The Digital Transformation: Challenges in Technological, Scientific and Social Development

Proceedings of the 7th FDIBA Conference
Sofia, Bulgaria, 30 November – 1 December, 2017

FDIBA Conference Proceedings

Volume 1, December 2017
ISSN 2535–132X (Print)
ISSN 2535–1338 (Online)

Die Digitale Transformation – Herausforderungen im technologischen, wirtschaftlichen und sozialen Wandel

Tagungsband der 7. FDIBA Konferenz
Sofia, Bulgarien, 30. November – 1. Dezember, 2017

Tagungsbände der FDIBA Konferenz

Volumen 1, Dezember 2017
ISSN 2535–132X (Gedruckt)
ISSN 2535–1338 (Online)

Technische Universität Sofia
Karlsruher Institut für Technologie
Hochschule für Ökonomie und Management Essen
Technische Universität Braunschweig
Otto von Guericke Universität Magdeburg

With the financial support of the German Academic Exchange Service
DAAD
Mit der Förderung des Deutschen Akademischen Austauschdienstes

Technical Program Committee

Chair: Alexandar Tsenov

Thomas Heupel
Sebastian Zug
Peter Lang
Jivka Ovcharova
Hristomir Yordanov

Organisation Committee

Chair: Hristomir Yordanov

Alexandar Tsenov
Yordan Yordanov
Evegniy Petrov
Marin Marinov
Nadezhda Kazakova

Reviewers

Virginiya Dimitrova
Vasil Galabov
Zwetelina Gankova - Ivanova
Veska Georgieva
Sasho Guerov
Thomas Heupel
Velko Iltchev
Stefan Kirianov
Marin Marinov
Thomas Mücke
Tashko Nikolov
Sanel Purgic
Stefan Stefanov
Aleksandar Tsenov
Elena Varbanova
Hristomir Yordanov
Sebastian Zug

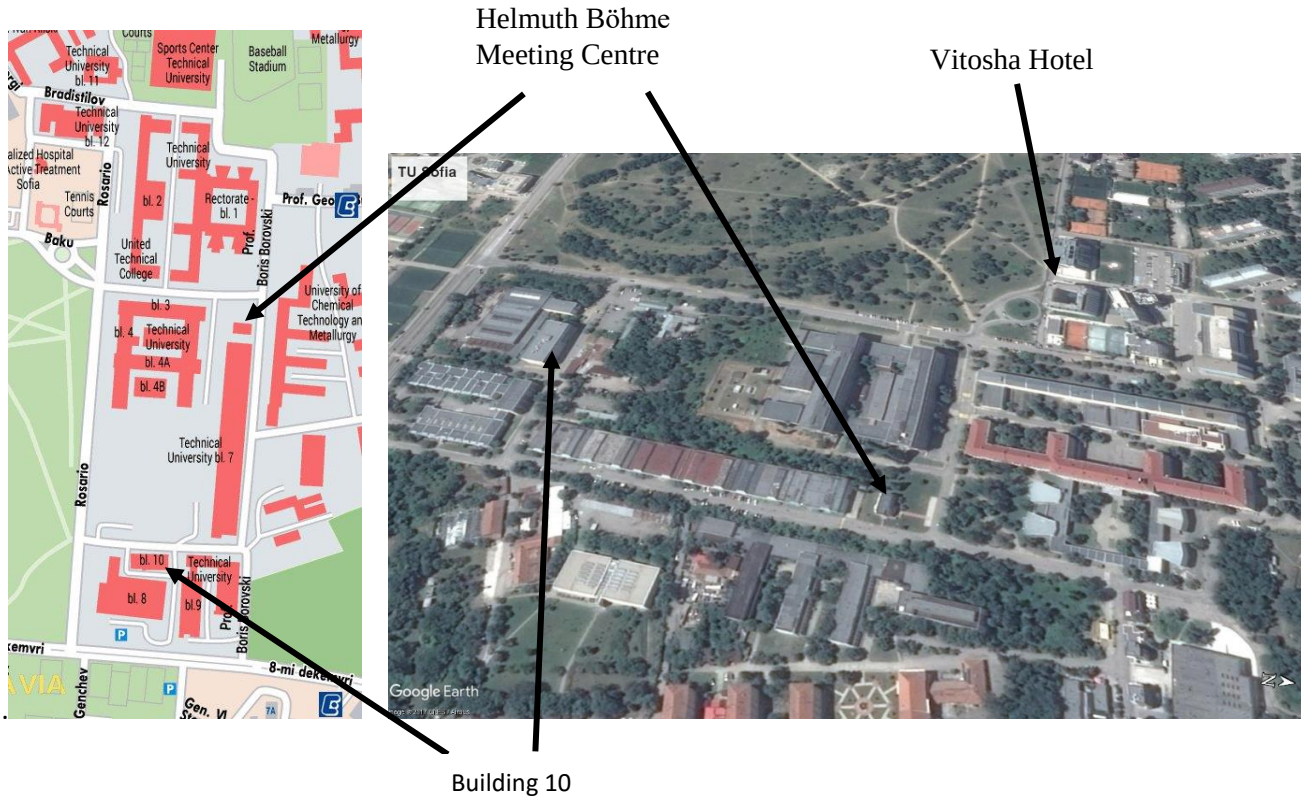
The Digital Transformation: Challenges in Technological, Scientific and Social Development
Proceedings of the 7th FDIBA Conference, Sofia, Bulgaria, 30 November – 1 December, 2017
FDIBA Conference Proceedings, 2017, Volume 1
Alexandar Tsenov, Thomas Heupel, Sebastian Zug, Peter Lang, Jivka Ovcharova, Hristomir Yordanov (Eds.)
Book size A4-ISO (210 × 297 mm), 140 pages
Paperback
Published by TU–Sofia Publishing house
Printed by TU–Sofia Publishing house
Sofia, Bulgaria
November 2017
ISSN 2535–132X (Printed)
ISSN 2535–1338 (Online)

FDIBA Conference Program

30. 11. – 1. 12. 2017

Venue

The conference will be held in two locations in the campus of the Technical University of Sofia. The first is the conference room of the Helmut Böhme Meeting Centre, where the opening and the sessions on Economics and Management and Mechanical Engineering will take place. The sessions on Computer Science and Education will be held in room 10100 in Building 10. The same building, room 10202, will host the poster session. The banquet will be held at the Vitosha Hotel restaurant.



Wireless Internet Access

Free Wi-Fi will be provided to all conference participants during the event at both locations. The name of the network is FDIBA Conference 2017. Please take into consideration that data is not encrypted over this network and using only secure protocols is recommended. You can also use the eduroam network if your institution participates in eduroam (not available in Helmut Böhme Meeting Centre).

Sessions

Thursday, 30. 11. 2017

The Helmut Böhme International Meeting Centre (IBZ)	
17:00	Official Opening Ceremony
17:30	**Plenary Talk** **Prof. Jivka Ovtcharova, KIT** *Spielerische Umsetzung von Industrie 4.0 bei KMU: Risiken minimieren und Chancen erkennen*
19:30	**Banquet** **Vitosha Hotel**

Friday, 01. 12. 2017 morning

Raum		10100		IBZ	10202
9:00	Session INF1 Chair: Velko Iltchev Co-chair: Sebastian Zug	**Sebastian Zug, Anja Hawlitschek, Till Krenz** *What are the key features of future Remote Labs? A critical evaluation of an existing one*	Session BWL1 Chair: Stefan Stefanov Co-chair: Thomas Heupel	**Invited talk:** **D-r Peter Greulich** *Quality Security and Product Liability with examples from the European Automotive industry*	
9:20		**Veska Georgieva, Plamen Petrov, Antoniya Mihaylova** *GUI for Circular and Elliptic Objects Detection in Digital Images*			
9:40		**Petar N. Nikolov, Vassil T. Galabov** *Experimental realization of Controlled Square Root of Z Gate Using IBM's Cloud Quantum Experience Platform*		**Serap Demiröz, Thomas Heupel** *Digital Transformation and its radical changes for external management accounting: a consideration of small and medium-sized enterprises*	
10:00		**Vanesa Georgieva** *Graph partitioning – comparison of naïve and evolutionary approach*		**Ralitza Dimitrova** *The Requirement to Deceive – a Serious Shortcoming of the Bulgarian Computer Fraud Regulation*	Poster Session Chair: Zvetelina Gankova **Julieta Kaleicheva, Valentin Mishev, Zdravka Karaguiozova**, *Effect of Nanoadditives on the Characteristics of Bainitic Transformation in Spheroidal Graphite Cast Irons* **Ognyan Kamenov, Magdalina Uzunova** *Bidirectional travelling surface waves in a convecting fluid* **Magdalina Uzunova** *Solitary and Cnoidal Waves in the Equation of a Convecting Fluid* **Zwetelina Gankova - Ivanova, Svetlozar Todorov** *Transparent Governance for Corruption Prevention* **Julieta Kaleicheva, Zdravka Karaguiozova, Valentin Mishev** *Enhancement of surface properties of steel samples using new technique for electroless nickel plating* **Zwetelina Gankova – Ivanova** *Shadow Economy, transparent E-Governance* **Georgi Yanachkov, Simona Hesapchieva** *Using Multilink Type Suspension as a Front Suspension For a Motorcycle* **Galina Nikolcheva, Julieta Kaleicheva** *Grouping of technological features using a clustering algorithm* **Stefan Kiranov** *Application of UniSim Design to study the static and dynamic properties of a chemical technology equipment*
10:20		**Coffee Break**		**Coffee Break**	
10:40	Session INF2 Chair: Sebastian Zug Co-chair: Veska Georgieva	**Blagovest Achanov** *Spatial random graphs generation- comparison of naïve and evolutionary approach*	Session BWL2 Chair: Thomas Heupel Co-chair: Stefan Stefanov	**Gero P. Gunkel** *Emanzipationsstrategien für übernahmeaffine Start-Ups*	
11:00		**Konstantina Dimitrova** *Knowledge discovery for business process optimization at Enterprise 2.0*		**Dirk Henrich, Thomas Heupel** *Steht Berlin 30 Jahre nach dem Mauerfall vor der „Technologiefalle“? Die Risiken der Digitalisierung für den Nachfolgeprozess!*	
11:20		**Kristian Nikolov, Boris Kirov, Vassil Galabov** *Automated Genome Indexing for Metagenomic Analysis/Automatisierte Genom-Indexierung für metagenomische Analyse*		**Sonja Schönauer, Christian Roth, Thomas Heupel** *Erhält das Beyond Budgeting vor dem Hintergrund der Digitalen Transformation sowie des anstehenden Generationenwechsels von Babyboomern zu Digital Natives eine neue Chance?*	
11:40		**Angel Garabitov** *Passive Voicequalitätanalyse Mit Melodischer Struktur Der Vorher Aufgezeichnete Telefonkonversation*		**Nikolay Kambosev** *X-Matrix Combining With Lean/Six Sigma For Company Great Results*	

12:00	**LUNCH** **Vitosha Hotel**

Friday, 01. 12. 2017 afternoon

Raum		10100		IBZ	10202
14:00	Session INF3 Chair: Veska Georgieva Co-chair: Velko Ilchev	**Virginiya Dimitrova** *Framework for implementing, evaluating and analyzing some algorithms for non-photorealistic image processing*	Session MB Chair: Vassil Galabov Co-chair: Thomas Mücke	**Thomas Mücke** *Challenges for new business models based on big-data analysis in the machining industry*	
14:20		**Vanya Markova** *Approach for knowledge transfer for building autonomous agent behaviour*		**Miglena Dontschewa, Denise Stamatov, Marin B. Marinov** *Konzept und Entwicklung einer VR-Anwendung*	
14:40		**Stefan Hensel, Marin B. Marinov** *Sensor data fusion for topological localization of rail vehicles*		**Sasho Guergov, Aleksandar Tsenov** *Nicht-traditionelles Verfahren zur Organisierung von Informationssteuerungssysteme für Intelligente Produktionsmaschinen*	
15:00		**Julia Popowa, Lucas Pfeiffer** *German as a foreign language and the mastery of relevant technical language as a requirement for a successful engineering education at FDIBA*		**Sasho Guergov, Aleksander Tsenov** *On the issue of technology process for intelligent machine tools*	
15:20		**Thomas Scheicher, Thomas Heupel** *Ways to identify influence factors of virtual reality-systems on the acceptance in teaching - differences in generations X, Y and Z.*		**Stefan Hensel, Marin B. Marinov** *Estimation of Magnetic Field Maps With Mobile Platforms*	
16.00				**ROUND TABLE**	
17.00	**CONFERENCE CLOSING**				

This page is intentionally left blank

Contents

Plenary talk **1**

Play-like Implementation of Industry 4.0 at SMEs: Minimizing Risks and Identifying Opportunities
Jivka Ovcharova 1

Session INF1 **3**

What are the key features of future Remote Labs? A critical evaluation of an existing one
Sebastian Zug, Anja Hawlitschek, Till Krenz 3

GUI for Circular and Elliptic Objects Detection in Digital Images
Veska Georgieva, Plamen Petrov, Antoniya Mihaylova 7

Experimental realization of Controlled Square Root of Z Gate Using IBM's Cloud Quantum Experience Platform
Petar N. Nikolov, Vassil T. Galabov 11

Graph partitioning – comparison of naive and evolutionary approach
Vanesa Georgieva 15

Session INF2 **19**

Spatial random graphs generation — comparison of naive and evolutionary approach
Blagovest Achanov 19

Knowledge discovery for business process optimization at Enterprise 2.0
Konstantina Dimitrova 23

Automated Genome Indexing for Metagenomic Analysis
Kristian Nikolov, Boris Kirov, Vassil Galabov 27

Passive voice quality analysis with melodic structure of previously recorded telephone conversations
Angel Garabitov 29

Session INF3+EDU **33**

Framework for implementing, evaluating and analyzing some algorithms for non-photorealistic image processing
Virginia Todorova Dimitrova 33

Approach for knowledge transfer for building autonomous agent behaviour
Vanya Markova, Ventseslav Shopov 37

Sensor data fusion for topological localization of rail vehicles
Stefan Hensel, Marin B. Marinov 41

German as a foreign language and the mastery of relevant technical language as prerequisites for a successful engineering education at FDIBA
Julia Popowa, Lucas Pfeiffer 45

Ways to identify influence factors of virtual reality-systems on the acceptance in teaching - differences in generations X, Y and Z
Thomas Scheicher, Thomas Heupel 47

Session BWL1 **51**

Quality Security and Product Liability with Examples from the European Automotive Industry
Invited talk
Peter Greulich 51

