyć losowo zbiór punktów przestrzeni rozwiązań

$$P = \{\pi_1, \pi_2,..., \pi_s\}$$

dla ustalonej stałej liczebności zbioru ekstremów lokalnych s. Stosujac algorytm typu popraw (np. tabu search), startując kolejno z rozwiązań ze zbioru P, wyznaczyć (np. stosując obliczenia równolegle) zbiór minimów lokalnych

$$LM^* = \{\pi_1^*, \pi_2^*,..., \pi_s^*\} \tag{1}$$

Definiujemy nowy graf $G^* = (V^*,A^*)$, gdzie $V^* = LM^*$, a $A^* = \{(u, v) : u,v \in V^*\}$. Oczywiście, $A^* \subseteq A$ oraz $V^* \subseteq V$ (G^* jest podgrafem grafu otoczeń G). Miarę d^* odległości wierzchołków w G^* definiujemy następująco: $d^*(u,v) = d(u,v),\ u,v \in V^*$.

Empiryczny rozkład minimów lokalnych. Wyznaczamy wierzchołek centralny $v°$ w grafie G^*. Obliczamy odległości $d^*(v°,v),\ v \in A^*$ pomiędzy wierzchołkiem centralnym, a pozostałymi minimami lokalnymi ze zbioru LM^*. Obliczamy średnia μ oraz odchylenie standardowe σ z tych odległości. Następnie, badamy położenie minimów lokalnych względem centrum $v°$, tj. zależność pomiędzy odległością od centrum, a liczba minimów lokalnych. Weryfikujemy hipotezę H_0: elementy zbioru minimów lokalnych LM^* ze względu na miarę d^* maja rozkład normalny $N(\mu, \sigma)$. W przypadku pozytywnej weryfikacji, przeszukiwanie przestrzeni rozwiązań ograniczamy do rozwiązań, które są odległe od $v°$, nie więcej niż $\delta \bullet \sigma$ (promień obszaru poszukiwali), gdzie δ jest pewnym, ustalanym eksperymentalnie parametrem. Do przeszukiwania tego obszaru będą stosowane algorytmy metaheurystyczne.

Wyniki przeprowadzonych eksperymentów obliczeniowych przedstawiono w pracy [1].

5. Literatura

[1] Bożejko, W., Gnatowski, A., Uchronski, M., Wodecki, M.: Metoda lokalnych poszukiwan na ograniczonym obszarze przestrzeni rozwiazan. Materiały XX Sympozjum Zastosowania Teorii Systemów, Zakopane, 26–30 wrzesien 2017 (2017).

[2] Consoli, P.A., Mei, Y., Minku, L.L., Yao, X.: Dynamic selection of evolutionary operators based on online learning and fitness landscape analysis. Soft Computing 20(10), 3889–3914 (2016).

[3] Diaconis, P.: Group representations in probability and statistics. Lecture notes – monograph series, Institute of Mathematical Statistics (1988).

[4] Eiben, A.E., Michalewicz, Z., Schoenauer, M., Smith, J.E.: Parameter Control in Evolutionary Algorithms. In: Parameter Setting in Evolutionary Algorithms, pp. 19–46. Springer Berlin Heidelberg, Berlin, Heidelberg (2007).

[5] Humeau, J., Liefooghe, A., Talbi, E.G., Verel, S.: ParadisEO–MO: From fitness landscape analysis to efficient local search algorithms, vol. 19 (2013)

[6] Humeau, J., Liefooghe, A., Talbi, E.G., Verel, S.: ParadisEO–MO: from fitness landscape analysis to efficient local search algorithms. Journal of Heuristics 19(6), 881–915 (12 2013)

[7] Hutter, F., Leyton–brown, K.: ParamILS : An Automatic Algorithm Configuration Framework. Journal Of Artificial Intelligence Research 36, 267–306 (2009).

[8] Jana, N.D., Sil, J., Das, S.: Selection of appropriate metaheuristic algorithms for protein structure prediction in AB off-lattice model: a perspective from fitness landscape analysis. Information Sciences 391–392, 28–64 (2017).

[9] Kolter, J., Maloof, M.: Dynamic weighted majority: a new ensemble method for tracking concept drift. In: Third IEEE International Conference on Data Mining. pp. 123–130. IEEE Comput. Soc (2003).

[10] Liefooghe, A., Derbel, B., Verel, S., Aguirre, H., Tanaka, K.: Towards Landscape-Aware Automatic Algorithm Configuration: Preliminary Experiments on Neutral and Rugged Landscapes. In: Hu, B., M., L.I. (eds.) Evolutionary Computation in Combinatorial Optimization. EvoCOP 2017. Lecture Notes in Computer Science, vol 10197, pp. 215–232. Springer International Publishing, Cham (2017).

[11] Lopez-Ibaneż, M., Dubois-Lacoste, J., Perez Caceres, L., Birattari, M., Stutżle, T.: The irace package: Iterated racing for automatic algorithm configuration. Operations Research Perspectives 3(January), 43–58 (2016).

[12] Lu, H., Shi, J., Fei, Z., Zhou, Q., Mao, K.: Measures in the time and frequency domains for fitness landscape analysis of dynamic optimization problems. Applied Soft Computing 51, 192–208 (2017)

[13] Malan, K.M., Engelbrecht, A.P.: Fitness Landscape Analysis for Metaheuristic Performance Prediction. In: Richter, H., Engelbrecht, A. (eds.) Recent Advances in the Theory and Application of Fitness Landscapes, chap. 4, pp. 103–132. Springer Berlin Heidelberg, Berlin, Heidelberg (2014).

[14] Naudts, B., Suys, D., Verschoren, A.: Epistasis as a Basic Concept in Formal Landscape Analysis. Proceedings of the Seventh International Conference on Genetic Algorithms pp. 65–72 (1997).

[15] Nowicki, E., Smutnicki, C.: A fast taboo search algorithm for the job shop problem. Management Science 42(6), 797–813 (6 1996).

[16] Reidys, C.M., Stadler, P.F.: Neutrality in fitness landscapes. Applied Mathematics and Computation 117(2–3), 321–350 (2001).

[17] Rose, H., Ebeling, W., Asselmeyer, T.: The density of states — A measure of the difficulty of optimisation problems. In: Guervós, J.J.M., Adamidis, P., Beyer, H.G., Schwefel, H.P., Fernandez-Wlacanas, J.L. (eds.) Parallel Problem Solving from Nature — PPSNIV. PPSN 1996. Lecture Notes in Computer Science, vol 1141, vol. 2439, pp. 208–217. Springer Berlin Heidelberg, Berlin, Heidelberg (1996).

[18] Vassilev, V.K., Fogarty, T.C., Miller, J.F.: Information Characteristics and the Structure of Landscapes. Evolutionary Computation 8(1), 3160 (3 2000).

[19] Wolpert, D.H., Macready, W.G.: No free lunch theorems for optimi- zation. IEEE Transactions on Evolutionary Computation 1(1), 67–82 (1997).

Problem optymalizacji procesu replikacji danych z uwzględnieniem okien czasowych

Wojciech BOZEJKO[1], Piotr NADYBSKI[2], Mieczysław WODECKI[3]

[1] *Politechnika Wrocławska*

[2] *Państwowa Wyższa Szkoła Zawodowa im. Witelona w Legnicy*

[3] *Uniwersytet Wrocławski*

Streszczenie: *Wykonanie kopii zapasowych danych, raplik maszyn wirtualnych zwiększa niezawodność systemu. Nadawca, poprzez kanały o ograniczonej przepustowości rozsyła informacje do różnych odbiorców. Każda informacja ma ustalony rozmiar oraz najwcześniejszy i najpóźniejszy termin dostarczenia. W pracy jest rozpatrywany problem, w którym należy przydzielić informacje do kanałów oraz w każdym z nich ustalić kolejność wysyłania, aby łączna kara za niedotrzymanie terminów dostarczenia była minimalna.*

Słowa kluczowe: rozsyłanie w sieci, okna czasowe, algorytm przeszukiwania z tabu.

1. Wprowadzenie

W pracy rozpatrujemy system. w którym generowane i uruchamiane sa maszyny wirtualne. Każda maszyna musi zostać zreplikowana, w postaci zapasowej kopii. w innej lokalizacji dostępnej za pośrednictwem łacza (kanału) o ograniczonej przepustowości. Obraz każdej maszyny ma ustalony rozmiar. który determinuje czas transmisji. Każdą kopie należy wykonać, w z góry ustalonym oknie czasowym. Niedotrzymanie terminów (najwcześniejszego lub najpóźniejszego) wiąże się z kara, która zależy od z wielkości przekroczenia ustalonego terminu a oraz współczynnika kary. Specyfika pracy systemu nie pozwala na wykonywanie kopii on-line, dlatego tez zadania wykonania kopii zapasowej są kolejkowane i uruchamiane w chwili mniejszej aktywności użytkowników. Należy zatem przydzielić zadania wykonania kopii do odpowiednich kanałów oraz dla każdego z nich ustalic kolejność wysyłania tak, aby zminimalizować sumę kar za niedotrzymanie ustalonych terminów. W literaturze są rozpatrywane głównie problemy związane z kosztami przechowywania oraz ryzykiem niedostępności (utraty) danych (Mansuori i in. [3], Thanasis i in. [5]). Z kolei Hadij w pracy [2], przedstawiono algorytmy uwzględniania zarówno przepustowości sieci i opóźnieniami jak i koszty związanymi z przesyłaniem danych.