The Requirement to Deceive – a Serious Shortcoming of the Bulgarian Computer Fraud Regulation
Ralitza Dimitrova 53

Digital Transformation and its radical changes for external management accounting: a consideration of small and medium-sized enterprises
Serap Demiröz, Thomas Heupel 57

Session BWL2 **61**
Accretion strategies referencing start-ups inclined to affiliation
Gero Peter Gunkel . . . 61
Berlin faces the technology trap 30 yeas after the fall of the wall — the digital transformation influences the follow-up process
Dirk Henrich, Thomas Heupel . . . 65
Do Digital Transformation and generation change from baby boomer to digital natives enable Beyond Budgeting a new chance?
Sonja Schönauer, Christian Roth, Thomas Heupel . . . 69
X-Matrix Combining with Lean/Six Sigma for Company Great Results
Nikolay Kambosev . . . 73

Session MB **77**
Challenges for New Business Models Based on Big-Data Analysis in the Machining Industry
Thomas Mücke . . . 77
Concept and Development of a VR-Application
Miglena Dontschewa, Denise Stamatov, Marin B. Marinov . . . 81
Non-traditional procedure for the organization of information control systems for intelligent production machines
Sasho Guergov, Aleksandar Tsenov . . . 85
On the issue of technology process for intelligent machine tools
Sasho Guergov, Aleksandar Tsenov . . . 89
Estimation of Magnetic Field Maps With Mobile Platforms
Stefan Hensel, Marin B. Marinov . . . 93

Poster Session **97**
Effect of Nanoadditives on the Characteristics of Bainitic Transformation in Spheroidal Graphite Cast Irons
Julieta Kaleicheva, Valentin Mishev, Zdravka Karaguiozova . . . 97
Bidirectional travelling surface waves in a convecting fluid
Ognyan Y. Kamenov, Magdalina V. Uzunova . . . 101
Solitary and Cnoidal Waves in the Equation of a Convecting Fluid
Magdalina Uzunova . . . 105
Transparent Governance for Corruption Prevention
Zwetelina Gankova–Ivanova, Svetlozar Todorov . . . 109
Enhancement of surface properties of steel samples using new technique for electroless nickel plating
Julieta Kaleicheva, Zdravka Karaguiozova, Valentin Mishev . . . 113
Shadow Economy, transparent E-Governance
Zwetelina Gankova-Ivanova . . . 117
Using Multilink Type Suspension as a Front Suspension for a Motorcycle
Georgi Yanachkov, Simona Hesapchieva . . . 121
Grouping of technological features using a clustering algorithm
Galina Nikolcheva, Julieta Kaleicheva . . . 124
Application of UniSim Design to study the static and dynamic properties of a chemical technology equipment
Stefan Kiranov . . . 127

Play-like Implementation of Industry 4.0 at SMEs: Minimizing Risks and Identifying Opportunities

Spielerische Umsetzung von Industrie 4.0 bei KMU: Risiken minimieren und Chancen erkennen

Plenary Talk

Jivka Ovcharova

Institute for Information Management in Engineering, Karlsruhe Institute of Technology
Karlsruhe, Germany, e-mail: jivka.ovcharova@kit.edu

Zusammenfassung — Für die Digitalisierung gibt es keine Lösung von der Stange, sowohl für die großen, als auch für die kleinen und mittelständischen Unternehmen (KMUs). Jeder Betrieb ist einzigartig, jede Lösung individuell. Gerade hier weisen KMUs Riesenvorteile auf. Sie sind ,,von Natur" her pragmatisch und anpassungsfähig - noch aus den geringsten Optionen und Ressourcen versuchen sie das Beste herauszuholen. Die Dynamik und Variabilität der Veränderungen trifft in besonderem Ausmaß deren ,,Keimzelle" — die Arbeitsgestaltung. Die wichtigste Chance wird daran gesehen, dass die digitalen Technologien durch Mensch-Maschine-Kollaboration in Co-Working-and-Sharing-Spaces und im globalen Multi-Unternehmertum angetrieben werden. ,,Right-Brainers" - Menschen mit digitaler Qualifikation, der Fähigkeit des vernetzten Denkens und Handelns und mit dem Blick für das große Ganze – werden wie nie zuvor gefragt.

In einer Art ,,Sandkasten" werden neue Ideen auf einfache und spielerische Weise ausprobiert und durch weitreichende Bildungs- und Qualifizierungsmaßnahmen untermauert. Die spielerische Umsetzung symbolisiert dabei einen Ort um zusammen zu kommen, einen Ort zu denken, auszuprobieren und spielerisch zu schaffen. Ein Sandkasten wird als Gestaltung- aber auch Marketing- und Einkauf Think Tank angesehen, wo Menschen bereichs- und unternehmensübergreifend ein gemeinsames Instrumentarium nutzen um zusammenarbeiten, um Lösungen zu kreieren, diese umzusetzen und zu vermarkten, ganz nach dem Motto ,,face it¡', d.h. auf den Punkt kommen, begreifen, pragmatisch und schnell vorgehen, durch Handlung.

Das Profit der Unternehmen wird dadurch gesichert, dass im einem ,,Sandkasten" Investitionen nur dann betätigt werden, wenn der Einsatz durch messbare Vorteile belegt wird, ganz nach dem Motto ,,Einsatz-Profit-Beschaffung" anstatt der üblichen Vorgehensweise: ,,Beschaffung-Einsatz-Profit". Das Investitionsrisiko wird minimiert, bevor Ideen in einer breiteren Skala operativ umgesetzt werden. Dieses Konzept wird gerade bei der Firma ELABO aus Crailsheim, Baden-Württemberg im Rahmen des Shopfloor Execution System (SES)-Einsatzes erfolgreich umgesetzt.

Abstract — There is no off-the-peg solution for digitalization, for both large as well as small and medium-sized enterprises (SMEs). Every business is unique, every solution individually crafted. In this sense, SMEs demonstrate huge advantages. Those companies are "inherently" pragmatic and adaptable – they try to make progress even with limited resources and options. The dynamic and variable nature of these changes is linked to the "core" advantage of SMEs – the work design. The most important opportunity for promoting digital technologies is found in human-machine collaboration in co-working-and-sharing spaces and in global multi-entrepreneurship. The demand for "right-brainers"—people with digital skills, network thinking and action, as well as the ability to see the "big picture"—is higher than ever before.

In a kind of a sandbox, new ideas are tried out in a simple and play-like manner, which is underpinned by overarching education and qualification measures. The play-like implementation symbolizes a place to come together, to think, to try out, and to playfully create a place. The sandbox is seen as a design, as well as a marketing and buying think tank, where people use a common tool across departments and companies to work together in order to create, implement and market solutions. This is based on the motto "face it!", i.e. get to the point, understand, act in a pragmatic and fast way.

In the sandbox model, enterprises secure their profits because investments are only activated if the new venture has proved to have measurable advantages. This process exists under the motto "use – profit – procurement" instead of the usual procedure "procurement – use – profit." Investment risk is minimized before ideas are operationally implemented on a larger scale. This concept is currently being successfully applied by ELABO from Crailsheim, Baden-Württemberg as part of their Shopfloor Execution System (SES) deployment.

This page is intentionally left blank

What are the key features of future Remote Labs? A critical evaluation of an existing one

Was muss ein Remote Lab zukünftig leisten?

Sebastian Zug* and Anja Hawlitschek† and Till Krenz‡

* Otto-von-Guericke University, zug at ovgu.de
† Magdeburg-Stendal University of Applied Sciences, anja.hawlitschek at hs-magdeburg.de
‡ Otto-von-Guericke University, till.krenz at ovgu.de

Abstract — The authors are convinced that the upcoming fourth generation of Remote Labs has to adress the heterogeneity of the students especially. In support of this expectation we investigated an existing implementation and discuss the technical and didactical consequences.

Zusammenfassung — Die vierte Generation von Remote Labs wird sich in stärkerem Mae als bisher der Heterogenität der Studenten annehmen müssen. Zur Begründung dieser These wurde eine existierende Implementierung evaluiert und davon ausgehend die technischen und didaktischen Konsequenzen diskutiert.

I. INTRODUCTION

A key feature of the university program for engineering education is a systematic deepening of the theoretical knowledge by practical training sessions. Students have to design and to manufacture exemplary objects, to execute experiments in the laboratory or to implement small IT applications. These supplemental elements of the education process provide specific competencies (programming languages, usage of tools) and train the understanding of engineering concepts and strategies (i.e., divide and conquer). Larger student projects offer the opportunity to improve management skills and illustrate the challenges of interdisciplinary teams. It was shown that these experiences generate an additional motivation for students [1].

Nevertheless, the implementation of practical tasks is accompanied by technical and didactical challenges. First of all, the tasks have to be designed and prepared carefully to meet the capabilities and the prior knowledge of the students [2]. Additionally, the provision of the required hard- and software is often expensive and make a continuous monitoring and maintenance service necessary. Depending on the available number of instances an additional coordination effort for organizing the access to the devices is necessary.

A common approach to cope with these disadvantages is the concept of remote labs, an internet-based interface to educational hardware setups. Based on a camera stream, the student can control a machine, an experimental setup or a computer and evaluate the success of his/her commands. This position and time-independent access increases the utilization of the device and reduces the maintenance effort because the system avoids the error-prone direct access of the students. A remote lab combines the comfort of a simulation environment but provides a realistic experience by covering all situations that may occur in a real development process [3]. But of course, the preparatory overhead increases again, the lecturer has to implement an appropriate back- and front-end for a successful integration. To reduce this effort different frameworks were developed offering a generic and hardware independent devices access [4]. Based on the homogeneous interfaces universities started to offer common remote labs.

The working group Embbeded Smart Systems (ESS) of the OVGU implemented such a modular Remote Lab and integrated it within lectures for embedded systems since two years. Each season approximately 60 students participate in the course "Principals and Components of embedded Systems" that imparts the basics of software development for micro-controllers and peripherals. While analyzing the teaching evaluation report (a form that is filled by students at the end of each semester for every lecture), we got positive feedbacks but criticisms too. On the one hand, the students emphasized the possibility to access the hardware permanently and praised the availability of all data-sheets at one point. The first aspect is visible in Fig. 1. 26 percent of student's activities are outside of the usual office time. Even though the evening and night hours are the least used periods, these hours make up about 20 Percent. Hence, it is obvious, that one of the main goals was reached, the students plan the tasks now according to individual working habits and personal timetables.

**This work was partitially supported by the German Federal Ministry of Education and Research (BMBF, Funding number: 16DHL1033) and the German Academic Exchange Service (DAAD, Funding number: 57363292).

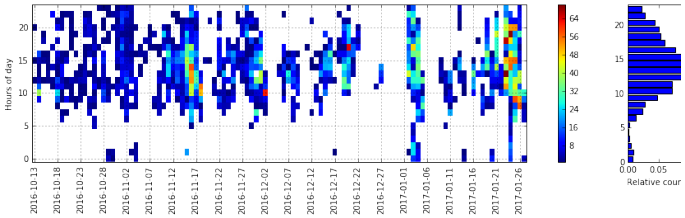


Fig. 1. Representation of log-ins to the system during the semester distributed about the hours of a day.

On the other hand, the participants were not completely convinced by the idea of the remote laboratory at all. They argued, that some of the tasks would require direct access to the hardware. In this situation, a real visit of the laboratory would be necessary to check the behavior of the system on-site. Additionally, the students addressed the susceptibility to errors of the system and asked for a closer interaction with the lectures. Consequently, we started a detailed investigation by evaluating the log files of the web front end and by aggregating additional information based on a questionnaire. This paper presents the main results of this analysis and derives resulting consequences for a further improvement of our setup.

II. Existing Remote Lab for Embedded Systems

The undergraduate lecture for "Principles and Components of Embedded Systems" (PKeS) is given for students of Computer Science and Electrical Engineering. It intends to close the gap between lectures in computer architectures, electrical basics and software engineering for embedded devices. For this purpose the lecture addresses topics as Analog to Digital Converter (ADC), Pulse Width Modulation (PWM) or Real Time Operation Systems (RTOSs). The theoretical knowledge is consolidated by a set of 6 practical exercises implemented on an 8-Bit-Micro-controller ATmega 2560 manufactured by Atmel.

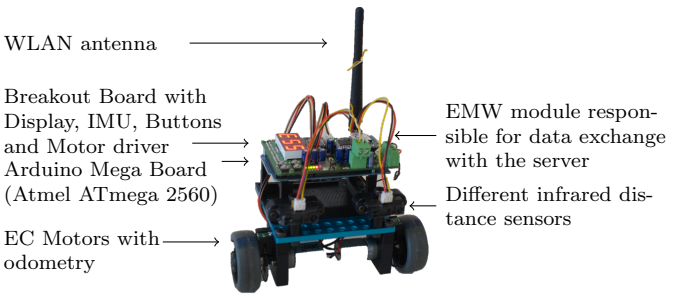


Fig. 2. One of the remote programmed robot systems with an AVR controller as a target

We integrated the controller (integrated into an open sources Arduino board) in a small robot system, containing different types of sensors (Infra-red distance sensors, buttons, and an Inertial Measurement Unit (IMU)), two motors, a 3x7 segment display, as visible in Fig. 2. The practical tasks range from a mandatory "Hello World Task" over a display driver up to robot's escape from a maze. Each task has to be implemented in assembly of c/c++ programs by single students or groups of two persons within two or three weeks. An experienced tutor evaluates the result and the quality of the code during the exercises.

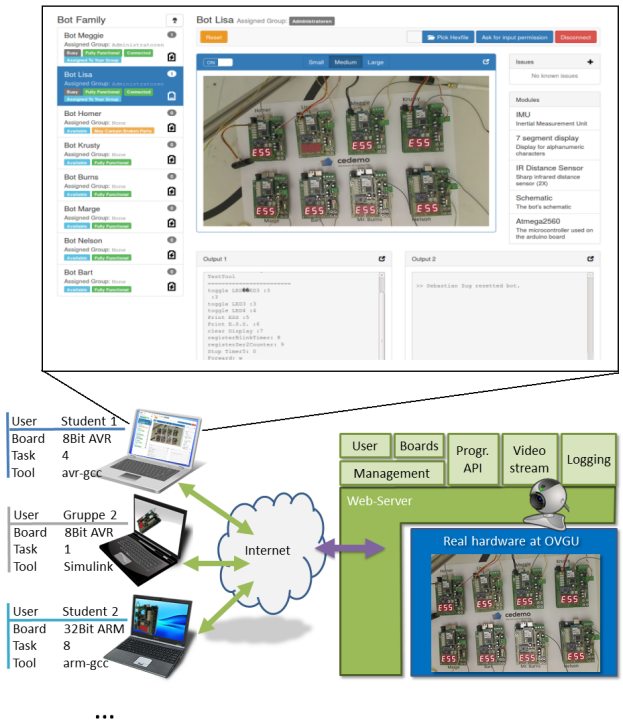


Fig. 3. Core elements of the existing implementation and screenshot of the web-frontend, it contains the video stream, management sections, textual interfaces for debugging purposes and a link collection to hardware data sheets

During the last decade, the robot platform was changed four times. For all implementations the reliability of the platforms was a challenging problem. When students worked directly with the systems, we missed a careful dealing with the setup. Robots crashed with the walls, were "adapted and improved" by students or left with low voltage batteries. Additionally, during the days before a deadline, the tutors have to manage the increased demand for robots.