Z podobnymi zagadnieniami mamy do czynienia w systemach wytwarzania dokładnie na czas (ang. Just in Time), gdzie koszty powoduje nie tylko zbyt późne, ale także zbyt wczesne wykonanie zadania. Tego typu zagadnienia (w skrócie E/T) są inspiracja do sformułowania wielu NP-trudnych problemów szeregowania z nieregularnymi funkcjami celu. W praktyce do ich rozwiązywania są stosowane niemal wyłącznie algorytmy aproksymacyjne. Obszerny opis takich problemów oraz metod ich rozwiązywania jest przedstawiony w monografii Smutnickiego [4].

2. Sformułowanie problemu

Ogólnie, rozpatrywany w pracy problem można przedstawić następująco. Dany jest zbiór n informacji $I = \{1, 2,\dots ,n\}$, które należy przesłać do odbiorców kanałami z m elementowego zbioru $K = \{1, 2,\dots, m\}$. Z i-ta informacja ($i \in \mathrm{I}$) jest związany czas przesyłania p_i, najwcześniejszy e_i oraz najpóźniejszy d_i moment przesyłu, a także współczynniki funkcji kary za zbyt wczesne u_i lub zbyt późne w_i przesłanie. Jeżeli ustalony jest moment rozpoczęcia przesyłania i-tej informacji S_i,a $C_i = S_i + p_i$ jest momentem jej zakończenia, $E_j = max\{0,e_j — C_j\}$ przyspieszeniem, $T_j= \max\{0, C_j-d_j\}$ spóźnieniem, natomiast $f_i(C_i)=u_jE_j+w_jT_j$ kosztem (kara) wysłania informacji. Należy wiec przydzielić informacje do kanałów oraz wyznaczyć kolejności ich wysyłania, w każdym z kanałów, aby zminimalizować sumę kosztów. Musza być przy tym spełnione następujące ograniczenia: (i) wysyłanie informacji nie może byc przerwane, (ii) w dowolnym momencie przez kanał może być wysyłana co najwyżej jedna informacja, (iii) dowolna informacja może być przesłana tylko przez jeden kanał, (iv) w kanale informacje są przesyłane bezpośrednio jedna po drugiej.

Dla ustalenia uwagi zakładamy, ze przesyłanie informacji rozpoczyna sie w chwili 0. W skrócie problem ten będziemy oznaczali przez VME/T.

Ciąg zbiorów informacji

$$Q = (Q_1, Q_2,\dots, Q_m),$$

takich, że

$$Q_i \cap Q_j = \emptyset, i \neq j, i, j \in K \quad oraz \quad \sum_{i=1}^{m} Q_i = I$$

nazywamy przydziałem informacji do kanałów. Przez Z oznaczamy zbiór wszystkich takich przydziałów. Niech

$$\pi = (\pi_1, \pi_2, \ldots, \pi_m),$$

będzie ciągiem permutacji takim, ze π_i jest n_i elementową ($n_i=|Q_i|$) permutacja (kolejnością wysyłania) informacji ze zbioru Q_k, $k \in K$, tj. przez k-ty kanał. Przez $P(Q)$ oznaczamy zbiór wszystkich takich permutacji.

Zbiór rozwiązań dopuszczalnych problemu rozsyłania informacji w sieci VME/T można więc zdefiniować następująco:

$$ZP=\{(Q,\pi) : Q \in Z, \pi \in P(Q)\}. \quad (1)$$

Rozwiązanie problemu sprowadza się więc do wyznaczenia rozwiązania optymalnego $\Theta^* \in$ ZP takiego, że

$$F(\Theta^*)=min\{ F(\Theta): \Theta \in P(Q)\}$$

gdzie $F(\Theta) = \sum_{k=1}^{m} F_k(\Theta)$, $F_k(\Theta) = \sum_{j \in Q_k} f_j(\Theta)$, a $f_j(\Theta)$ jest kosztem przesłania informacji i.

3. Metoda rozwiązania

Dla ustalonego przydziału informacji do kanałów. wyznaczenie optymalnej (tj. minimalizującej sumę kosztów) kolejności ich wysyłania jest problemem NP-trudnym (dla pojedynczego kanału sprowadza sie to do NP-trudnego problemu rozpatrywanego w pracy [6]). Do rozwiązania rozpatrywanego w pracy (ogólniejszego) problemu VME/T będziemy wiec stosowali algorytm przybliżony, którego idee można przedstawić następująco.

Algorytm

> wyznaczyć rozwiązanie startowe Θ=(Q,π), $\Theta \in$ZP;
> Θ*:=Θ
> **repeat**
>> Krok 1: wygenerować z Q nowy przydział Q′$\in$Z;
>> Krok2: dla przydziału Q′ wyznaczyć kolejność π′$\in$P(Q′) wysyłania informacji przez poszczególne kanały; (Θ′=(Q',π') jest nowym rozwiązaniem);
>> **If** F(Θ′)<F(Θ*) **then** Θ*:=Θ′;
>> Θ:=Θ′
>
> **until** {Warunek zakończenia}

Krok 1 może być zrealizowany na wiele sposobów. Ogólnie, sprowadza się to do wymiany (przeniesienia) pewnych informacji pomiędzy kanałami. Należy jedynie podkreślić, że już dla dwóch kanałów liczba możliwych przydziałów wynosi 2^n. Z kolei realizacja Kroku 2 wymaga osobno, dla każdego z kanałów, ustalenia kolejności wysyłania informacji (problem NP-trudny). Dlatego w tym kroku będziemy stosowali algorytm metaheurystyczny oparty na metodzie przeszukiwania z tabu. W następnym rozdziale przedstawiamy własności umożliwiające eliminacje pewnych obszarów przestrzeni rozwiązań dzięki czemu można znacznie poprawić efektywność algorytmu.

4. Własności problemu

Realizacja Kroku 2 algorytmu będzie polegała na wyborze z najbardziej obciążonego kanału informacji o największym koszcie i przeniesieniu jej na pewna pozycje do najmniej obciążonego kanału.

Niech $\Theta = (Q,\pi)$ będzie pewnym rozwiązaniem problemu VME/T. Rozpatrujemy dwa kanały k i l ($k \neq l$, $k,l \in K$). Z rozwiązania Θ generujemy nowe rozwiązanie $\Theta' = (Q',\pi')$ przez przeniesienie jednej informacji z kanału k do l. Niech s będzie pewną pozycja w permutacji π_k, a t pozycją w permutacji π_l. Ruch typu transfer (w skrócie t - ruch) przenosi informacje znajdująca sie na pozycji s w kanale k na pozycje t w kanale l generując w ten sposób nowe rozwiązanie Θ' (w skrócie będziemy pisali, ze $\Theta' = \tau_l^k(s,t)(\Theta)$). Łatwo sprawdzić, ze tak wygenerowane rozwiązaniem $\Theta' = (Q',\pi')$ jest dopuszczalnym dla problemu VME/T. Poniżej przedstawiamy twierdzenia z których wynika, ze pewne pozycje, na które wstawiamy przenoszona informacje, można pominąć.

Bloki w permutacji
Dla ustalenia uwagi, niech β będzie permutacją zadań, tj. kolejnością wysyłania informacji przez pewien kanał. Dowolne zadanie $\beta(i)$ w permutacji $\beta \in \Phi_n$ jest przyspieszone (ang. early), jezeli $C_{\beta(i)} < e_{\beta(i)}$, terminowe (ang. on time), jeśli $e_{\beta(i)} \leq C_{\beta(i)} \leq d_{\beta(i)}$ oraz spóźnione (ang. tardy), gdy $C_{\beta(i)} \geq d_{\beta(i)}$.

Permutacje β można podzielić na bloki (subpermutacje)

$$B = [B^1, B^2, \ldots, B^v],$$

z których każda jest ciągiem zadań spełniających ograniczenia:

1.
$$B^i = (\beta(a^i), \beta(a^i+1), \ldots, \beta(b^i - 1), \beta(b^i)),$$
$$a^i = b^{i-1} + 1,\ 1 \leq i \leq v,\ b^0 = 0,\ b^v = n.$$

2. Każde zadanie $j \in B^i$ spełnia dokładnie jeden z warunków:

$$e_j > C_{\beta(b^i)} \quad \text{lub} \quad \text{(C1)}$$

$$e_j > C_{\beta(b^{i-1})} + p_j \ \wedge d_j \geq C_{\beta(b^i)} \quad \text{lub} \quad \text{(C2)}$$

$$e_j > C_{\beta(b^{i-1})} + p_j \quad \text{(C3)}$$

Jest oczywiste, że warunek C1 spełniają zadania przyspieszone, C2 zadania terminowe, a C3 zadania spóźnione. Każda z subpermutacji podziału B jest wiec odpowiednio blokiem zadań przyspieszonych (E-blokiem) lub terminowych (T-blokiem) lub też zadań (D-blokiem) spóźnionych. W pracy [6] przedstawiono algorytm rozbicia permutacji na bloki o złożoności obliczeniowej O(n).

Niech δ będzie subpermutacja permutacji β. Elementy w δ występują zgodnie z reguła, WSPT (ang. Weighted Shortest Processing Time), jeżeli są uszeregowane według nierosnących wartości w_j /p_j. Podobnie elementy w δ występują zgodnie z reguła WLPT (ang. Weighted Longest Processing Time). jeżeli są uszeregowane według niemalejących wartości u_j/p_j.

Jeżeli w D-bloku zadania występują, zgodnie z reguła WSPT, to jest to optymalne uszeregowanie zadań tego bloku. Podobnie, jeżeli B jest blokiem zadań przyśpieszonych (E-blokiem), to uszeregowanie elementów zgodnie z zasada, WLPT jest optymalne. Niech B będzie pewnym podziałem β na bloki. Permutacja β jest uporządkowana, jeżeli zadania w D- blokach występują, zgodnie z reguła WSPT, a w E-blokach - zgodnie z reguła WLPT. Można wówczas udowodnić, że żadna zmiana kolejności elementów w dowolnym bloku nie generuje rozwiązania o mniejszym koszcie.

Kryteria eliminacyjne

Rozpatrujemy ruchy typu transfer przenoszące pewna, informacje t z kanału k do kanału l. Dla uproszczenia zapisu przyjmujemy, że permutacja naturalna n = (1, 2,... , h) jest kolejnością wysyłaniu informacji w kanale k. Dla informacji v, niech p,e,d będą odpowiednio: czasem oraz najwcześniejszym i najpóźniejszym terminem przesłania. Permutacja n została rozbita na bloki. Rozpatrujemy pewien blok $B = (a, a + 1,... ,b - 1,b)$ oraz ruchy polegające na wstawieniu informacji t na poszczególne pozycje w tym bloku. Pokażemy że pewne z tych pozycji można pominąć ponieważ tak wyznaczone kolejności wysyłania maja większy koszt kosztowne od innych. Dowody twierdzę pomijamy. bowiem zasadnicze idee są podobne do dowodów twierdzeń przedstawionych w pracy [6].

Twierdzenie 1. Jeżeli B jest blokiem zadań przyspieszonych, a t wstawianym do B zadaniem oraz $C_b + p \leq e$ oraz $min\{d_i : i = 1, 2,... ,n\} + p \leq e$, to t należy wstawić na taka pozycje, aby elementy były uporządkowane zgodnie z reguła, WLPT.

Twierdzenie 2. Jeżeli B jest blokiem zadań terminowych lub spóźnionych, a t wstawianym do B zadaniem oraz $C_b + p \leq e$, to t należy wstawić na ostatnia, pozycje w bloku B.

Twierdzenia powyższe („kryteria eliminacyjne") obejmują przypadek. gdy zadanie t po wstawieniu na dowolna pozycje, w bloku B jest przyspieszone. Podobne twierdzenia można udowodnić dla przypadku. gdy odpowiednio zadanie to jest terminowe albo spóźnione.

5. Eksperymenty obliczeniowe

Przedstawiony w pracy algorytm został zaimplementowany w języku C# oraz uruchomiony na komputerze z procesorem Intel Xeon. Przykłady testowe zostały wygenerowane zgodnie z procedurą zamieszczona w pracy Beasley'a [1]. Dodatkowo, podobnie jak współczynnik kary w_j wygenerowano także u_j. Ponadto, losowo wyznaczono najwcześniejsze momenty rozpoczęcia wysyłania informacji e_j = random(0,70/100) * d_j). W Tabeli 1 zamieszczono wyniki algorytmu konstrukcyjnego (bazującego na metodzie zachłannej) oraz przedstawionego w pracy algorytmu przeszukiwania z tabu. Dla każdej pary: liczby maszyn wirtualnych oraz lokalizacji zapasowych, wygenerowano 10 losowych zestawów danych. Średnie z poszczególnych wyników umieszczono tabeli. Względna wielkość poprawy wyznaczono następująco:

$$\delta = \frac{AK - ATS}{AK}$$

gdzie *AK* jest wynikiem algorytmu konstrukcyjnego, a *ATS* - algorytmu przeszukiwania z tabu. Średnia poprawa wyników algorytmu konstrukcyjnego wynosi około 18%.

Tabela 1. Względna poprawa rozwiązania konstrukcyjnego.

Liczba maszyn	Liczba kanałów	Algorytm konstrukcyjny	Algorytm z TS	Poprawa δ
50	5	121336	88494	0.27
100	5	614495	462223	0.25
200	5	2555773	1862234	0.27
500	5	15562223	12762905	0.17
50	10	136998	119124	0.13
100	10	673991	579248	0.14
200	10	2244920	19563661	0.11
500	10	15531563	14052978	0.10

6. Podsumowanie

W pracy rozpatrujemy pewien NP-trudny problem wykonania replikacji maszyn wirtualnych, tj. przesłania ich obrazów, poprzez siec, do różnych miejsc ich składowania. Przedstawiamy własności problemu, algorytm oraz wyniki przeprowadzonych eksperymentów obliczeniowych.

7. Literatura

[1] Beasley E., Weighted tardiness, OR-Library, http://people.brunel.ac.uk/ mastjjb/jeb/info.html

[2] Hadji, M., Scalable and Cost-Efficient Algorithms for Reliable and Di- stributed Cloud Storage. International Conference on Cloud Computing and Services Science, 15-37, 2015.

[3] Mansouri, Y., Toosi, A.N., Buyya, R., Brokering algorithms for optimi- /ing the availability and cost of cloud storage services. Proceedings of the 2013 IEEE International Conference on Cloud Computing Technology and Science, 581-589, 2013.

[4] Smutnicki C., Optimization and control in just-in-time manufacturing system, Prace naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej, Seria Monografie, 1997.

[5] Thanasis, P.G., Bonvin, N., Aberer, K.: Scalia: an adaptive scheme for efficient multi-cloud storage. Proceedings of the International Confe- rence on High Performance Computing, Networking, Storage and Ana- lysis, 1-20, 2012
[6] Wodecki M., A block approach to earliness-tardiness scheduling pro- blems, International Journal on Advanced Manufacturing Technology, 40, 797-807, 2009.

Metoda lokalnych poszukiwań na ograniczonym obszarze przestrzeni rozwiazań

Wojciech BOZEJKO[1], Andrzej GNATOWSKI[2], Mariusz UCHRONSKI[3], Mieczysław WODECKI[4]

[1,2,3]*Politechnika Wrocławska*

[4]*Uniwersytet Wrocławski*

Streszczenie: W pracy przedstawiamy pewne empiryczne metody badania przestrzeni rozwiązań problemów optymalizacyjnych. Próbkując zbiór rozwiązań dopuszczalnych wyznaczamy histogram częstości występowania minimów lokalnych i na jego podstawie weryfikujemy hipotezę statystyczna dotycząca rozkładu (normalnego) występowania tych minimów. Dzięki temu możemy elastycznie zmieniać „promień" obszaru poszukiwań. Przeprowadzając eksperymenty obliczeniowe na przykładach problemu job shop otrzymaliśmy obiecujące wyniki.

Słowa kluczowe: przestrzeń rozwiązań, minima lokalne, job shop, przeszukiwanie z tabu

1. WPROWADZENIE

W tym rozdziale, bardzo skrótowo, przedstawiamy pewne definicje i oznaczenia, które będą stosowane w dalszej części pracy. Dokładnie są one opisane w pracy Bożejko i in. [2]. Zamieszczony jest tam także obszerny przegląd literatury.

Niech Ω będzie zbiorem wszystkich permutacji elementów ze zbioru $I = \{1, 2, \ldots , n\}$. Rozpatrujemy problem optymalizacyjny:

$$\min\{F(\alpha) : \alpha \in \Omega\},$$

gdzie funkcja celu $F : \Omega \rightarrow R^{+}$. Jeżeli problem ten jest NP-trudny, wówczas do jego rozwiązania stosowane są algorytmy oparte na metodzie lokalnego poszukiwania, których zasadniczym elementem jest otoczenie tj. odwzorowanie:

$$N : \Omega \rightarrow 2^{\Omega},$$

przyporządkowujące każdemu elementowi $a \in \Omega$ podzbiór $N(\alpha) \subseteq \Omega$.