Consequently, we decided to transform the project in a remote lab. Based on an extension of robots hardware and a self-implemented front- and back-end the robots offer a programming interface to a web-page. The server infrastructure contain robot and user management modules, the programming API, algorithms for video stream processing and encoding as well as a logging data base system.

The screenshot of the webpage is visible in the upper part of Fig. 3. On the left side, eight controllers are available at the moment. Their names and robot's current situation, like energy state or mentioned failures are visible. The students are free to choose one of the available devices. The system

interrupts the connection automatically after 20min or if no actions are registered for 10min. One the right side the web page lists some links to the relevant datasheets and tutorials. The center of the page integrates the video stream. In this screenshot, taken during the implementation of the first task, we concentrate all controllers in front of one camera. For this task the access to the whole robot periphery is not needed. With the three buttons on the right side the students can upload their `.hex` file and activate the light in the laboratory.

III. EVALUATION

A. Objectives and procedures

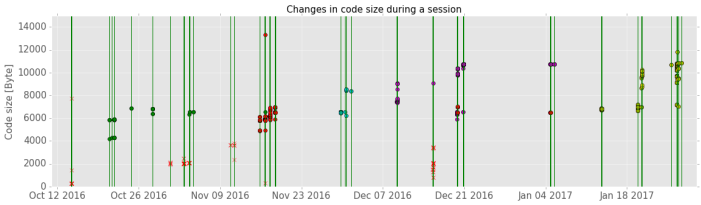
With an empirical study we want to examine the heterogeneity of our students concerning previous knowledge on embedded systems, the behavior of the students in the system and attitudes towards the Remote Lab in comparison with a Hands-on Lab. Accordingly, our research questions were:

- Are there any correlations between the previous knowledge and the learning performance?
- Is there a relation between user behavior and learning performance?
- How do students evaluate the Remote Lab in comparison with a Hands-on Lab? Are there any challenges we should consider for further versions of the Remote Lab or in further research?

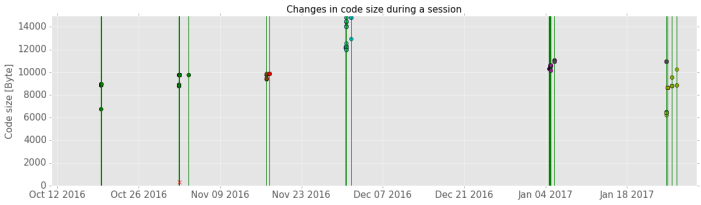
To examine our research questions, we carried out an evaluation with 34 participants of the course PKeS. Therefore we used the log-files of the practical exercises in the Remote Lab and a digital survey, administered after the end of the semester. In the log-files students' interactions with the system were recorded, for example the logins and the upload and size of the code. In the survey the students could rate there prior knowledge concerning embedded controllers/boards and robot applications on rating scales from 1 (very low) to 5 (very high). Furthermore, they were asked to evaluate the Remote Lab in comparison to a Hands-on Lab. On the one hand, they could rate advantages or disadvantages of Remote Labs on a rating scale. On the other hand, they had the possibility to write down opinions and remarks as elaborated text.

B. Evaluation of the Log-Files

In this paper we focus the evaluation part on the heterogeneity of the students related to prior knowledge and coding experience. Fig. 4 and 5 illustrate this aspect based on two exemplary data sets. The first one contrasts the programming activities based on the flash operations by a student with a low prior knowledge and an experienced one. While the x-axis represents the time window of the semester the y-axis displays the size of the executable code transferred to the robot. The student of the development progress visible in (b) implements the task in fewer sessions than the student of (a). Additionally, the spread of the code size in lower in the second case. The style of programming of both students



(a) Student with very low prior knowledge



(b) Experienced Student

Fig. 4. Comparison of the individual progress of different students for tasks 1-5

was investigated in Fig 5 more in detail. The diagrams depict the change of the code size related to the duration this step. The second student developed the program in small steps, while the first had obviously copied a lot of lines. We cannot analyze the source of this operation, but we recognized that many students prefer to search for existing implementations from previous semesters as to write their own code.

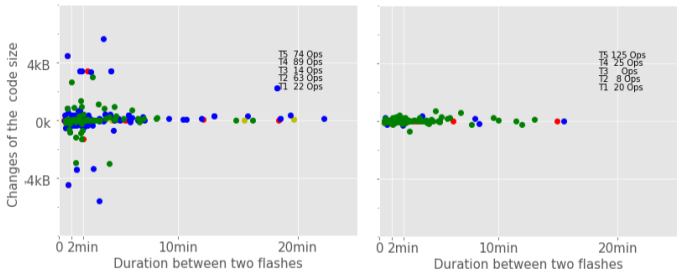


Fig. 5. Number of flash operations vs. corresponding code size changes for the exemplary students

C. Analysis of a Questionnaire

The prior knowledge of our students on embedded controllers/boards (M: 1.92, SD: 1.10) as well as robot applications (M: 1.68, SD: 1.02) is rather low. However, there is also a quarter of students which state that their prior knowledge is rather high or very high. There is a correlation between prior knowledge on embedded controllers and robot application ($r = .56$, $p < .001$). Someone who has high knowledge on robot applications has high knowledge on embedded controllers as well and vice versa. Both items also correlate with the learning performance. The better the knowledge about robot applications ($r = -.34$, $p < .05$) and embedded controllers/boards ($r = -.39$, $p < .05$) before the course, the better the grade in the exam afterwards. The results of a rating scale comparing the working experience in the Remote Lab to a regular Laboratory (Figure 6) exhibit perceived advan-

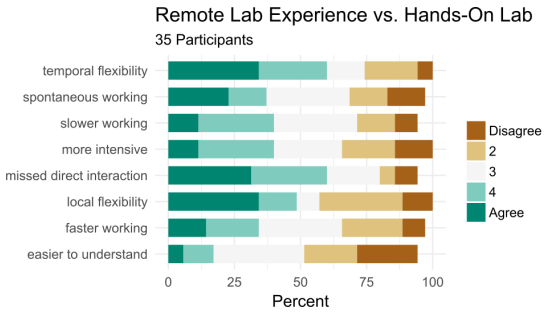


Fig. 6. Students compare their Remote Lab experience to regular Laboratories

tages of the Remote Lab setting, as gained temporal and local flexibility. At the same time problems are highlighted that pose hurdles to overcome in the upcoming iterations of the Remote Lab, such as bridging the missing direct contact to the robots by improving their reliability and extending the didactic integration in order to better the learning experience in the Remote Lab.

These outcomes are also supported by open text answers given by the participants. The answers offer some additional insights into the the perceived advantages and disadvantages of the Remote Lab. They repeatedly include the additional flexibility and improved interaction with the robots due to the the Remote Lab situation. Technical problems are specifically named in these answers, for example the instability of the video stream, reliability of the robots and missing debugging capabilities in the code editor. These answers are particularly useful, as they highlight the importance of improvements that are already prepared in the upcoming implementation of the Remote Lab.

IV. CONCLUSIONS AND FUTURE REQUIREMENTS

The results of our evaluation implicate that students are willing to learn with Remote Labs and do not particularly miss the hands-on approach. However, the students mentioned some disadvantages of Remote Labs, especially aspects of usability. If the system is not working properly decreasing motivation and low technology acceptance might be the result. Therefore, the usability of the Remote Lab should be as high as possible.

Our results also indicate a relation between prior knowledge and learning. To overcome the determined heterogeneity, the Remote Lab has to provide additional learning material tailored for the specific needs. This may include an extended documentation, instructional videos or tutorials. As part of the ongoing project the feasibility of multiple adaptive alternatives will be tested and implemented. In the upcoming iteration of the Remote Lab students will be classified into groups of beginner and more-skilled students. Both of these groups receive the same set of exercises, but the beginners will start of with more pre-populated clues about how to solve the tasks. The more advanced students on the other hand get the chance to start from scratch, allowing them to put their established competences to use [5]. Extended forms of adaptive learning, as the automatic detection of students that are stuck and the administration of tailored assistance will be studied on the basis of the use logs and further questionnaires. Another vector for adaptation that will be tested in future implementations are Learning Analytic Dashboards that allow the teachers and students to reflect upon the learning experience and (self-)adapt accordingly.

In order to establish an infrastructure that motivates the re-usability in other contexts the integration of the Remote Lab interface into e-Learning platforms (e.g. Moodle) is another ambition of this project. If the Remote Lab interface was available as a plugin for popular e-Learning environments this would allow for a better didactic integration into the topical context of the respective course. The existing structural support of these platforms could be used as well, for example the time tabling functions could be used in order to reserve time slots in the laboratory, communication with the teachers and within the group could rely on the existing capabilities and forums and wikis could aid in the cooperative development and exchange of knowledge and competences.

To sum up, we would like to make some suggestions for implementing Remote Labs in the future:

- To increase learning performance it is suitable to implement adaptive Remote Labs for tailored instruction.
- The usability of Remote Labs is of particular importance to sustain motivation.
- An integration of a Remote Lab into a learning management system provide additional benefit for learner and teacher.

However, our research is still at the beginning. In further research we have to gain much more detailed information on challenges as well as advantages of Remote Labs and an optimal instructional design for specific objectives and learners.

References

[1] D. Lopato, "Undergraduate research experiences support science career decisions and active learning," *CBE Life Science Education*, vol. 6, no. 4, pp. 297–306, 2007.

[2] S. Kalyuga, P. Ayres, P. Chandler, and J. Sweller, "The expertise reversal effect," *Educational Psychologist*, vol. 38, no. 1, pp. 23–31, 2003.

[3] B. Balamuralithara and P. C. Woods, "Virtual laboratories in engineering education: The simulation lab and remote lab," *Computer Applications in Engineering Education*, vol. 17, no. 1, pp. 108–118, 2009.

[4] P. Tröger, A. Rasche, F. Feinbube, and R. Wierschke, "SOA meets robots-A service-based software infrastructure for remote laboratories," *International Journal of Online Engineering*, vol. 4, no. 2, pp. 24–30, 2008.

[5] T. Clarke, P. Ayres, and J. Sweller, "The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications," *Educational Technology Research and Development*, vol. 53, no. 3, pp. 15–24, 2005.

GUI for Circular and Elliptic Objects Detection in Digital Images

GUI für kreisförmigen und ellipsenförmigen Objektserkennung in den digitalen Bildern

Veska Georgieva*, Plamen Petrov†, Antoniya Mihaylova*,
* Faculty of Telecommunications, Technical University of Sofia,
Sofia, Bulgaria, e-mails: vesg@tu-sofia.bg, antonianankova@gmail.com
† Faculty of Mechanical Engineering, Technical University of Sofia,
Sofia, Bulgaria, e-mail: ppetrov@tu-sofia.bg

***Abstract* —** The paper presents software for automatic detection of objects with circular and elliptic forms and its graphical user interface (GUI). It works in the MATLAB environment and uses IMAGE TOOLBOXES defined functions. The circle objects can be detected and located on the base of Circular Hough Transform (CHT). For elliptic shape detection the Randomized Hough Transform (RHT) is used. The program proposes also an interactive option to choose specific parameters and works with the region of interest (ROI) in the processed image. The proposed GUI can be applied to real medical and microscopy images attempt to make diagnostic more precise. It can be used in computer vision systems for detection of different objects with circular and elliptic form. The presented GUI is suitable also to engineering education for studying of this processing.

***Zusammenfassung* —** Der Artikel präsentiert eine Software zur automatischen Erkennung von Objekten mit kreisförmigen und elliptischen Formen und deren grafische Benutzeroberfläche (GUI). Es funktioniert in der MATLAB-Umgebung und verwendet IMAGE TOOLBOXES definierte Funktionen. Die Kreisobjekte können auf der Basis von Circular Hough Transform (CHT) erkannt und lokalisiert werden. Für die elliptische Formdetektion wird die randomisierte Hough Transform (RHT) verwendet. Das Programm schlägt auch eine interaktive Option vor, um bestimmte Parameter auszuwählen und arbeitet mit dem interessierenden Bereich (ROI) im verarbeiteten Bild. Die vorgeschlagene GUI kann auf reale medizinische und mikroskopische Bilder angewendet werden, um die Diagnose präziser zu machen. Es kann in Computer-Vision-Systeme für die Erkennung von verschiedenen Objekten mit kreisförmigen und elliptischen Form verwendet werden. Die dargestellte GUI eignet sich auch für die Ingenieurausbildung für das Studium dieser Verarbeitung.

I. Introduction

Circle and ellipse detection has been an important task in image processing and pattern recognition. Many approaches have been developed to solve this problem. There are many techniques, which are based on Hough transform (HT) and its variants [1-4]. The HT and its variants have been successfully used in many problems. A major advantage of HT is its insensitivity to imperfect data. HT is able to detect the geometry shapes such as lines, circles and ellipses in various situations. But computationally it is inefficient, which increases with the complexity of shapes [5].

In this paper, we present software for automatic detection of objects with circular and elliptic forms and its graphical user interface (GUI). The circle objects can be detected and located on the base of Circular Hough Transform (CHT). For elliptic shape detection the Randomized Hough Transform (RHT) is used. The algorithm proposes also an interactive option to choose specific parameters and works with the region of interest (ROI) in the processed image.

The remainder of this paper is organized as follows. In the Section 2, the briefly theoretical part is given. In Section 3, the GUI for circle and elliptic objects detection is presented. Section 4 describes tasks carried out from the main program and some experimental results, obtained by computer simulation. Concluding remarks are given in Section 5.