Grafy oraz generatory otoczeń

Poszczególne elementy otoczenia $N(\alpha)$ są generowane przez ruchy. tj. „niewielkie" zmiany kolejności elementów (zamiany lub przestawienia). Zbiór takich ruchów nazywamy generatorem otoczenia $N(\alpha)$. W przypadku. gdy rozwiązaniami dopuszczalnymi są permutacje. generator otoczenia można utożsamić z pewnymi podzbiorami zbioru wszystkich permutacji Ω. Będziemy go oznaczali przez $G(\alpha)$.

Dla otoczenia N, definiujemy graf otoczenia $G(N) = (V, A)$, gdzie zbiór wierzchołków $V = \Omega$ .a zbiór łuków

$$A=\{(\alpha,\beta): \beta \in N(\alpha), \alpha \in \Omega\}$$

Na zbiorze permutacji (wierzchołkach grafu G). w literaturze. zdefiniowano wiele miar określających odległości pomiędzy wierzchołkami. Dla ustalenia uwagi. niech $d(u, v)$ będzie taka miara. Załóżmy. że

$$LM^* = \{\pi_1^*, \pi_2^*, ..., \pi_s^*\} \tag{1}$$

jest zbiorem minimów lokalnych pewnego problemu. Definiujemy graf $G^* = (V^*, A^*)$, gdzie

$$A^* = LM^*, \text{ a } V^* = \{(u, v) : u,v \in A^*\}$$

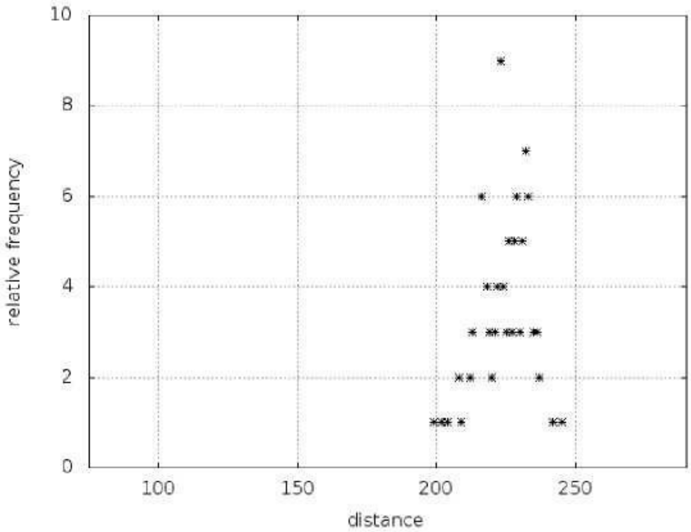
Miarę d^* odległości wierzchołków w G^* definiujemy następująco:

$$d^*(u,v) = d(u,v),\ u,v \in V^*.$$

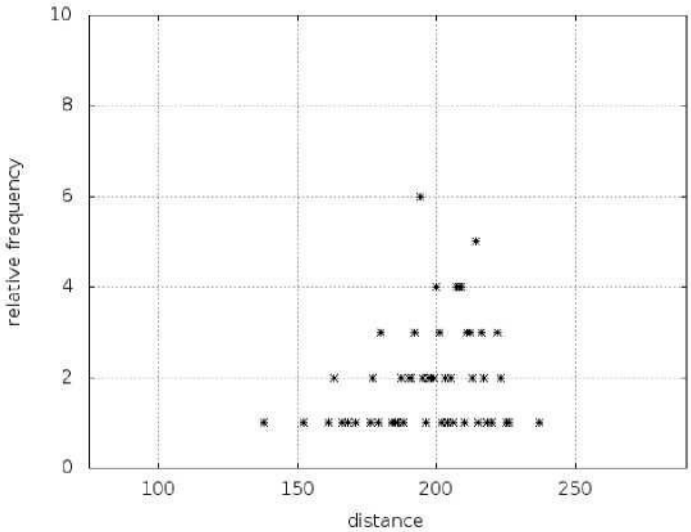
Empiryczny rozkład minimów lokalnych

Wyznaczamy wierzchołek centralny π^*_{cen} w grafie G^*, oraz odległości $d^*(\pi^*_{cen}, v)$, $v \in A^*$. Nastepnie. średnia i oraz odchylenie standardowe a z tych odległości.

Dla pewnych przykładów Taillarda [3] zbadano rozkłady położenia 100 minimów lokalnych. Za odległość w grafie Z^* przyjęto miarę Calye'a (minimalna liczba transpozycji). Na rysunkach 1 i 2 przedstawiono. dla przykładu Tail11, wykresy częstości występowania minimów względem centrum grafu G^* oraz najlepszego znanego obecnie rozwiązania. Na osi X oznaczono doległoś. a na Y - liczbę minimów.



Rysunek 1: Rozkład minimów lokalnych względem centrum grafu, dla przykładu Tail11.



Rysunek 2: Rozkład minimów lokalnych względem najlepszego znanego obecnie rozwiązania, dla przykładu Tail11.

W pierwszym przypadki minima są skupione w odległości 250, a w drugim 200 odpowiednio od wierzchołka centralnego oraz najlepszego rozwiązania. Ponadto, jak łatwo zauważyć, oba wykresy można dokładnie aproksymować przez zmienna losowa, o rozkładzie normalnym.

W ogólnym przypadku weryfikujemy hipotezę H_0 : elementy zbioru minimów lokalnych LM_*, ze względu na miarę d^* mają, rozkład normalny $N(\mu, \delta)$. W przypadku pozytywnej weryfikacji. przeszukiwanie przestrzeni rozwiązań ograniczamy do rozwiązań, które są odległe od wierzchołka centralnego π^*_{cen} , nie więcej niz $\delta \bullet \alpha$. gdzie δ jest parametrem.

1 Przeszukiwanie ograniczenego obszaru przestrzeni

Na bazie analizy opisanych w poprzednim rozdziale eksperymentów obliczeniowych przyjęliśmy następujące założenia:

- minima lokalne są skoncentrowane w pewnym obszarze przestrzeni rozwiązań,
- zależność pomiędzy odległością od rozwiązania optymalnego, a liczba minimów ma rozkład normalny.

Dzięki temu, korzystając ze wstępnych obliczeń, wyznaczamy obszar poszukiwali w zbiorze rozwiązań. W zależności od potrzeb obszar ten możemy zmniejszać lub powiększać. Idę algorytmu wyznaczania elementu suboptymalnego w zbiorze Ω można przedstawić następująco:

ALGORYTM INTS

Dane:

zbiór minimów lokalnych LM^*;

Krok 1:{ *Empiryczny rozkład minimów lokalnych*}

Wyznaczyć wierzchołek centralny π^*_{cen} w grafie G^*;
obliczyć średnią μ oraz odchylenie standardowe σ rozkładu
N(μ, δ) - zależności pomiędzy odległością od π^*_{cen},
a liczbą, minimów lokalnych;

Krok 2: {*Przeszukiwanie*}

$\vartheta = \delta \bullet \alpha$, gdzie $1 < \delta < 5$;
$\pi := \pi^*$
wyznaczyć $\pi^* \in LM^*$ takie, że
$C_{max}(\pi^*) = \min\{ C_{max}(\pi^*_i): (\pi^*_i \in P\}$:
repeat
    wygenerować $N(\pi)$ - otoczenie π zawierające elementy
    odległe od π nie więcej niż o ϑ;
    wyznaczyć element $\pi \in N(\pi)$ taki, że
    $C_{max}(\pi') = \min\{ C_{max}(\pi): (\pi \in N(\pi) \}$;
if $C_{max}(\pi') < C_{max}(\pi^*)$ then $\pi^* := \pi'$;
    $\pi := \pi'$;
if {zmiana średnicy ϑ } **then** {zmień parametr δ};
until {Warunek zakończenia}

W Kroku 2 algorytmu warunkiem zakończenia może być: maksymalna liczba iteracji, czas obliczeń, wyznaczenie rozwiązania o pewnej wartości, itp. Z kolei zmiana średnicy może nastąpić, jeżeli po wykonaniu pewnej liczby iteracji nie uzyskano poprawy najlepszego bieżącego rozwiązania π^*. W kroku tym, w procesie przeszukiwania, można stosować np. dowolna metaheurystyke.

3. EKSPERYMENTY OBLICZENIOWE

Algorytm INTS, który powstał na bazie zamieszczonego w pracy [1] algorytmu NTS został zaimplementowany w języku C++ i uruchomiony na klastrze BEM we Wrocławskim Centrum superkomputerowym. Obliczenia, w ramach grantu nr 96, wykonano na przykładach Taillarda dla problemu job shop. Otrzymane zbiorcze wyniki zamieszczono w Tabeli 1. W kolumnie pierwszej jest nazwa przykładu, a drugiej - rozmiar, gdzie *n* jest liczba zadań, a *m* liczba maszyn. W kolejnych dwóch kolumnach przedstawiono średnie procentowe odchylenie w stosunku do najlepszego znanego rozwiązania, dla różnych wartości parametru δ. W kolumnie *best* zamieszczono najlepsze wyniki uzyskane dla różnych wartości parametru δ. Z kolei ostatnia kolumna zawiera wyniki algorytmu INTS z mechanizmem adaptacyjnym, który polega na modyfikacji zbioru minimów lokalnych LM^*, gdy zostaje wyznaczone lepsze minimum lokalne. W tym przypadku otrzymano najlepsze rozwiązania. Średnio są one gorsze (względnie) od najlepszych obecnie znanych o około 0.45%. Należy podkreślić, że najlepsze rozwiązania wyznaczono stosując różne algorytmy i różne czasy obliczem.

Tabela 1: Średni błąd względny, w stosunku do najlepszych rozwiązań.

			1NTS			
problem	*n x m*	NTS	Ξ	4 = 5	*best*	*adapt*
TA01-10	15 *x* 15	0.5344	0.3615	0.5262	0.2565	0.2565
TA11-20	20 *x* 15	1.0371	0.9945	1.2134	0.7937	0.7937
TA21-30	20 *x 20*	1.0101	0.8371	0.7468	0.5275	0.5156
TA31-40	30 *x* 15	0.9911	0.9006	0.8866	0.6629	0.6518
TA41-50	30 *x 20*	1.7772	1.8130	1.6803	1.2241	1.2241
TA51-60	50 *x* 15	0.0915	0.0915	0.0915	0.0915	0.0915
TA61-70	50 *x* 20	0.1479	0.1715	0.1479	0.0244	0.0244
TA71-80	100 *x 20*	0.0089	0.0089	0.0089	0.0089	0.0089
sśrednio		0.6998	0.6473	0.6627	0.4487	0.4458

4. PODSUMOWANIE

W pracy przedstawiono metodę ograniczania obszaru przeszukiwania przestrzeni rozwiązań dla problemów optymalizacyjnych, których rozwiązania dopuszczalne mogą być reprezentowane przez permutacje. Bazuje ona na empirycznym badaniu rozkładu minimów lokalnych. Uzyskane rozwiązania, dla problemu job shop, są niewiele gorsze od najlepszych obecnie znanych. Stanowią one inspiracja do dalszych badan nad przestrzeniami rozwiązań pojedynczych przykładów.

5. LITERATURA

[1] Bozejko W., Uchroilski M., Wodecki M., A neuro-tabu search algorithm for the job shop problem, Proceedings of ICAISC 2010, Lecture Notes in Artificial Intelligence No. 6114, Springer (2010), 387-394.

[2] Bozejko W., Gnatowski A., Smutnicki C. Wodecki M., Wykorzystanie własności krajobrazu przestrzeni rozwiaza! w konstrukcji algorytmów metaheurystycznych, Materiały X Sympozjum Zastosowania teorii Systemów, Zakopane 2017.

[3] Taillard E., Benchmarks for basic scheduling problems, European Journal of Operational Research

Wielomianowy algorytm przydziału dla wielomaszynowego gniazda produkcyjnego

Wojciech BOŻEJKO[1], Andrzej GNATOWSKI[1], Mieczysław WODECKI[2]

[1]*Politechnika Wrocławska, Katedra Automatyki, Mechatroniki i Systemów Sterowania ul. Janiszewskiego 11-17 Wrocław, Polska, {wojciech.bozejko, andrzej.gnatowski}@pwr.edu.pl*

[2]*Uniwersytet Wrocławski, Instytut Informatyki Joliot-Curie 15, Wrocław, Polska, mwd@ii.uni.wroc.pl*

Streszczenie: W pracy rozpatrywany jest system wytwarzania z przezbrojeniami. Składa się on z wielomaszynowego gniazda, na które w ustalonej kolejności trafiają operacje. Problem sprowadza się do wyznaczenia takiego przydziału operacji do maszyn o różnych parametrach, który minimalizuje ich łączny czas wykonywania; przy założeniu, że jednocześnie pracować może co najwyżej jedna maszyna. Liczba wszystkich możliwych przydziałów jest wykładnicza. Prezentujemy wielomianowy, względem liczby zadań, algorytm wyznaczania optymalnego przydziału operacji do maszyn.

Słowa kluczowe: algorytm dokładny, optymalizacja dyskretna, gniazdo produkcyjne

1. WPROWADZENIE

Zagadnienia szeregowania zadań w elastycznych systemach produkcyjnych stanowią szczególnie interesujący przedmiot badań. Optymalizacji podlega w nich nie tylko przydział operacji do maszyn, ale również kolejność ich wykonywania. Ze względu na NP-trudność większości wariantów problemu, podstawowym narzędziem ich rozwiązywania są algorytmy heurystyczne. Jak pokazano na przykładzie elastycznego problemu gniazdowego (Chaudhry & Khan, 2016), w literaturze dominują rozwiązania hybrydowe opierające się na łączeniu różnych algorytmów w celu osiągnięcia jak najlepszych rezultatów.

W niniejszej pracy rozważa się problem przydziału operacji (zadań) do maszyn w wielomaszynowym gnieździe produkcyjnym, w którym kolejność wykonywania operacji jest ustalona. Problem ten jest nazywany dalej problemem przydziału (AP). Gniazdo produkcyjne z AP, może być elementem większego systemu produkcyjnego, jak na przykład elastycznych systemów gniazdowych opisanych w (Bożejko, Hejducki, Uchroński, & Wodecki, 2012; Bożejko, Pempera, & Wodecki, 2015, 2016). Tym samym zaproponowany algorytm stałby się częścią rozwiązania hybrydowego, w którym optymalizacja odbywałaby się na dwóch poziomach:

1. kolejność wykonywania operacji;
2. przydział operacji do maszyn.

Dwupoziomowe metaheurystyki były z powodzeniem stosowane do rozwiązywania praktycznych problemów szeregowania (Aringhieri, Landa, Soriano, Tànfani, & Testi, 2015). Alternatywnie, jako właściwy problem optymalizacji można przyjąć sam poziom pierwszy, a przestrzeń rozwiązań poziomu drugiego rozważać jako tzw. wielkiego otoczenia (Ahuja, Ergun, Orlin, & Punnen, 2002). Problem podobny do rozważanego w pracy AP, ale z o liczbą maszyn graniczoną do dwóch, rozwiązywano w (Bozejko, Gnatowski, Klempous, Affenzeller, & Beham, 2017), cykliczny problem przydziału w dwumaszynowym gnieździe omawiano w (Bożejko, Gnatowski, Idzikowski, & Wodecki, 2016), z kolei wariant dla problemu przepływowego z czasami rozładunku przedstawiono w (Cheng, Wang, & Sriskandarajah, 1999).

2. SFORMUŁOWANIE PROBLEMU

Niech $\mathcal{J} = \{1,2,3,\dots,n\}$ oznacza zbiór operacji (zadań) wykonywanych kolejno (tj. w kolejności $1,2,\dots,n$) w gnieździe. Gniazdo produkcyjne składa się z m maszyn, numerowanych kolejno $1,2,\dots,m$, tworzących zbiór $\mathcal{M} = \{1,2,\dots,m\}$. Każda operacja $i \in \mathcal{J}$ musi być wykonywana nieprzerwanie na dokładnie jednej maszynie, określonej przydziałem $P(i) \in \mathcal{M}$, przez $p_i^{P(i)}$ czasu. Przydział wszystkich operacji do maszyn opisywany jest przez krotkę

$$P = (P(1), P(2), \dots, P(i), \dots, P(n)), \tag{1}$$

przy czym $P(i)$ oznacza przydział operacji i, a $\mathcal{P}$ to zbiór wszystkich możliwych przydziałów. Pomiędzy każdymi dwiema operacjami $i, j \in \mathcal{M}, i < j$ wykonywanymi na tej samej maszynie $a = P(i) = P(j)$, następuje przezbrojenie zajmujące $s_{i,j}^a$ czasu. Niech $s(P,i)$ oznacza dla ustalonego przydziału $P \in \mathcal{P}$ czas przezbrojenia wykonywanego przed operacją $i \in \mathcal{J}$

$$s(P,i) = \begin{cases} s_{\max\{j\in\mathcal{J}:\ j<i \wedge P(i)=P(j)\},i}^{P(i)} & \text{gdy: } \exists j \in \mathcal{J}\ \big(j < i \wedge P(i) = P(j)\big); \\ 0 & \text{w przeciwnym wypadku.} \end{cases} \tag{2}$$