II. Theoretical part

The CHT is useful for detecting circles of known radius as well to detect circles of various radii. This method is based on creating an accumulator matrix of size of the original image to be processed. The local maxima in accumulator space are obtained by voting procedure. Parameter space is defined by the parametric representation used to describe circles in the picture plane [6], which is given by

$$r^2 = (x - x_0)^2 + (y - y_0)^2 . \qquad (1)$$

It implies that the accumulator space is three-dimensional (for three unknown parameters x_0, y_0 and r) and defines a locus of points (x, y) centered on an origin (x_0, y_0) with radius r. Points corresponding to x_0, y_0 and r, which has more votes, are considered to be a circle with center (x_0, y_0) and radius r.

The advantage of RHT in comparison with classical HT is that some curves can be fully determined by a certain number of points on the curve [7, 8]. For example, a straight line can be determined by two points, and an ellipse (or a circle) can be determined by three points. So the time efficiency can be improved and the storage requirement of the original algorithm will be reduced. The whole process generally consists of three steps [9]:

1. Fit ellipses with randomly selected points.
2. Update the accumulator array and corresponding scores.

3. Output the ellipses with scores higher than some predefined threshold.

The general equation, defined for ellipse in [9] is given by

$$a(x-p)^2 + 2b(x-p)(y-q) + c(y-q)^2 = 1 \quad (2)$$

with restriction: $ac - b^2 > 0$.

RHT starts by randomly selecting three points on the ellipse. By first step the tangents of these three points are founded. They can be found by fitting a straight line using least squares technique for a small window of neighboring pixels.

The next step is to find the intersection points of the tangent lines. With the ellipse parameters determined from previous stage, the accumulator array can be updated correspondingly. The RHT does not keep "grid of buckets" as the accumulator array in comparison with classical HT. Rather, it first calculates the similarities between the newly detected ellipse and the ones already stored in accumulator array. As long as the similarity exceeds some predefined threshold, replace the one in the accumulator with the average of both ellipses and add 1 to its score. Otherwise, initialize this ellipse to an empty position in the accumulator and assign a score of 1.

Once the score of one candidate ellipse exceeds the threshold, it is determined as existing in the image and should be removed from the image and accumulator array so that the algorithm can detect other potential ellipses faster. The algorithm terminates when the number of iterations reaches a maximum limit or all the ellipses have been detected.

III. GUI for Circular and Elliptic Objects Detection

The GUI for circular and elliptic objects detection is presented in Fig.1. It is divided in several areas, where the user applies different settings, concerning different modes for object detection, theirs parameters and areas for universal application.

The area "Select Image" is for entering an image file name with an image file extension. The user can navigate among the folders in the work folder and choose image by using "Open" button.

The Hough transform's settings are selected in area "Mode". Two specific modes "Circle" and "Ellipse" are used in regard to implementation of CHT and RHT, respectively.

By selecting of Sobel or Canny operator edge detection can be implemented.

In area "Radius" the maximal and minimal radius of the circles can be determinated. So we can find all the circles with radius r pixels in the range $[R_{min}, R_{max}]$. In the case of ellipse detection these parameters are associated with "minMajorAxis" and "maxMajorAxis".

In area "Object Polarity" we can indicate whether the circular objects are brighter or darker than the background. If the circular objects are brighter than the background, the option "bright" will be selected. If the circular objects are darker than the background, the option "dark" will be chosen.

The area "Fine Settings" can be used for selecting sensitivity factor and edge gradient threshold. Sensitivity factor is the sensitivity for the circular Hough transform accumulator array, specified as the comma-separated pair consisting of 'Sensitivity' and a nonnegative scalar value in the range [0, 1]. As we increase the sensitivity factor, more circular objects, including weak and partially obscured circles can be detected. Higher sensitivity values also increase the risk of false detection. Edge gradient threshold sets the gradient threshold for determining edge pixels in the image, specified as the comma-separated pair consisting of "Edge Thresh" and a nonnegative scalar value in the range [0, 1]. By "0" we can set the threshold to zero-gradient magnitude. Specifying "1" we can set the threshold to the maximum gradient magnitude.

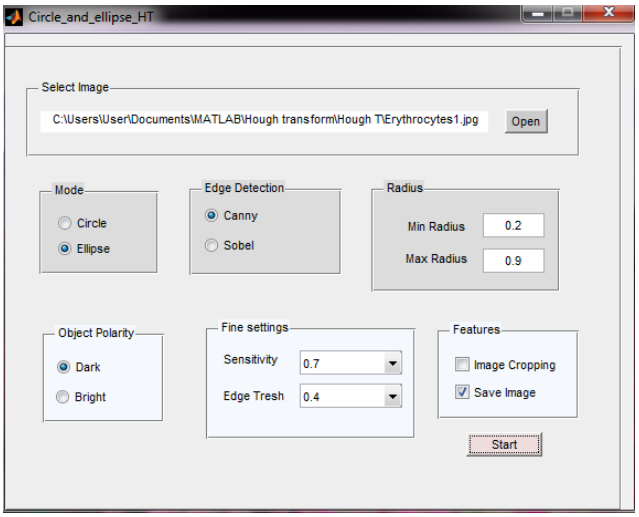


Fig. 1. Overview of GUI for circular and elliptic objects detection

By setting the threshold to a lower value we can detect more circular objects (with both weak and strong edges). By increasing the value of the threshold, fewer circles with weak edges can be detected.

The area "Features" present the universal application such as saving the result of processing and working with selected region of interest, which can be selected by "Image Cropping".

After choosing all input information the procedure of processing begins, when the user clicks on button "Start". Then the final result is shown – original image, and processed images.

IV. Tasks Carried Out from the Main Program

The main algorithm that works behind is shown in Fig.2. By acting of component from GUI a callback-function from the main program can be implemented.

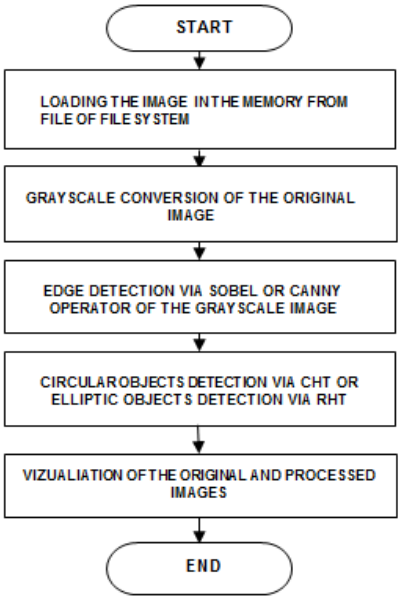


Fig. 2. The flowchart of the main algorithm

Every graphic component can be treated to object. Every object can be referred to handle. The objects referred a complex of attributes, which can be manipulated from the software. The multifarious attributes can be leaved by using in MATLAB environment, such as "Enabled", "Value", "Visible", "On", "Off" etc. Every attribute can be enabling in the presence of corresponding handle or reference to the object.

By initialization of the graphic application every graphic component can be reiterated to a cycle of events for the MATLAB environment. It submits addresses of the callback-functions, associated to a given event, which are important. By its identification the corresponding callback-function can be called out. The input data validation is one of the important tasks of the main program. If an error concerned with wrong information occurs, the execution is canceled. Another essential purpose of the main program is presenting the input information in appropriate data structures. It is necessary for the next steps in the processing strategy. In this step the processing with appropriate input data is made.

Because of separately processing of images, the program can work with single image as well with image sequence. Input image is the image in which we detect circular or ellipse objects, specified as a grayscale, true color, or binary image.

V. Experimental Results

Some results from simulation, which illustrate the work of the program, will be presented in the following.

The original grayscale microscopy image of embryos of size 570x610 pixels is given in Fig.3. The processed images after edge detection by Sobel operator and CHT are presented in Fig.4 and Fig.5, respectively.

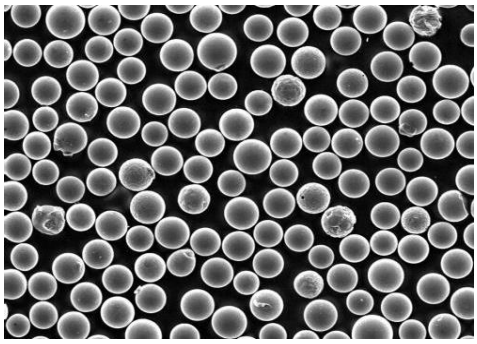


Fig. 3. Original microscopy image of embryos

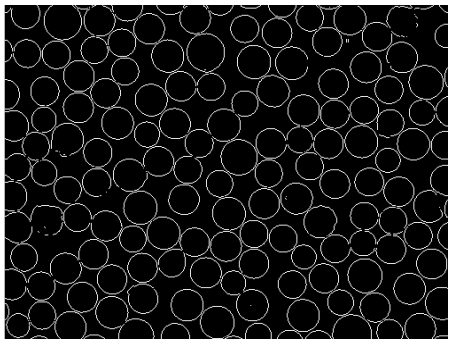


Fig. 4. Processed microscopy image of embryos after edge detection by Sobel operator

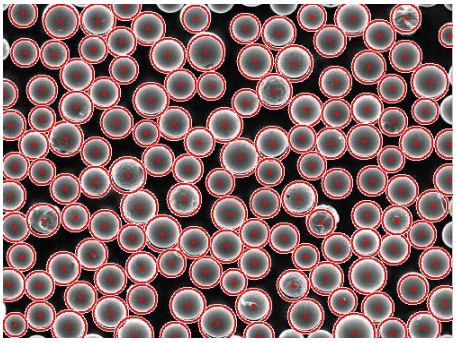


Fig. 5. Processed microscopy image after CHT with detected all embryos

Some results of elliptic objects detection are presented on the next figures below. The original microscopy image of erythrocytes of size 290x174 pixels is given in Fig.6. The processed images after edge detection by Sobel operator and RHT are presented in Fig.7 and Fig.8, respectively.

Fig. 6. Original microscopy image of erythrocytes

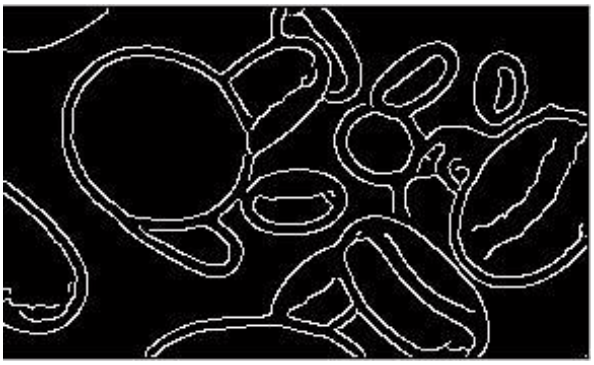


Fig. 7. Processed microscopy image of erythrocytes after edge detection by Sobel operator

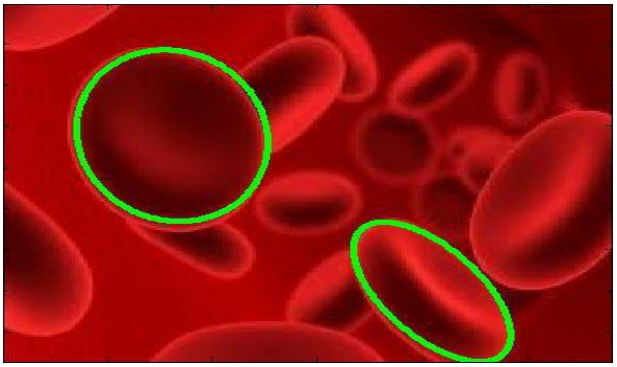


Fig. 8. Processed microscopy image after RHT with detected erythrocytes for selected parameters

Using the Circular Hough Transform, it is possible to find the coordinates of the centers of the embryos with circle form, as well as their radii.

By processing via Randomized Hough Transform the coordinates of the centroids and angles between the major and minor axis by elliptic objects are found.

The results, obtained by detection of circular and elliptic cells are evaluated by calculating the measures precision, recall and F-measure [10]. The measures were calculated for each image and, then, averaged over all images. The obtained results are given in Table 1.

TABLE I. EXPERIMENTAL RESULTS FOR DETECTED CELLS

	Performance Measures		
	Precision	***Recall***	***F-measure***
Value in [%]	97.67	95.23	96.43

We have made also some experiments with images of industrial objects with circular and elliptic shapes. The obtained simulations results, which are evaluated by calculating the measures precision, recall and F-measure, are similar to microscopy images [11, 12].

The processed images for detection of industrial objects with circular form using a monocular camera are presented in Fig.9, Fig.10 and Fig.11, respectively [12].

Fig. 9. The original image of an industrial object with circular shape

Fig. 10. Processed image of an industrial object after edge detection using Sobel operator

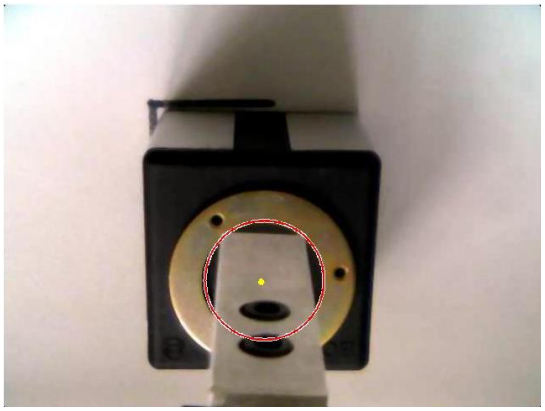


Fig. 11. The processed image with detected object

VI. CONCLUSION

In this paper software for automatic detection of objects with circular and elliptic forms and its graphical user interface (GUI) is presented. The circle objects are detected and located on the base of Circular Hough Transform (CHT). For elliptic shape detection the Randomized Hough Transform (RHT) is used. The experimental results have shown that the presented software can detect cells in the microscopy images accurately with average precision, recall and F-measure of 97.67%, 95.23% and 96.43%, respectively, which is satisfactory. Moreover, with the automatic detection, we can eliminate the subjectivity because, unlike in manual detection, is guaranteed that the same criteria are always used to detect cells. Because of the image quality, in the case of strong noise and very small size, some objects can be detected very difficult. For this reason a preprocessing of the images is needed. In addition, the program could be used in others applications. It can be used in computer vision systems for detection of different objects with circular and elliptic form. The presented GUI is suitable also to engineering education for studying of this processing.

REFERENCES

[1] P.V.C. Hough,"Method and means for recognizing complex patterns", *US Patent*, 3069654, 1962.

[2] H. Li, M.A. Lavin, R.J. Le Master, "Fast Hough transform: a hierarchical approach", *J. Computer. Vision Graphics Image Process.* 36, pp.139–161, 1986.

[3] A.S. Aguado, M.E. Montiel, M.S. Nixon, "On using directional information for parameter space decomposition in ellipse detection", *Pattern Recognition* vol.28 (3) pp.369–381, 1996.

[4] A. Sewisy,F. Leberl, "Detection ellipses by 2nding lines of symmetry in the images via an Hough transform applied to straight lines", *Image Vision Computing* vol.19 (12), pp. 857–866, 2001.

[5] Si-Cheng Zhang, Zhi-Qiang Liu, "A robust, real-time ellipse detector", *Pattern Recognition* vol.38, pp 273 – 287, 2005.

[6] S. Pei, J. Horng, "Circular detection based on Hough transform", *Pattern Recognition Letters*, vol.16, pp. 615-625, 1995.

[7] T. C. Chen and K. L. Chung, "An efficient randomized algorithm for detecting circles," *Computer Vision and Image Understanding*, pp. 172–191, 2001.

[8] L. Xu, E. Oja, and P. Kultanen, "A new curve detection method: Randomized Hough Transform," *Pattern Recognition Letters*, pp. 331–338, 1990.

[9] S. Inverso, "Ellipse Detection Using Randomized Hough Transform", www.saminverso.com/res/vision/EllipseDetectionOld.pdf, May 20, 2002.

[10] G.Faustino, M.Gattass, P. Carvalho, S. Rehen and C. de Lucena, "Automatic embryonic stem cells detection and counting in fluorescence microscopy images', *Monografias em Ciência da Computação*, ISSN 0103-9741, No. 04, 2009.

[11] V.Georgieva, "An Approach for Automatic Sorting of Tablets and Capsules", Proceeding of the 24th International Conference - ADP, Bulgaria, ISSN 1310 – 3946, pp.327-332, 2015.

[12] V.Georgieva, P. Petrov, "An Approach for Position Detection of Industrial Objects", Proceeding of International Scientific Conference on Information, Communication and Energy Systems and Technologies, ICEST'2013, ISBN: 978-619-167-182-3, pp.581-584, Ohrid, Macedonia, 2013.

Experimental realization of Controlled Square Root of Z Gate Using IBM's Cloud Quantum Experience Platform

Experimentelle Realisierung der kontrollierten Quadratwurzel des Z-Gatters mit der Cloud-Quantum-Erfahrungsplattform von IBM

Petar N. Nikolov*, Vassil T. Galabov†

*FDIBA, Technical University - Sofia
Sofia, Bulgaria, petar.nikolov@fdiba.tu-sofia.bg, vtg@tu-sofia.bg

***Abstract* —** Quantum computers form a technological cluster with huge growth in the last few years. Although this technology is of still very limited size – perhaps the reason it is not seen as a technology which may be mass produced or of public use in the near future– it is one of the most promising developments with a potential to change the world. The IBM Quantum Experience Platform makes it possible for every person around the world, without limitation as to geographical location, to get acquainted with the technology of quantum computing. It is a resource for both researchers and enthusiasts entering the quantum world. With the development of the platform, IBM has proven that the programming and writing code executable on a quantum computer can be easy and accessible (it is a cloud platform) even to people lacking any deep knowledge of quantum mechanics.

The construction of the Controlled Square Root of Z gate is achieved using only existing predefined gates in the Composer tool. This newly created gate could be used for further work on quantum algorithms and opens a new feasible way to write quantum code.

***Zusammenfassung* —** Quantencomputer bilden in den letzten Jahren einen technologischen Cluster mit einem enormen Wachstum. Obwohl diese Technologie immer noch sehr begrenzt ist - vielleicht der Grund, warum sie nicht als Technologie betrachtet wird, die in naher Zukunft massenhaft oder öffentlich genutzt werden kann - ist sie eine der vielversprechendsten Entwicklungen mit einem Potenzial, die Welt zu verändern. Die IBM Quantum Experience Platform ermöglicht es jedem Menschen auf der ganzen Welt, ohne Einschränkung der geographischen Lage, die Technologie des Quantencomputers kennenzulernen. Es ist eine Ressource für Forscher und Enthusiasten, die in die Quantenwelt eindringen. Mit der Entwicklung der Plattform hat IBM bewiesen, dass der Programmier- und Schreibcode, der auf einem Quantencomputer ausführbar ist, einfach und zugänglich sein kann (es ist eine Cloud-Plattform), auch für Menschen, die keine tiefe Kenntnis der Quantenmechanik haben.

Die Konstruktion der kontrollierten Quadratwurzel des Z-Gates wird unter Verwendung nur vorhandener vordefinierter Gatter im Composer-Werkzeug erreicht. Dieses neu geschaffene Gate könnte für weitere Arbeiten an Quantenalgorithmen verwendet werden und eröffnet eine neue Möglichkeit, Quantencode zu schreiben.

I. Introduction

A qubit is the equivalent of a bit in classical computing. It takes a value of 0 or 1 when measured [4]. Unmeasured, however, the qubit is in superposition $|\psi>$, which gives the mathematical representation of the qubit at any given time as a two-dimensional state space with orthonormal basis vectors $|0>$ and $|1>$:

$$|\psi> = a_0|0> + a_1|1> \quad (1)$$

The quantum probabilities a_0 and a_1 represent the chance that a given quantum state will be observed when the superposition has collapsed. These probabilities are complex numbers that satisfy the conditions:

$$| a_0^2 | + | a_1^2 | = 1 \quad (2)$$

$$|| \, |\psi > \, || = <\psi|\psi > = 1 \quad (3)$$

When the qubits are connected into strings, the result is a quantum register. The length of the string determines the amount of information this register can store. The superposition of a register $|\psi_n>$ means that each qubit in this register is in superposition.

$$|\psi_n> = \sum a_i|i> \quad (4)$$

So a register of n qubits is in superposition of all 2^n possible bit strings, that could be represented using n bits (4), where i is a bit string of 0s and 1s.

The operations over quantum registers are performed by Quantum logic gates [3]. Quantum logic gate applied to a quantum registers maps one quantum superposition to another, together allowing the evolution of the system to some desired final state.

The mathematical representation of quantum logic gates involves transformation matrices, or linear operators, applied to a quantum register by tensoring the transformation matrix with the matrix representation of the register. All linear operators

that correspond to quantum logic gates must be unitary (if a matrix U is unitary, U^{-1} is equal to the conjugate transpose $\hat{U}^{\top}$):

$$U^{-1} = \hat{U}^{T} \tag{5}$$

II. Quantum Controlled Square Root of Z Gate

Quantum computing uses controlled operations – multi-qubit operations that change the state of a qubit based on the values of other qubits [5].

Knowing that the Z gate has the following matrix representation:

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{6}$$

and given that Z = √Z√Z, a conclusion can be made that the √Z has the following matrix representation:

$$\sqrt{Z} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \tag{7}$$

where i is the imaginary unit ($i^2 = -1$).

The controlled-√Z gate is a gate that operates on two qubits in such a way that the first qubit serves as a control qubit for the second one. It has the matrix representation

$$C\sqrt{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \tag{8}$$

III. Experimental Results from the IBM's Quantum Experience Platform

IBM's Quantum Experience Platform is a 5-qubit quantum computer available on the cloud. In this platform writing a quantum code can be achieved by a tool called Compozer. In this tool we have a few limitations, one of which is the available gates we can use for computation. The gates we are going to use for the experimental realization of the controlled-√Z gate (Fig. 2, Fig. 3 and Fig. 5) are as follows:

- Identity Gate – It performs an idle operation on the qubit.

- Pauli-X Gate – It performs a π-rotation around the X axis on a qubit. (Bit flip)

- Controlled NOT Gate – It flips the target qubit if the control is in state 1.
- Phase gate that is square root of S gate, where S gate is √Z gate.

- Phase gate that is transposed conjugate of square root of S gate

- Measurement in the standard basis (Z)

Adding a control qubit to a √Z gate creates a simple quantum circuit. It applies a √Z gate to the target qubit if the control qubit is in state |1> . In general we follow the rules that:

HXH = Z
$SXS^{\top} = Y$
HH = I
$SS^{\top} = I$

Where H is Hadamard Gate, X is Pauli-X Gate, Z is Pauli-Z Gate, S is Phase Gate that is √Z gate, $S^{\top}$ is the conjugate transpose of S Gate, Y is Pauli-Y Gate.

Every quantum operation is a unitary matrix and every unitary matrix can be realized as a quantum operation [2]. Every unitary matrix has a square root, so every quantum operation also has a square root. It is possible to factor a controlled quantum operation into pieces [1], so we will be able to find matrices A,B and C and a phase vector $e^{i\Theta}$ such that [5]:

$$e^{i\theta}A \cdot X \cdot B \cdot X \cdot C = U \tag{9}$$

and

$$A.B.C = I \tag{10}$$

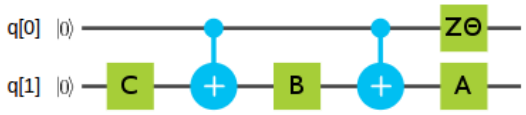


Fig. 1. Controlled U Gate

In Fig. 1 we see that the measured output will be the same as the input when the q[0] input state is |0> because of (10). And the ZΘ does not fire, because phase-shift gates act only on |1> state. In our case the appropriate values for the A, B, C and Θ are as follows [5, 6]:

- We can skip the C
- $B = \sqrt[4]{Z^{\top}}$
- $A = \sqrt[4]{Z}$
- $Z\Theta = \sqrt[4]{Z}$

In IBM's Quantum Experience Platform the √Z gate is represented by the S gate. The $\sqrt[4]{Z}$ is represented by the T gate and its conjugate transpose is represented by the $T^{\top}$ gate. The final result for the Controlled Square root of Z gate is shown in the Fig. 2.

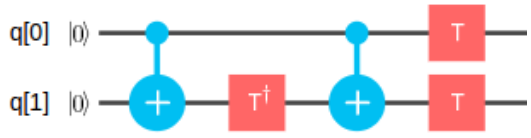


Fig. 2. Controlled Square Root of Z Gate

Because the constructed gate acts on two qubits, four different input combinations could be possible, depending on the qubit states before applying the gate. When the initial states of the two qubits are |11> a phase shift for the q[1] with π/2 radians is observed (Fig. 3 and Fig. 4).

The Quantum Sphere in Fig. 4 and Fig. 6 represents geometrically the pure state space of qubit in Hilbert space, where the north and the south pole correspond to the standard basis vectors |0> and |1>.

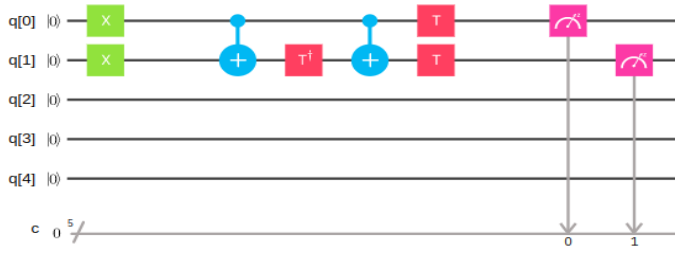


Fig. 3. The IBM Q – Controlled Square Root of Z Gate Circuit |11>

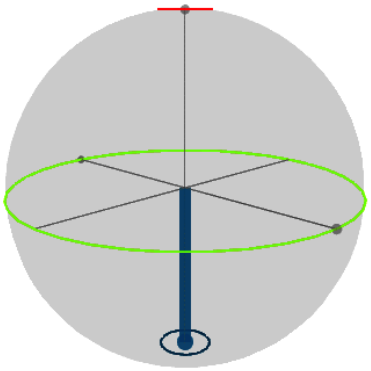


Fig. 4. The IBM Q – Quantum State: Quantum Sphere |11>

When the initial states of the two qubits are |01> we do not have phase shift for the q[1] because the control qubit does not fire (Fig. 5 and Fig. 6).

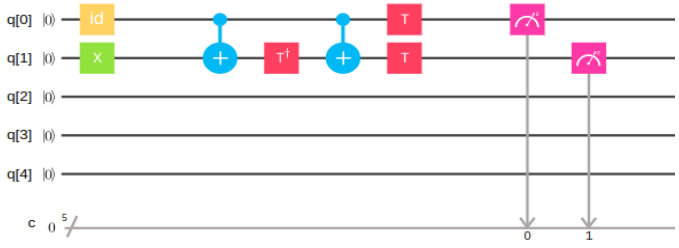


Fig. 5. The IBM Q – Controlled Square Root of Z Gate Circuit |01>

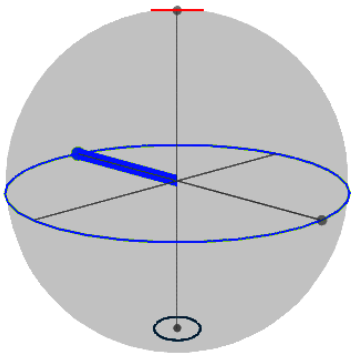


Fig. 6. The IBM Q – Quantum State: Quantum Sphere |01>

Conclusion

Our experiment for the realization of Controlled Square Root of Z Gate on IBM's Quantum Experience Platform was presented in this paper. Ideally, this work will provide a workable solution for the construction of other quantum gates on this platform in the future, which might contribute for the research and development of new quantum algorithms.

Acknowledgment

We acknowledge the use of IBM's Quantum Experience Platform for this work. All the views expressed are those of the authors and do not reflect the official position or policy of IBM. We are very grateful to the group for innovative bioinformatics research at Sofia Tech Park BioInfoTech Bioinformatics Lab for their support.

References

[1] A. Barenco, Ch. Bennett, R. Cleve, D. DiVichenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, H. Weinfurter "Elementary gates for quantum computation", *Physical Review A*, pp. 7-17, March 1995

[2] D. Deutsch, A. Barenco, A. Ekert, "Universality in Quantum Computation", *Proc. R. Soc. London A*, pp.3-6, June 1995

[3] D. Deutsch, "Quantum Computational Networks", *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, no. 1868, pp. 73-90, Sep. 1989

[4] S. Aaronson, Lecture Notes: "Lecture 9: Quantum." [Online] Available: http://www.scottaaronson.com/democritus/lec9.html , 2006

[5] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press UK, 2010

[6] C. Gidney, "Using Quantum Gates instead of Ancilla Bits", [Online}, http://algassert.com/circuits/2015/06/22/Using-Quantum-Gates-instead-of-Ancilla-Bits.html , 2015

This page is intentionally left blank

Graph partitioning – comparison of naive and evolutionary approach

Graphpartitionierung – Vergleich zwischen naiver und evolutionärer Ansatz

Vanesa Georgieva

MSc Student at FDIBA, Technical University Sofia
Sofia, Bulgaria, vanesa.georgieva@fdiba.tu-sofia.bg

***Abstract* —** In this paper is examined the problem for graph partitioning. An innovative solution of this problem is presented, which uses an evolutionary genetic approach combined with one of the first algorithms for graph partitioning – Fiduccia-Mattheyses. This solution is theoretically explained and practically implemented. A comparative analysis between the evolutionary and the naive approach is also described.