Problem sprowadza się do znalezienia takiego przydziału P^* spośród m^n możliwych przydziałów ze zbioru $\mathcal{P}$, że łączny czas wykonywania wszystkich operacji

$$C_{max}(P) = \sum_{i \in \mathcal{J}} \left(p_i^{P(i)} + s(P,i) \right), \tag{3}$$

jest minimalny

$$P^* \in \underset{P \in \mathcal{P}}{\operatorname{argmin}} C_{max}(P). \tag{4}$$

3. ALGORYTM WIELOMIANOWY

Zaproponowany algorytm rozwiązujący w czasie wielomianowym problem przydziału (AP) wykorzystuje schemat programowania dynamicznego. Opiera się on na podziale rozwiązanego problemu na podproblemy, cechujące się własnością optymalnej podstruktury; to jest ich optymalne rozwiązanie jest funkcją optymalnych rozwiązań ich podproblemów ((Cormen, Leiserson, Rivest, & Stein, 2007), s. 380–381). Każdy z podproblemów opisany jest przez m parametrów, tworzących wektor

$$\vec{x} = (x_1, x_2, \dots, x_m); \tag{5}$$

spełniający warunki

$$\forall i \in \mathcal{M} \ (x_i \in \mathcal{J} \cup \{0\}); \tag{6}$$

$$\forall i,j \in \mathcal{M} \ (x_i = x_j \Rightarrow (x_i = 0 \vee i = j)). \tag{7}$$

Funkcja $w(\vec{x})$ wektorowi $\vec{x}$ przyporządkowuje najkrótszy możliwy czas wykonywania operacji dla podproblemu składającego się z operacji $1,2,\dots,\max_{i \in \mathcal{M}} x_i$, w którym: (1) jeśli na pewnej maszynie $i \in \mathcal{M}$ nie jest wykonywana żadna operacja, to $x_i = 0$; (2) z kolei jeśli na maszynie $i \in \mathcal{M}$ jest wykonywana co najmniej jedna operacja, to jako ostatnia wykonywana jest operacja x_i. Niech $\mathcal{P}^i$ oznacza zbiór wszystkich możliwych przydziałów problemu składającego się operacji $1,2,\dots,i$. Wtedy, $w(\vec{x})$ można zdefiniować formalnie jako

$$w(\vec{x}) = \min \ \left\{ C_{max}(P) : P \in \mathcal{P}^{x_{\omega(\vec{x})}} \wedge \forall a \in \mathcal{M} \left(\left(x_a = 0 \wedge \forall i \in \left\{1,2,\dots,x_{\omega(\vec{x})}\right\} (P(i) \neq a) \right) \oplus \left(x_a = \max \ \left\{ j \in \left\{1,2,\dots,x_{\omega(\vec{x})}\right\} : P(j) = a \right\} \right) \right) \right\}, \tag{8}$$

gdzie $\oplus$ to symbol alternatywy rozłącznej, a ω to funkcja zwracająca dla zadanego wektora $\vec{x}$ maszynę na której wykonywana jest ostatnia operacja

$$\omega(\vec{x}) = \max \ \{a \in \mathcal{M} : \forall b \in \mathcal{M} \ (x_a \geq x_b)\}. \tag{9}$$

Z ograniczeń problemu wiadomo, że jeżeli wykonywana jest choć jedna operacja, to istnieje dokładnie jedna taka maszyna, czyli

$$|a \in \mathcal{M} : \forall b \in \mathcal{M} \ (x_a \geq x_b)| = 1. \tag{10}$$

Analogicznie, zdefiniowano funkcję

$$\psi(\vec{x}) = \max \ \{a \in \mathcal{M}: (\forall b \in \mathcal{M}\ (x_a = x_b)) \vee (\forall b \in \mathcal{M}\ (x_a < x_b \Leftrightarrow x_b = \max_{i \in \mathcal{M}} x_i))\} \quad (11)$$

zwracająca dla zadanego wektora $\vec{x}$ maszynę na której wykonywana jest przedostatnia operacja (lub 0, jeśli wszystkie operacje są wykonywane na jednej maszynie; pominięto trywialny przypadek $m = 0$). Na przykład dla podproblemu $\vec{x} = (0,0,7,3,0,5,1)$, mamy $\omega(\vec{x}) = 3$ i $\psi(\vec{x}) = 6$.

Zbiór możliwych wektorów $\vec{x}$ (a tym samym podproblemów), można podzielić na trzy rozłączne podzbiory, spełniające warunki:

1. $\vec{x}$ takie, że: $x_{\omega(\vec{x})} = 0$;
2. $\vec{x}$ takie, że: $x_{\omega(\vec{x})} - x_{\psi(\vec{x})} > 1$;
3. $\vec{x}$ takie, że: $x_{\omega(\vec{x})} - x_{\psi(\vec{x})} = 1$.

Niech $\vec{x}$ spełnia **pierwszy warunek**. Wtedy $\forall i \in \mathcal{M}\ (x_i = 0)$ — czyli nie jest wykonywana żadna operacja, a zatem $w(\vec{x}) = 0$. Jeżeli $\vec{x}$ spełnia **drugi warunek**, to operacje od $x_{\psi(\vec{x})} + 1$ do $x_{\omega(\vec{x})}$ są wykonywane na maszynie $\omega(\vec{x})$. Czas wykonywania tych operacji, wraz z przezbrojeniami, wynosi

$$\sum_{i=x_{\psi(\vec{x})}+2}^{x_{\omega(\vec{x})}} \left(p_i^{\omega(\vec{x})} + s_{i-1,i}^{\omega(\vec{x})}\right) + p_{x_{\psi(\vec{x})}+1}^{\omega(\vec{x})}. \quad (12)$$

Rozważmy podproblem opisany wektorem

$$\vec{y} = (x_1, x_2, \dots, x_{\omega(\vec{x})-2}, x_{\omega(\vec{x})-1}, x_{\psi(\vec{x})} + 1, x_{\omega(\vec{x})+1}, x_{\omega(\vec{x})+2}, \dots, x_m).$$

Wtedy

$$w(\vec{x}) = w(\vec{y}) + \sum_{i=x_{\psi(\vec{x})}+2}^{x_{\omega(\vec{x})}} \left(p_i^{\omega(\vec{x})} + s_{i-1,i}^{\omega(\vec{x})}\right). \quad (13)$$

Dla **warunku trzeciego**, tj. $x_{\omega(\vec{x})} - x_{\psi(\vec{x})} = 1$, podproblemy również definiowane są przez modyfikację (zmniejszenie) wartości $x_{\omega(\vec{x})}$. Jako ostatnia na maszynie $\omega(\vec{x})$ może być wykonana operacja ze zbioru $\{0,1,2, \dots, x_{\omega(\vec{x})-1}\}$, za wyjątkiem $\{x_1, x_2, \dots, x_m\}$, czyli ostatnich operacjach wykonywanych na pozostałych maszynach określanych przez $\vec{x}$. Ostatecznie, $x_{\omega(\vec{x})} \in u(\vec{x})$, gdzie

$$u(\vec{x}) = \{0\} \cup (\{1,2, \dots, x_{\omega(\vec{x})-1}\} \backslash \{x_1, x_2, \dots, x_m\}). \quad (14)$$

Dla uproszczenia wzorów, niech

$$\mu(\vec{x}, y) = (x_1, x_2, \dots, x_{\omega(\vec{x})-1}, y, x_{\omega(\vec{x})+1}, \dots, x_m). \quad (15)$$

Wtedy

$$w(\vec{x}) = \min \ \left\{w(\mu(\vec{x}, i)) + p_{x_{\omega(\vec{x})}}^{\omega(\vec{x})} + s_{i,x_{\omega(\vec{x})}}^{\omega(\vec{x})} : i \in u(\vec{x})\right\}, \quad (16)$$

zakładając odtąd, że

$$\forall a \in \mathcal{M}\ \forall i \in \{1,2, \dots\}\ (s_{0,i}^a). \quad (17)$$

Podsumowując, najkrótszy czas wykonania operacji dla podproblemu $\vec{x}$, wynosi

$$w(\vec{x}) = \begin{cases} w\left(\mu(\vec{x}, x_{\psi(\vec{x})} + 1)\right) + \sum_{i=x_{\psi(\vec{x})}+2}^{x_{\omega(\vec{x})}} \left(p_i^{\omega(\vec{x})} + s_{i-1,i}^{\omega(\vec{x})}\right) & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} > 1; \\ \min\left\{w(\mu(\vec{x}, i)) + p_{x_{\omega(\vec{x})}}^{\omega(\vec{x})} + s_{i,x_{\omega(\vec{x})}}^{\omega(\vec{x})} : i \in u(\vec{x})\right\} & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} = 1; \\ 0 & \text{gdy: } x_{\omega(\vec{x})} = 0, \end{cases} \tag{18}$$

pamiętając o założeniu z równania (17).