***Zusammenfassung* —** In diesem Papier wird das Problem für die Graphpartitionierung untersucht. Eine innovative Lösung dieses Problem wird vorgestellt, die einen evolütionären genetischen Ansatz mit einem der ersten Algorithmen für die Graphpartitionierung - Fiduccia-Mattheyses kombiniert. Diese Lösung wird theoretisch erklärt und praktisch umgesetzt. Eine Vergleichsanalyse zwischen dem evolutionären und dem naiven Ansatz wird ebenfalls beschrieben.

I. Introduction

A. Motivation

The graphs are fundamental data structure in computer science. When we want to represent our data as a graph and the graph become too large to be processed on one machine we should divide it into smaller pieces. Here comes the graph partitioning. Graph partitioning is used today for example in Facebook for friendship recommendation, where the graph vertices represent the users and the graph edges the personal relationships between them or in Google for PageRank computations, where websites are represented by graph vertices and the hyperlinks between the websites are represented by the graph edge.

Graphs that have billions of vertices can become too large and this brings a lot of issues. For example the computing of the desired partition on a single machine can be difficult due to memory or the processing and the finding of the best partition can take very long, because of the large amount of different partitions combinations. For solving such problems, we should determine such graph partitioning algorithm, which improves the run-time performance by partitioning, uses less memory and has the possibility to process a large amount of data. The algorithm should also be competitive with the naive methods like random graph partitioning.

In the interest of finding such an algorithm the main goal of this paper is to study the application of evolutionary approach for graph partitioning and to compare it performance with the naive approach for graph partitioning.

B. Related works

In 1970 Kerninghan and Lin were maybe the first who defined and investigated the graph partitioning problem [2]. They invented an algorithm that was improved from Fiduccia and Mattheyses in 1982 [3]. Today there are a lot of studies and different solutions of the graph partitioning problem. For example Shopov and Markova presented the applicability of some k-medoids algorithms for graph clustering in the Social Network Analysis [8]. A clustering is also partition of vertices, but there is no balance constraint and the number of partitions k is not given.

In the following part the graph concept and the graph partitioning problem will be briefly described and related works for graph partitioning technics will be overviewed. In part three the main methods and algorithms, which will be compared in this paper will be presented: the naive approach and the evolutionary approach using the Fiduccia-Mattheyses algorithm. In the last part some experiments will be described, which will are used to compare the evolutionary and the naive approaches for graph partitioning.

II. Graph partitioning problem

There are a lot of approaches solving the graph partitioning problem such as local search algorithms, obtaining partitions, multilevel approach, evolutionary algorithms etc. For the achievement of the main goal of this work were used algorithms from the area of the local search algorithms and the evolutionary algorithms and this is why they will be examined in this part.

According to [1] the main goal of the local search algorithms is to partition a graph by improving the objective function (the count of edges, moving between the partitions) by moving vertices between the partitions. The main goal of the evolutionary algorithms is to find a population of candidate solutions (individuals) and to evolve them toward better solutions.

A. Graph theory

The graph theory is the study of graphs and their properties. It is found from the Swiss mathematician Leonhard Paul Euler. [4]

A graph is a collection of vertices (points) that are connected by edges (lines). A simple graph G consists of a

nonempty finite set V(G) of elements called vertices (or nodes), and a finite set E(G) of distinct unordered pairs of distinct elements of V(G) called edges. We call V(G) the vertex set and E(G) the edge set of G. An edge {v, w} is said to join the vertices v and w, and is usually abbreviated to vw. For example simple graph [4] G whose vertex set V(G) is {w, v, w, z} has edge set E(G), consisting of the edges uv, uw, vw and wz. In any simple graph there is at most one edge joining a given pair of vertices. However, many results that hold for simple graphs can be extended to more general objects in which two vertices may have several edges joining them. In addition, we may remove the restriction that an edge joins two distinct vertices, and allow loops - edges joining a vertex to itself. The resulting object, in which loops and multiple edges are allowed, is called a general graph - or, simply, a graph. Thus every simple graph is a graph, but not every graph is a simple graph.

There are different types of graphs depending on the number of edges, number of vertices, interconnectivity, and their overall structure. In this work we will talk about connected weighted undirected graphs. This means that there should be a path between each pair vertices, to each edge should be assigned a positive number (weight) and the edges should have no direction, they should be bidirectional. (Fig.1)

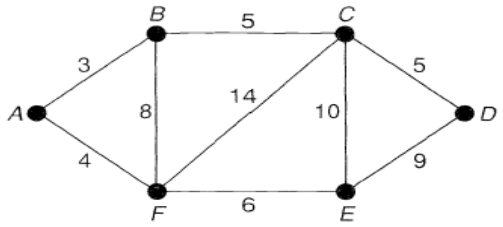


Fig. 1 Undirected weighted connected graph [4]

B. Graph partitioning

The data in the Graph partitioning problem (GPP) is presented as a graph G = (V, E). A set of n vertices is indicated with V and E indicates a set of n edges, so the graph can be partitioned into k partitions under specific criteria. The criteria is that each partition should be of about the same size and there should be a few connections between them. By fulfilling the criteria the capacity of the edges between the separate subgraphs should be minimized [5]. This problem is hard and its practical solutions are based on different approaches, called heuristics, which should solve, learn or discover problem, their practical methods are not optimal or perfect, but have immediate goals.

C. Kernighan-Lin algorithm

Probably the first, which investigated and defined the graph partitioning problem, Kernighan and Lin invented an algorithm [2], proposing simple heuristic found to produce globally optimal partition by moving vertices from one partition to another and to reduce its cost. The simplest version of this classic algorithm is to partition the graph into two partitions. They start with randomly chosen initial partition. Each partition should have about the same number of vertices. The vertices between these two partitions are exchanged by determining the cut size of each cut stage and the cut, which has maximum cut size, is saved. After all necessary exchanges are performed, the best cut, which is saved, is used as the output of the algorithm.

This idea was improved by Fidducia and Mattheyses in 1984 and is briefly described in the following subpart.

D. Fidducia-Mattheyses algorithm

Fiduccia and Mattheyses proposed improvements to Kernighan-Lin algorithm such that one pass of their version can be proved to run in linear time [3]. Like the Kernighan-Lin method, the Fiduccia-Mattheyses method performs passes in which each vertex is cut and moved at most once from one partition to another, and the best bisection observed during an iteration (if the corresponding reduction in the number of edges cut is positive) is used as input for the next iteration. However, instead of selecting pairs of vertices, the Fiduccia-Mattheyses method selects single vertex for movement.

E. Evolutionary algorithms

Evolutionary algorithms are stochastic search and optimization heuristics derived from the classic evolution theory, which are implemented on computers in the majority of cases [7]. The main goal of these algorithms is to select only the best individuals from the population, which cover specific criteria, and to reproduce only them. In this way the descendants will inherit properties only from the best individuals and the individuals, which not cover the criteria, will disappear. These population dynamics follow the basic rule of the Darvin's theory of evolution, which can be described in short as the "survival of the fittest".

III. Methods and algorithms

In this part will be explained the main two approaches for solving the graph partitioning problem, which will be compared in this paper: the naive approach and the evolutionary approach. We will begin with the naive approach.

A. Naive approach

The naive approach is a technique for estimation in which the last periods are used as current period's forecast. This method is usually used only for comparison with the forecasts, which are generated by better, developed to a high degree of complexity, techniques. In this paper it will be compared with evolutionary approach for GP.

1) Algorithm

The first part of the naive approach is dividing the vertices V of the graph G into all possible even k partitions with partition size q.

$$q = V/k \quad (q \in N) \tag{1}$$

The number of all possible partitions p of the graph is equal to the combination of all graph vertices n, which should be set into q positions, without repetitions. The order of the vertices in the partitions is not important.

$$p = C(n,q) = n!/q!(n-q)! \tag{2}$$

After the partitioning the graph into p partitions, we should remove the inappropriate ones. Inappropriate partition is partition, which contains disconnected subgraphs. We should check, if there is a path between each pair vertices in each partition and if there is no path, we should remove this partition.

After this check, if there is only one partition left that contains only connected vertices, we accept it as a final

partition of graph G and save the result as an output of the algorithm. Such a result can exist, if there is only one edge (v_n, v_m) that connects the subgraphs in the best partition. All other partitions will contain disconnected subgraphs.

If there is more than one partition left after the check for disconnected graph, we should find the best partition. The first requirement for best partition is fulfilled and the next requirement is the partition with the least connections between the subgraphs in it is to be found. For this purpose we need to calculate the gain g for each partition. The partition gain represents the count of vertices neighbors' nbc, which are distributed in the opposite partition. The smaller this count is, the less the connection between the subgraphs in the partitions are. The goal is to find the partition with the max gain:

$$g_{max} = nbc_{min} \quad (3)$$

When we find the partition with the highest gain, it is set as the best partition and will be saved as an end result and output of the partitioning of graph G.

```
Find the number p of all possible partitions of graph G such that p = n!/q!(n-q)!
For (i := 1 to p)
    Determine a balanced partition ci of the nodes into sets Ai and Bi with partition
    size = |V|/k = q
End For
For (i := to number c)
   Check connectivity of all vertices in subgraph Ai and subgraph Bi
   If (Ai OR Bi  contains disconnected vertices) remove ci
End For
Let gv is empty list
For (i := 1 to size of c)
   Compute gain value g of ci
   Add g to gv[i]
End For
Best partition = the partition with max gain from gv
```

Fig. 2 Naive approach- Pseudo code

B. *Evolutionary genetic approach based on the Fiduccia-Mattheyses algorithm*

The main goal of this work is to present and compare with the naive approach, a solution of the GPP using evolutionary genetic approach, which processed large amount of data and improve the run-time performance and memory usage. In this subpart will be presented such a solution: the Fidducia-Mattheyses algorithm used together with evolutionary genetic approach.

When we use an evolutionary approach, we will need some terms, coming from the biology, which will be briefly described.

1) Evolutionary representations and operators

As described in [6] population is the number of all of the individuals of the same groups, which exist in a particular area, and have the capability of interbreeding. It is presented by a set of individuals, each of them represented graph *k*-partition. Each of these individuals will be created by random distribution of the graph vertices, which will create the chromosome of the individuals. The best individual from the population will be saved by each new generation. In general generation means aggregate of the functions and phenomena which attend reproduction.

As in the nature, as so in the evolutionary algorithm, there are constantly new populations and generations. That is why we need to set a start parameter for the maximum generation, which we want to reach. For example we can set the value of the maximum generation equals 100, so we want to find the best individual in the period of 100 epochs. When we choose the maximum generation that should be reached, we should find how big our partition can be. The dividing of the vertices V of the graph G into even k partitions should have partition size pSize (Eq.1)

2) Algorithm

To get the initial population, which we want to improve, we will use the algorithm of Fiduccia-Mattheyses. We start with adding random individuals to the first population, created by combinations from all graph vertices. The population is an array of individuals, containing k-partitions with pSize. Than for each individual in this population we want to find the best partition. The best partition for each individual is found in the following way. Firstly we execute the Fiduccia-Mattheyses algorithm for the current individual and calculate its fitness value. The fitness value is the cut size - the number of edges, which are removed from the graph, so it can be divided into k parts. The cut size should be as less as possible. If the fitness value of the individual, derived from the Fiduccia-Mattheyses (FM) is bigger than the fitness value of the current individual, we replaced the genes in the chromosome of the current individual with these from the chromosome of the individual derived from FM. We iterate the FM for each individual until no fitness improvement is found.

Now when we have initialized the initial population, we can start with its evolution. We will repeat the cycle of evolution until maximum generation is reached. Each cycle of evolution contains some steps, which will be described in the following paragraphs.

The first step is the step of the random selection. Here we select two random individuals from the population and called them parents. From these two parents we want to generate an offspring that can evolve. The generation of the offspring happens with the usage of the uniform crossover between the selected parents, which is the next step in the evolution.

During the step of the uniform crossover we want to create an individual with mixing ratio between the given parents, which enables the parent chromosomes to contribute the gene level of the offspring. First we compared the two parents and if they had similar genes, they are added to the new genes of

```
Initialize population with populationSize = vertices count/k
   For (i := to populationSize)
      Create individual
      Create FM individual as iterate Fiduccia Mattheyses algorithm until the best
      fitness value for the current individual is found
      While (fitness value of the current individual < fitness value of FM individual)
          Iterate Fiducciia Mattheyses algorithm until the best fitness value for the
          current individual is found
       End While
  End For
Start creating generations
Generation = 0
While (Generation < MaxGenerations)
  Select random parents from the current populatin
  Create FM individual as iterate Fiduccia Mattheyses algorithm until the best fitness
  value for the offspring is found
  While (fitness value of the offspring < fitness value of FM individual)
      Iterate Fiduccia Mattheyses algorithm until the best fitness value for the
      offspring is found
   End While
  Replace the worst individual from the old population with the offspring
  Increase Generation with 1
 End While
 Best partition = the last individual from the population
```

Fig. 3 Evolutionary approach- Pseudo code

the offspring. The mismatched genes are saved as their order is permutated. The first half of the mismatched genes goes to the first partition and the other half go to the other. In the end we get the new offspring. After that we execute the Fiduccia-

Mattheyses algorithm to the new offspring to find its best partition. In the end of the iteration we replaced the worst individual from the population (with the lowest fitness value) with the offspring. Than we receive the next generation.

We repeat this steps until the maximum generation is reached. After that we receive the best individual, derived from the population, as the optimal graph partition of graph G into k partitions.

IV. Experiments and Results

The evolutionary and the naive approaches for solving the graph partitioning problem was also practically implemented and the programming language that was used is Java. The implementation of the algorithms is used for comparison between them, because this is the main goal of this paper.

As an input value for the algorithms were used undirected, weighted, connected graphs with different number of vertices and edges between them. The conducted experiments are repeated several times for each approach and the differences between the execution time are insignificant. The execution time in the experiments will be presented in seconds and the memory usage in megabytes. The experiments will measure the coverage and the efficiency of the two approaches with the goal to prove, that the evolutionary approach can work with significantly larger graphs than the naive approach, works faster than it and uses less memory.

1) Experiment 1: The first experiment was done with relatively small graph that contains 17 vertices and 20 edges. This experiment has goal to show that there is a significant difference in the execution time and memory usage of the two approaches. The results are shown in the following graphic:

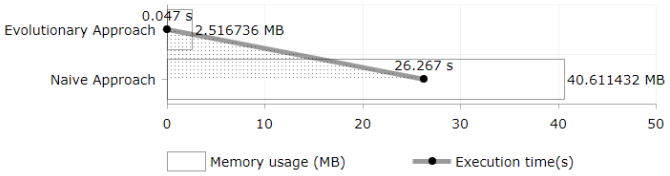


Fig. 4 Graph partitioning of graph with small number of vertices

As result of this experiment the both approaches output right partitions- there are no disconnected vertex in each partition. As we can see in the Fig.4 the execution time of the evolutionary approach takes an advantage over the naive approach with 28.22 seconds, which is 99% faster. The difference in memory usage is 38.09 megabytes-the evolutionary approach uses ~94% less memory than the naive approach. With this experiment we proved, that the evolutionary approach is more efficient than the naive approach: it is faster and uses less memory.

2) Experiment 2: The second experiment was done with a graph with large number of vertices and edges: 102 vertices and 101 edges and represents central cities in Russia. The goal of this experiment is to prove that the evolutionary approach can work with significantly larger graphs, but the naive approach cannot. The results are shown in the following graphic:

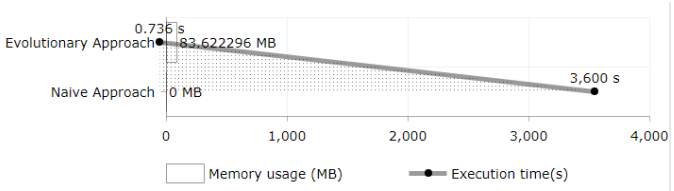


Fig. 5 Graph partitioning of graph with large number of vertices

In Fig.5 we can see that the evolutionary approach partition graph with large number of vertices for 0.736 seconds and uses 83 megabytes memory. It also outputs the right partitions- there are no disconnected vertex in each partition. The naive approach gives no results in this experiment. It was running for more than an hour (3600s) and there were no results. The reason is that there are a large number of combinations from graph vertices, which should be processed, so the naive approach can give appropriate results, and this will take a lot of time.

The experiments, which are described in this part, proved that the goal of this work is achieved. It was proved, that the evolutionary approach can work with significantly larger graphs than the naive approach, works faster than it and uses less memory.

V. Conclusion

The evolutionary approach can be easily adjusted to the graph partitioning problem. It has the advantages to change and customize any aspect of any given algorithm, which makes it more flexible and the progress is measured more easily. In this paper was presented a solution of the GPP using an evolutionary genetic approach combined with one of the first algorithms for GP-Fiduccia-Mattheyses. Different experiments were done, so it was proved that this approach is more efficient and has more coverage than the naive approach

Acknowledgements

I express my gratitude to Dr. Ventseslav Shopov and Dr. Vanya Markova for many valuable comments and help on a preliminary draft of this paper.

References

[1] C. Schulz, "*High quality graph partitioning",* Karlsruher Instituts für Technologie, 2013, pp 25-31.

[2] B. Kernighan and S. Lin. "*An efficient heuristic procedure for partitioning graphs*", The Bell System Tech J., The Bell System Technical Journal, February 1970.

[3] Fiduccia, C. M, Mattheyses. R. M., "*A linear-time heuristic for improving network partitions.*" 19th Design Automation Conference, 14-16 June 1982.

[4] R. Wilson., *Introduction to graph theory*, Fourth edition, Addison Wesley Longman Limited, 1996, pp 8-25

[5] A. Buluç , H. Meyerhenke, I. Safro, P. Sanders, C. Schulz, "*Recent advances in graph partitioning*", Lawrence Berkeley National Laboratory, November 2016, pp 2-3.

[6] X. Yu, M. Gen, *Introduction to evolutionary algorithms*, Springer London Dordrecht Heidelberg New York, 2010, pp 8-26

[7] F. Streichert, "*Introduction to evolutionary algorithms*", University of Tuebingen, April 2002

[8] V. Markova, V. Shopov, "*Graph partitioning methods in social network analysis*", Bulgarian Academy of Sciences, Proceedings of the International Conference on Information Technologies (InfoTech-2016), 20-21 September 2016, Bulgaria

Spatial random graphs generation — comparison of naive and evolutionary approach

Geometrische Zufallsgraphen — Vergleich zwischen naiver und evolutionärer Ansatz

Blagovest Achanov

MSc Student at Faculty of German Engineering and Industrial Management
Sofia, Bulgaria, blagovest.achanov@fdiba.tu-sofia.bg

***Abstract* —** This work introduces an evolutionary approach for generation of spatial random graphs based on Waxman graph and a naive approach for generation of undirected connected weighted graphs in given 2D space. Both approaches are experimentally compared by measuring their execution time, graph connectivity and maximal average length of edges, proving the practical usability of evolutionary algorithms combined with Waxman graphs for fast generation of connected undirected geometric graphs.

***Zusammenfassung* —** Diese Arbeit stellt ein naiver und ein evolutionärer Ansatz für Generierung von gewichtete ungerichtete verbundene Zufallsgraphen. Beide Ansätze werden mit verschiedenen Versuche gegen Ausführungszeit der Ecken und maximale Durschnittdistanz der Ecken. Die Versuche werden die Gelegenheit für praktische Anwendung der evolutionären Algorithmen für schnelle Generierung von geometrischen Graphen.

I. Introduction

Different situations such as the Seven Bridges of Königsberg problem [2], a network topology, an electric circuit or relationships in a social network can be displayed as graphs. When implementing such a system, it should be tested with different graphs that simulate some special conditions. Those graphs, which are generated using probability approaches are called random graphs.

The first studies in the area of random graph generation (RGG) are done in the late 1950s by Erdos, Renyi and Gilbert [3]. Both Erdos-Reniy and Gilbert models use an iterative way to generate graphs given number of nodes and don't support spatial graphs. Later, Waxman presented a model for generation of spatial random graphs. Tettamanzi [7], Cordella, De Stefano, Fontanella, and Marcelli, [5] worked on evolutionary algorithms applied to graphs. Shopov and Markova compared different approaches for RGG in [8] and implemented an evolutionary approach in [9].

The purpose of this work is to update the traditional Waxman graph to use an evolutionary approach and to compare it to a non-evolutionary approach, where all graphs meeting some requirements are generated (naive approach). Both approaches are going to be described in theory in part III and IV and practically compared in part V of this work. As result, the work should prove the practical usability of an evolutionary approach combined with Waxman graph.

II. Related works

Some of the publications, nominated in the introduction were combined and elaborated in this work. This part of the work introduces them in some detail.

A. Waxman graph

Waxman graphs are particular class of random graphs and are used for modelling network topologies for evaluating routing algorithms (see [1]). Vertices of the graph are randomly distributed over a rectangular area. Edges between nodes are added randomly using a probability mechanism, where the probability that two vertices are connected decreases exponentially depending on the Euclidean distance between the nodes. The number of vertices N is pre-determined [1]. The edge existence probability can be described on the following way:

$$P(\{u, v\}) = \beta e^{\frac{-d(u,v)}{L\alpha}} \quad (1)$$

U and v are vertices in the graph, L is the maximum Euclidean distance between any two nodes (see [9]), α and β are parameters in range (0, 1]. According to [6] larger values of α increase the densities of short edges and larger β values increase the density of longer edges. Waxman sets the values of α and β = 0.4 [1]. For the purpose of this work, values of α and β will be preset. Also, the maximum distance L will be calculated as the maximal possible distance between the nodes in the graph.

B. Evolution in graphs

EvoGeneSys [5] introduced in multilist structure and implemented evolutionary operators to work with it. They demonstrated the practical applicability of these methods on undirected unweighted graphs. The algorithm described there was not applied on weighted graphs with coordinates in 2D space. It was also not compared to a brute force approach.

This work uses as basis Waxman graph and updates the algorithms from [5] to work with spatial graphs. The usability of the evolutionary algorithm will be proved by comparing it to a naive approach, which will be introduced in the next part.

III. Naive Approach for RGG

The naive approach for graph generation, elaborated in this work, returns all possible graphs that meet specific requirements such as number of nodes, edges and width and height of a two dimensional space that contains the graph. All returned graphs are stored in a list. The algorithm can be split in 4 parts, illustrated in the following sections, where for each part is given a formula, giving all possible combinations:

A. Nodes generation

This part ensures, that a node is generated for each possible point in a rectangular 2D shape with integer values of coordinates. The total number of nodes for such a shape is:

$$vertexCount = (x_{max}-x_{min})(y_{max}-y_{min}) \quad (2)$$

B. Vertex selection

The vertex selection is used to combine different nodes in a graph. After generating all nodes (see previous paragraph), those vertices should be grouped in different graphs. For this reason, a combination is used- the total number of possibilities for nodes in graph is:

$$C(vertexCount, nodesInGraph) \quad (3)$$

C. Edges generation

The purpose of edges generation step is to generate all possible edges from set of nodes. According to its definition, each edge connects a pair of nodes [4]. So, in this case the total number of possible edges for each combination of nodes will be:

$$allEdges = C(nodesInGraph, 2) \quad (4)$$

For execution of the naive approach, a number of edges in graph is passed. To ensure the graph can be connected, this number must be at least nodesInGraph – 1 for an undirected graph. Then, the total number of edges combinations for a set of nodes will be:

$$C(allEdges, numberOfEdges) \quad (5)$$

D. Graphs generation

This step combines node sets and edge sets, in order to generate graphs. The following equation gives all possible graphs given shape coordinates, number of nodes and number of edges:

$$graphsCount = C(C((x_{max}-x_{min})(y_{max}-y_{min}), nodesInGraph), numberOfEdges)*C((x_{max}-x_{min})(y_{max}-y_{min}), nodesInGraph) \quad (6)$$

The set of graphs is generated following the pseudocode:

```
naiveGenerator( numberOfNodes, numberOfEdges,
rectangleCoordinates){
    Graph[] generatedGraphs;
    //generate all possible nodes in rectangle
    Nodes[] = generateNodes (rectangleCoordinates);
    //generate all combinations of nodes
    nodesForGraph[] = generateNodesCombinations(numberOfNodes,
    nodes);
    for(currentNodes : nodesInGraph){
        //generate all possible edges between nodes
        edges[] = generateEdges(currentNodes);
        //generate combinations of edges in graph
        edgesCombinations[]= generateEdgesCombinations(edges,
        numberOfEdges);
        for(currentEdges: edgesCombinations){
          Graph currentGraph = new Graph (currentNodes,
          currentEdges);
          generatedGraphs.add(currentGraph);
        }
    }
    return generatedGraphs;
}
```

Fig. 1 Naive approach- pseudocode

After all having all possible graphs, they can be evaluated given criteria and be reordered using it. For example, they can be filtered to a subset of connected graphs. However, the main generation approach is independent from this evaluation.

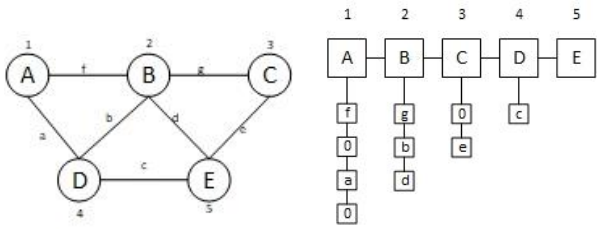
The naive approach has the complexity of ***O(C(C(shape,nodes), edges)* C(shape, nodes)***, where shape describes the number of all possible points with integer coordinates in the given 2D space. This complexity in O notation shows, that the speed of the algorithm will be significantly reduced by changing the total number of nodes in graph, all possible points or even the edges count. This assertion will be verified in the practical part of the work.

IV. Evolutionary approach for RGG

Another approach for RGG is to use the evolutionary operators in order to create new graphs. For this purpose, each graph should be represented in an appropriate structure, allowing an easy mutation and crossover. Such a structure was developed the in [5] multilist. For the purpose of this work, the multilist has to be updated to support evolutionary operations with spatial graphs respectively Waxman graph.

A. Multilist

Multilist (or ML) is a structure, used for representing graphs is as set of (N) lists, where the *main list* contains the nodes and the other lists (or *sublists*) contain the edges of a graph (see [5]). Each list of edges (or sublist) is connected to a node and contains the edges from this node. The size of each sublist depends on the position of the node in the main list and contains edges to the upcoming nodes from the main list following their order. As an example, the edge list from of the first node will contain N-1 edges. If an edge does not exist, it is represented as null. The edge list of the second node contains the connections to all upcoming nodes- or N-2 edges. The last node from the main list doesn't contain a sublist. Each edge is stored only in one list. The described multilist has a triangular structure and an example is given in the following figure:



(a) (b)

Fig. 2 (a) graph (b) multilist representation

The given graph in Fig 1. (a) is converted to a multilist in (b). Particulary, if we swap the indexes of vertex A and vertex B, a different multilist will be generated.

B. Crossover

Crossover using multilist structure according to [5] is done in two steps: each multilist is split in two sublists, sublists are swapped between and merged. As result 2 new individuals (offspring) is generated.