Wzór (18) może być wykorzystany do wyznaczenia optymalnego rozwiązania AP. W tym celu należy wprowadzić m dodatkowych, wykonywanych w ostatniej kolejności operacji, które spełniają następujące warunki:

$$\forall a \in \mathcal{M}\ \forall i \in \{n+1, n+2, \dots, n+m\}\ (p_i^a = 0) \tag{19}$$

$$\forall a \in \mathcal{M}\ \forall i \in \{n+1, n+2, \dots, n+m\}\ \forall j \in \{0,1, \dots, i-1\}\ \left(s_{j,i}^a = 0\right) \tag{20}$$

to jest, mają zerowe czasy wykonywania oraz przezbrojeń. Można łatwo pokazać, że

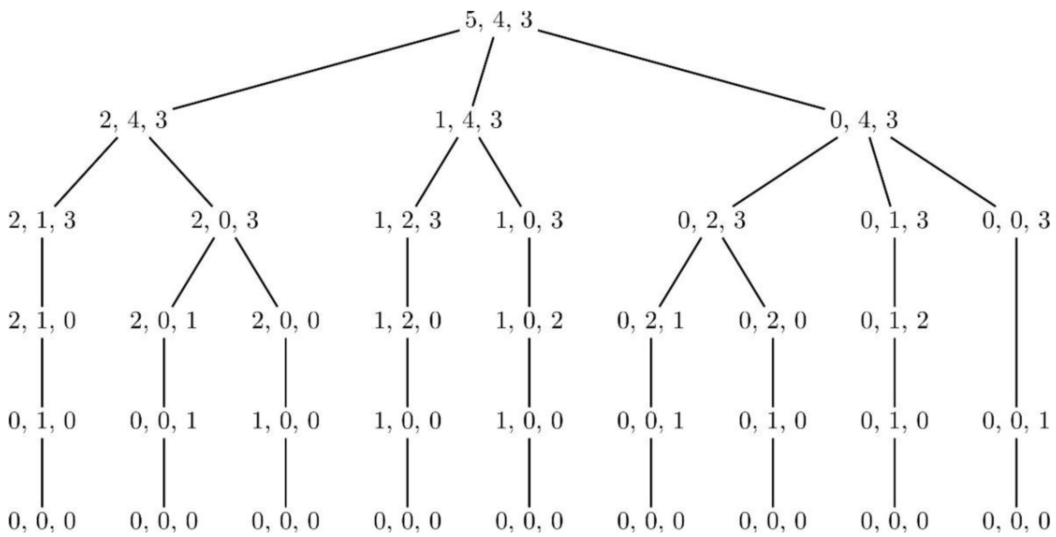
$$w\big((n+m, n+m-1, \dots, n+1)\big) = C_{max}(P^*), \tag{21}$$

(warunki z równań (6–7), nałożone na wektory opisujące podproblemy są spełnione, ponieważ wraz z dodaniem operacji, rozmiar nowego problemu to $n' = n + m$). Z kolei w celu obliczenia optymalnego przydziału P^*, należy prześledzić proces wyznaczania $w\big((n+m, n+m-1, \dots, n+1)\big)$. Rozważmy instancję pokazaną w tab. 1, składającą się z dwóch operacji oraz trzech maszyn.

Tabela 1: Przykładowa instancja problemu, $m = 3$, $n = 2$.

	p_i^a			$s_{i,j}^a$					
a	1	2	3	1		2		3	
j \ i	-	-	-	1	2	1	2	1	2
1	1	9	10	3	4	5	4	4	5
2	8	10	9	4	0	4	2	4	0

Wzór (21) przyjmuje dla niego postać $C_{max}(P^*) = w((5,4,3))$. Aby wyznaczyć wartość $w((5,4,3))$, niezbędna jest znajomość wartości $w((2,4,3))$, $w((1,4,3))$ oraz $w((0,4,3))$, z kolei dla $w((2,4,3))$, $w((2,1,3))$ oraz $w((2,0,3))$, itd.; rekurencyjna zależność kończy się na $w((0,0,0)) = 0$. Podsumowanie zależności pomiędzy podproblemami pokazano na rys. 1, gdzie w wierzchołkach zaznaczono kolejne wektory opisujące podproblemy.



Rysunek 1: Drzewo zależności podproblemów instancji z tabeli 1.

Jak widać, niektóre wierzchołki (a tym samym wartości $w(\vec{x})$) są niezbędne wielokrotnie, jak na przykład $(0,0,0)$, który jest wykorzystany 9 razy. Bazując na tym spostrzeżeniu, zdefiniowano skierowany graf $\mathcal{B} = (V, E)$, gdzie zbiór wierzchołków V i zbiór łuków E są zdefiniowane w następujący sposób:

- V to zbiór wektorów opisujących wszystkie podproblemy, rozważane rekurencyjnie przy wyznaczaniu za pomocą wzoru (18) wartości $w\big((n+m, n+m-1, \dots, n+1)\big)$:

$$V = c\big((n+m, n+m-1, \dots, n+1)\big), \qquad (22)$$

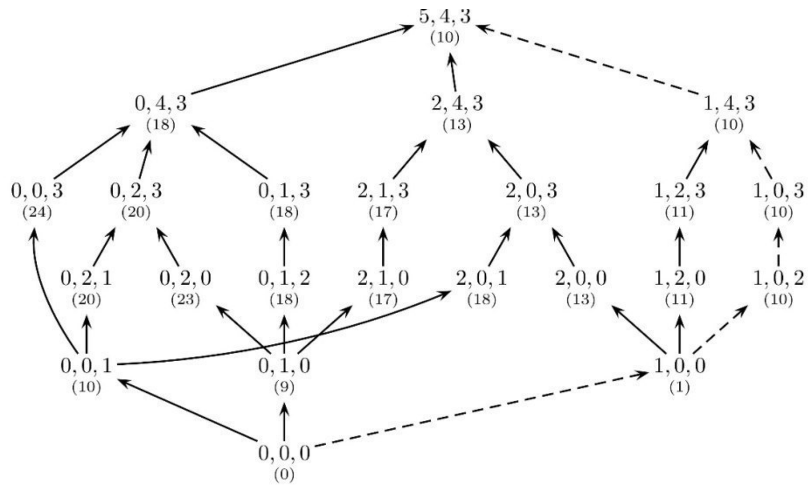
$$c(\vec{x}) = \begin{cases} \left\{\mu(\vec{x}, x_{\psi(\vec{x})}+1)\right\} \cup c\left(\mu\left(\vec{x}, x_{\psi(\vec{x})}+1\right)\right) & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} > 1; \\ \bigcup_{i \in u(\vec{x})} \left(\{\mu(\vec{x}, i)\} \cup c\big(\mu(\vec{x}, i)\big)\right) & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} = 1; \\ \{\emptyset\} & \text{w przeciwnym wypadku.} \end{cases} \qquad (23)$$

- E to zbiór wszystkich takich skierowanych łuków $(\vec{\imath}, \vec{\jmath})$, że wartość $w(\vec{\imath})$ jest *bezpośrednio* potrzebna do obliczenia $w(\vec{\jmath})$ za pomocą wzoru (18):

$$E = \{(\vec{\imath}, \vec{\jmath}) : \vec{\imath}, \vec{\jmath} \in V \wedge \vec{\imath} \in d(\vec{\jmath})\}, \qquad (24)$$

$$d(\vec{x}) = \begin{cases} \left\{\mu(\vec{x}, x_{\psi(\vec{x})}+1)\right\} & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} > 1; \\ \bigcup_{i \in u(\vec{x})} \{\mu(\vec{x}, i)\} & \text{gdy: } x_{\omega(\vec{x})} - x_{\psi(\vec{x})} = 1; \\ \{\emptyset\} & \text{w przeciwnym wypadku.} \end{cases} \qquad (25)$$

Każdy wierzchołek $\vec{a} \in V$ obciążony jest wagą $w(\vec{a})$. Na rys. 2 pokazano graf $\mathcal{B}(\vec{x})$ dla przykładu z tabeli 1 oraz $\vec{x} = (5,4,3)$.



Rysunek 2: Graf $\mathcal{B}$ dla przykładu z tabeli 1. Wagi wierzchołków zapisano w nawiasach.

W grafie $\mathcal{B}$, wyróżnić można ścieżkę κ, zdefiniowaną jako

$$\kappa = \left(\kappa_1, \kappa_2, \dots, \kappa_{|\kappa|}\right), \tag{26}$$

$$\kappa_i = \begin{cases} (n+m, n+m-1, \dots, n+1) & \text{gdy: } i = 1; \\ \max \quad \arg\min_{\vec{x} \in d(\kappa_{i-1})} \{w(\vec{x})\} & \text{gdy: } i \in \{2,3,\dots,|\kappa|\}. \end{cases} \tag{27}$$

Na rys. 2. zaznaczono przerywaną linią ścieżkę

$\kappa = \big((5,4,3),(1,4,3),(1,0,3),(1,0,2),(1,0,0),(0,0,0)\big)$ dla instancji z tabeli 1. Korzystając z algorytmu 1., na podstawie analizy ścieżki κ, wyznaczyć można optymalny przydział P^*.

Algorytm 1: Sposób wyznaczania optymalnego przydziału.

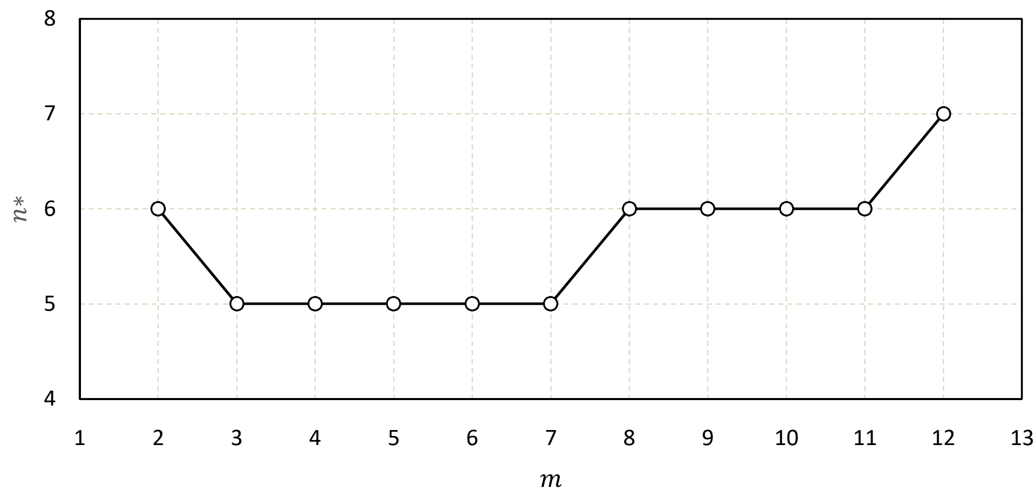
1: Wyznacz wierzchołki i wagi grafu $\mathcal{B}$
2: W kolejności topologicznej, wyznacz wszystkie wagi wierzchołków grafu $\mathcal{B}$
3: Wyznacz ścieżkę κ
4: **for** $i = 1,2,\dots,|\kappa|$ **do**
5:     **for** $j = 1,2,\dots,m$ **do**
6:         **if** $\kappa_{i_j} \neq 0$ **then**
7:             $P(\kappa_{i_j}) \leftarrow j$
8:         **end if**
9:     **end for**
10: **end for**

Na przykład, dla instancji z tab. 1, z $\kappa_1 = (5,4,3)$ wynika, że $P(5) = 1$, $P(4) = 2$, $P(3) = 3$, a z $\kappa_4 = (1,0,2)$, że $P(1) = 1$ oraz $P(2) = 3$. Ostatecznie $P = (1,3,3,2,1)$, przy czym tylko $P(1)$ i $P(2)$ są znaczące, ponieważ operacje 3,4,5 mają zerowe czasy wykonywania i przezbrojeń. Tym samym, dla omawianego przykładu, $P^* = (1,3)$.

Twierdzenie 1. Przydział n operacji do m maszyn w problemie przydziału (AP), minimalizujący łączny czas wykonywania operacji, można wyznaczyć w czasie $O((n+m)^{m+1})$.

Dowód. Dowód polega na wyznaczeniu złożoności obliczeniowej algorytmu 3. Krok pierwszy polega na zbudowaniu grafu $\mathcal{B}$, który odzwierciedla kolejność obliczeń wzoru (18). Z definicji, liczba wierzchołków grafu nie może być większa od liczby możliwych podproblemów spełniających warunki ze wzorów (6)–(7), to jest $(n+m)^m$. Stąd złożoność wykonania kroku 1. to $O((n+m)^m)$. Ze wzoru (18), w kroku 2., dla każdego z wierzchołków wykonywane są obliczenia o złożoności $O(n+m)$, stąd łączny czas wykonywania kroku 2. to $O((n+m)(n+m)^m) = O((n+m)^{m+1})$. Ponieważ $|\kappa| \leq n+m$, złożoność kroków 3.–10. wynosi $O((m+n)m)$. Tym samym, złożoność całego algorytmu to $O((n+m)^m) + O((n+m)^{m+1}) + O((n+m)m) = O((n+m)^{m+1})$.

Ze względu na złożoną strukturę grafu $\mathcal{B}$, oszacowanie z tw. 1 może być w praktyce zbyt luźne i nie dostarczać wystarczającej informacji na temat efektywności algorytmu. W celu dalszej ewaluacji algorytmu, porównano wyznaczoną numerycznie liczbę węzłów grafu $\mathcal{B}$, z liczbą rozwiązań rozważanych przez przegląd zupełny (m^n) dla $m < 13$ (z uwagi na tempo wzrostu $(n+m)^{m+1}$, rozważanie liczby maszyn większej niż 12 ma w obliczu mocy obliczeniowych współczenych komputerów ograniczoną celowość). Wyniki podsumowano na wykresie na Rys. 3.



Rysunek 3: Porównanie algorytmu wielomianowego do przeglądu zupełnego.

Przez n^* oznaczono najmiejszą liczbę operacji problemu m maszynowego, dla którego liczba węzłów grafu $\mathcal{B}$ jest mniejsza od liczby rozwiązań rozważanych przez przeglądu zupełnego (m^n). Z analizy uzyskanych danych wynika, że dla interesujących w praktyce rozmiarów problemu, przestawiony algorytm wielomianowy przegląda mniejszą liczbę węzłów (lub rozwiązań) niż przegląd zupełny.

4. PODSUMOWANIE

W pracy rozpatrywano problem przydziału dla wielomaszynowego gniazda produkcyjnego. Opierając się na schemacie programowania dynamicznego, zaproponowano wielomianowy względem liczby operacji algorytm dokładny o złożoności obliczeniowej $O((n+m)^{m+1})$, pomimo przestrzeni rozwiązań o wykładniczym rozmiarze m^n. Algorytm może być wykorzystany jako element wielopoziomowych heurystyk rozwiązywania niektórych elastycznych problemów szeregowania zadań.

5. LITERATURA

[1] Ahuja, R. K., Ergun, Ö., Orlin, J. B., & Punnen, A. P. (2002). A survey of very large-scale neighborhood search techniques. *Discrete Applied Mathematics*, *123*(1–3), 75–102. https://doi.org/10.1016/S0166-218X(01)00338-9

[2] Aringhieri, R., Landa, P., Soriano, P., Tànfani, E., & Testi, A. (2015). A two level metaheuristic for the operating room scheduling and assignment problem. *Computers and Operations Research*, *54*, 21–34. https://doi.org/10.1016/j.cor.2014.08.014

[3] Bożejko, W., Gnatowski, A., Klempous, R., Affenzeller, M., & Beham, A. (2017). Cyclic scheduling of a robotic cell. W *7th IEEE International Conference on Cognitive Infocommunications, CogInfoCom 2016 - Proceedings*. https://doi.org/10.1109/CogInfoCom.2016.7804579

[4] Bożejko, W., Gnatowski, A., Idzikowski, R., & Wodecki, M. (2016). Algorytm wielomianowy dla cyklicznego problemu przydziału operacji w dwumaszynowym gnieździe. W *Automatyzacja procesów dyskretnych, Teoria i zastosowania (red. A. Świerniak, J. Krystek), Tom I* (ss. 41–48). Gliwice.

[5] Bożejko, W., Hejducki, Z., Uchroński, M., & Wodecki, M. (2012). Solving the Flexible Job Shop Problem on Multi-GPU. *Procedia Computer Science*, *9*, 2020–2023. https://doi.org/10.1016/j.procs.2012.04.227

[6] Bożejko, W., Pempera, J., & Wodecki, M. (2015). Parallel Simulated Annealing Algorithm for Cyclic Flexible Job Shop Scheduling Problem. W *Lecture Notes in Artificial Intelligence* (T. 9120, ss. 603–612). Springer International Publishing. https://doi.org/10.1007/978-3-319-19369-4_53

[7] Bożejko, W., Pempera, J., & Wodecki, M. (2016). Szybki algorytm rozwiązywania cyklicznego problemu gniazdowego. W *Automatyzacja procesów dyskretnych, Teoria i zastosowania (red. A. Świerniak, J. Krystek), Tom I* (ss. 49–56). Gliwice.

[8] Chaudhry, I. A., & Khan, A. A. (2016). A research survey: Review of flexible job shop scheduling techniques. *International Transactions in Operational Research*, *23*(3), 551–591. https://doi.org/10.1111/itor.12199

[9] Cheng, T. C. E., Wang, G., & Sriskandarajah, C. (1999). One-operator-two-machine flowshop scheduling with setup and dismounting times. *Computers and Operations Research*, *26*(7), 715–730. https://doi.org/10.1016/S0305-0548(98)00087-2

[10] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2007). *Wprowadzenie do algorytmów*. WNT.

INDEX OF AUTHORS

DMMS