1) Split

To split a multilist, a split point in interval ($0 < S < N-1$) must be selected, where **N** is the number of vertices and **S** is the split point. The next figure gives an example for split operation of the graph in fig. 1 with split point 2:

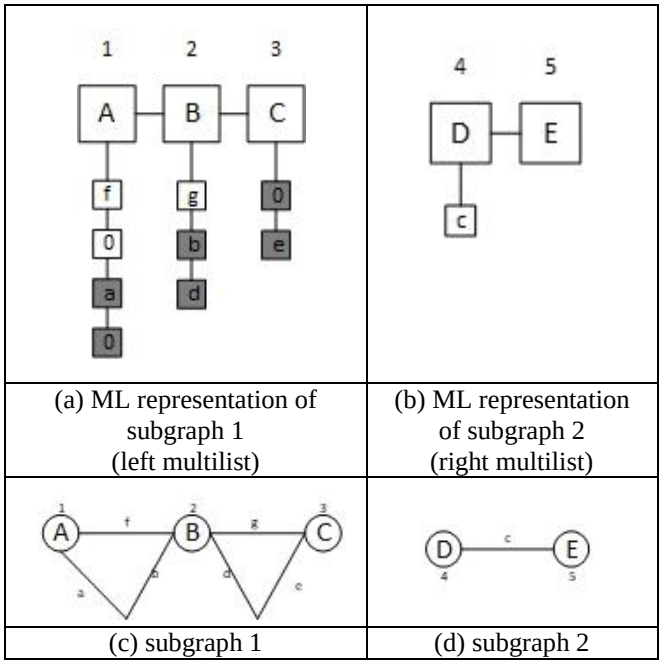


Fig. 3 Split operation

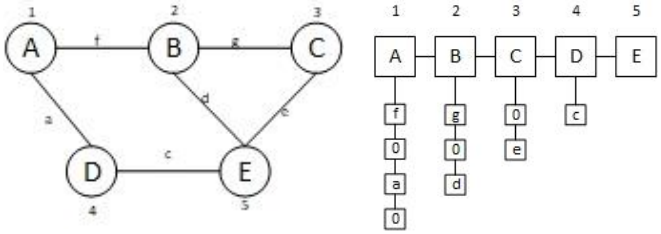
After splitting the multilist in two, some of the edges of the left multilist (colored in fig 2) will get invalid. They are used during merge operation to generate edges between the two sublists.

2) Merge

The purpose of merge operation is to generate a new individual by merging 2 multilists, describing 2 subgraphs. The merge is done naively by merging the 2 main lists. The left sublist contains edges to the right, so each edge has to be updated to lead to the new vertex and the distance should be recalculated. Important for merge operation is that the total number of vertices before splitting and after merging the graph does not change. This means that to the left multilist must be merged a right multilist with same number of nodes that was previously cut from the graph to produce the left multilist.

C. Mutation

Mutation is the second evolutionary operator. It appears in a random change in the individual (see [5]). Given a mutation rate, a mutation should appear not in all individuals. In this work, a mutation is done by randomly adding or removing and edge. Using the multilist, a random node and random edge from its sublist is selected. If the edge is null, a new edge is added. Otherwise the edge is removed.



(a) graph (b) multilist

Fig. 4 Mutation

The previous figure gives an example of mutation in the example graph from fig. 1. Due to mutation, the edge **b** between vertices **B** and **D** is removed. This also affects the multilist in (b).

D. Selection and fitness function

Selection operation is used to select a subgroup of individuals from the offspring that meet some conditions better than others. In the current case, those conditions are graph connectivity and maximal average edge length. The number of selected individuals is passed as parameter. The selected individuals are used as parents for the next epoch of the evolutionary algorithm (see Fig. 5).

```
evolGenerator(initNumGraphs, nodes, coords, alpha, beta, epochs,
offspringNum, crossoverPoint){
    Multilist[offspringNum] resultMultilist;
    Graph[] initPopulation = generateInitPopulation(nodes, coords,
    populationCount, alpha, beta);
    Multilist[] initMultilist = convertToMultilist(initPopulation);
    for( 0 .. epcohs){
        epochResult = doCrossover(resultMultilist, crossoverPoint);
        resultMultilist =doSelection(epochResult, offspringNum);
    }
    Graph[] resultGraphs = convertToGraph(resultMultilist);
    Return  resultGraphs;
}
```

Fig. 5 Evolutionary approach- pseudocode

V. Experiments and Results

In the practical part of this work both naive and evolutionary approach for generation of spatial random graphs were implemented in and a comparative analysis of them was done. In order to remove any language-dependent differences in performance both algorithms were implemented in Java. Multilists were represented using hash maps where the start vertex is the key and the value is a list of all edges from this node. This structure was selected because it uses key-value representation. In order to represent a graph, node, edge and graph classes were defined. The project was compiled and all experiments were executed on the same machine with 8 GB RAM and Intel i7-2640M CPU.

A. Experiment conditions and measures

In order to compare both algorithms, different experiment conditions are used. Dimensions of the rectangle are changed in order to test both approaches. Timer is implemented to measure the execution time of each algorithm. Maximal average edge distance is the other measured metric. Number of edges parameter is used only for the naive approach. In order to decrease its execution time, for number of edges is chosen to be equal to the number of vertices -1. This is used of the condition that the generated graphs are connected. In theory, the evolutionary approach can generate a connected graph with number of edges greater the number of nodes -1, but in most of the cases the generated graph contains exactly that number edges. This can be controlled by changing alpha and beta coefficients, but for the purpose of this work values are fixed. As described in [1], the values of α and β are set to be 0.4, the mutation rate is set to 0.15 and the offspring count is 20. Two values for count of epochs were chosen- 10 and 20. They're used to allow measure the time and result changes caused by the change of epochs. The following table shows the experiment conditions:

Experiment number	**# of nodes**	**# of edges**	**space**
1	5	4	3x3
2	5	4	4x4
3	5	4	4x5
4	5	4	10x10
5	5	4	20x20

Table 1 Experiment conditions- space and graphs

Experiments 1, 2 and 3 are executed for naive and evolutionary approaches. The last 2 experiments were done only for evolutionary approach, because the execution of them with a naive approach did not finish after 10 minutes. The results and conclusions from them are given in the following paragraph.

B. *Results*

The following diagram shows the execution time, measured for each of the experiments. Each experiment was executed five times. After that, the average time was calculated.

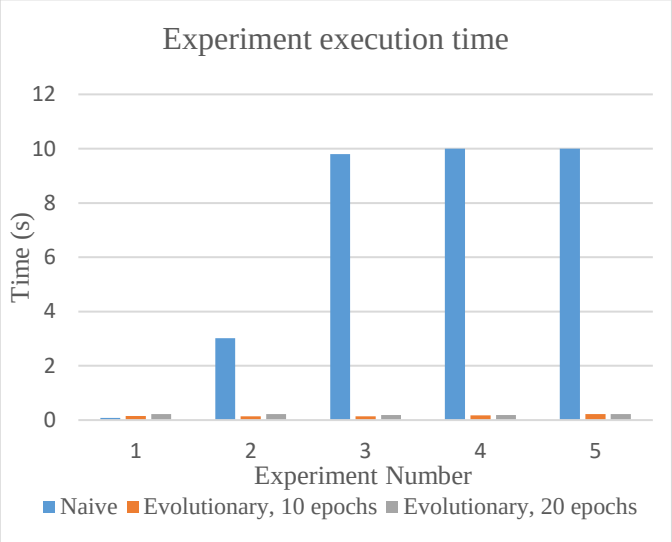


Fig. 6 Execution time

The measured time shows the rapid difference in the execution time of the naive approach in different cases. This behavior is expected and follows the description given in point III. The evolutionary approach works faster in 4 of 5 test cases. The increasing of space dimensions slows down the execution of the evolutionary method too, it's still much more appropriate that the naive approach. For test 4 and 5, the naive execution did not finish in 10 minutes and was not measured. The factor which increases the execution time of the evolutionary approach is the number of epochs. The execution of 20 epochs is almost always slower than the execution of 10 epochs. As a conclusion, the evolutionary approach is relevant for smaller and bigger spaces. The naive is not useable for dimensions over 4x5.

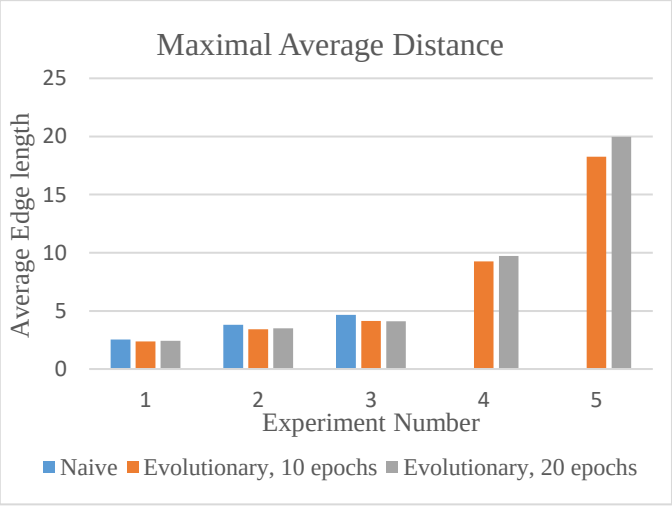


Fig. 7 Maximal average distance

Fig. 7 shows the metrics measured for maximal average distance on an edge in graph. The naive measures are control ones- they show the maximal possible distance. Compared to them, the results from the evolution are smaller. However, the difference in the results isn't very big. More, the increased number of epochs increased the result in more of the cases. This behavior is expected- the purpose of the evolutionary approach is not to give the best solution, but to give a result, which is optimal for some requirements. Theoretically the evolutionary approach can produce same result as the naive approach. The missing results for naive approach for experiments 4 and 5 are caused by the long execution time which was previously explained.

Analyzing both metrics, the evolutionary approach for spatial random graphs generation is applicable for graphs in small and bigger spaces. It produces relevant results for shorter execution time compared to naive approach which finds the best graph answering the distance condition, but it's not applicable for bigger dimensions or bigger graphs. Thus, only the evolutionary approach is practically applicable.

VI. Conclusion

The aim of this work was to create an evolutionary approach for generation of spatial random graphs and prove its usability in different conditions. For this, the Waxman graph was used. It was combined with an evolutionary algorithm using multilists following the approach described in [5]. During the work, a naive approach, generating all possible graphs given size and space was created. Both approaches were implemented and compared. The results showed the usability of the introduced evolutionary algorithms for generation of larger geometric random graphs. Analyzing the values of alpha, beta and mutation coefficients can improve the results of the evolution, which can be done in another work, related to this.

Acknowledgment

I would like to express my gratitude to Dr. Vanya Markova and Dr. Vetseslav Shopov for the support and advices during the research and the development of this work.

References

[1] B.M. Waxman, "Routing of multipoint connections", *IEEE Journal on Selected Areas in Communications*, 1988, pp 1617-1622

[2] See Shields, Rob, "Cultural Topology: The Seven Bridges of Königsburg 1736", *Theory Culture and Society*, 2012, pp 29.

[3] Roughan, Matthew, Tuke, Jonathan and Parsonage, Eric, "Estimating the Parameters of the Waxman Random Graph", *eprint arXiv:1506.07974*, 2015

[4] Robin J. Wilson, *Introduction to graph theory*, *4th edition*, Longman, 1996

[5] L. Cordella, C. De Stefano, F. Fontanella, A. Marcelli, "EvoGeneSys, a New Evolutionary Approach to Graph Generation", *Applied Soft Computing 13 (4)*, 2013

[6] Naldi, Maurizio, "Connectivity of Waxman topology models", *Journal of Computer communications 29.1* ,2005, pp 24-31.

[7] Tettamanzi A.G.B. "Drawing Graphs with Evolutionary Algorithms", *Adaptive Computing in Design and Manufacture*. Springer, 1998, London

[8] Ventseslav Shopov, Vanya Markova, "COMPARISON OF RANDOM GRAPH GENERATORS", *Proceedings of the International Conference on Information Technologies (InfoTech-2016)*, 2016

[9] Ventseslav Shopov, Vanya Markova, "EVOLUTIONARY APPROACH FOR SOLVING DYNAMIC GRAPH PROBLEMS", *Proceedings of the International Conference on Information Technologies (InfoTech-2016)*, 2016

Knowledge discovery for business process optimization at Enterprise 2.0

Knowledge discovery für Geschftsprozessoptimierung in Unternehmen 2.0 (Enterprise 2.0)

Konstantina Plamenova Dimitrova
Master Student at Technical University - Sofia, FDIBA
e-mail: konstantina.dimitrova@fdiba.tu-sofia.bg

Abstract — The goal of this paper is to explore the possibilities of deploying Knowledge Discovery into the Enterprise 2.0 for business process optimization and the corresponding benefits. For the purpose, an exemplary theoretical implementation of Knowledge Discovery Process using Decision Trees and Excel Macros is presented.

Zusammenfassung — Das Ziel dieses wissenschaftlichen Artikel ist Erforshung der Implementierung von Knowledge Discovery in dem Unternehmen 2.0 zum Zweck Geschäftsprozessoptimierung und davon resultierenden Gewinn. Zum Veranschaulichen ist eine theoretische Lösung von Knowledge Discovery Implementierung mittels Entscheidungsbaum (Decision Tree) und Excel Macros dargestellt.

I. Prerequisites and current environment

In the era of the Zetta-byte the trend of exponentially growing data is no longer a myth. In the past few years the business investing even more in knowledge and data management in order to process the already gathered data and preparing for the vast amounts of data that are yet to come [1, 2]. But efficient data storage methods are not enough. Support-oriented services, for example, independently from their focus, accumulate an enormous amount of unstructured data like call and chat logs, tickets, emails, photos that is significantly more challenging to process than the structured. Accumulating data without processing it and putting the gathered knowledge into practice can result in losses for the Enterprise 2.0 that vary in form.

The development of the Web and the Internet, two fundamental conditions, have not only lead to exponential data growth but also presented the opportunity for inventions, like tools for improving our working routines and techniques. The term Web 2.0, which was coined by OReilly [2], represents elaborated combination of technologies like AJAX, SaaS, tagging, blogging as an example of collective knowledge and the collaborative web, the social networks. They are used on daily bases in every aspect of a work process for communicating information in the professional sphere and outside, managing the information flow from every entry point via cloud-based enterprise platforms, harvesting collaborative knowledge from the employees and easing the information and data flow within the organization. This is what defines an Enterprise 2.0 an organization that takes advantage of the Web 2.0 technologies and uses them to fuel up the companys growth and improvement in the constantly changing and challenging business environment [3, 4]. Although the Web 2.0 technologies that are used in the Enterprise 2.0 like Yammer or other forms of blogs for collaborative knowledge exchange, facilitate the working processes in the organization, they also produce additional valuable data as a byproduct that needs further processing. For example, a logged ticket about a service issue together with the accompanying email or call log are stored in databases, but the communication between the employees for providing a resolution is saved as well. Thus, one single event can result in the induction of newly generated data of unstructured type and of a great importance to the Enterprise 2.0.

II. Knowledge Discovery

Processing the stored data is challenging task for every Enterprise 2.0. The difficulty of accomplishing this task is even higher when the data to be processed is unstructured. A support-oriented service with two communication channels for customer direct calls and emails, for example, can handle hundreds even thousands of tickets per day. What is more important than the current issue described in the ticker is whether it is a one-time event or a part of an ongrowing trend. This could be determined using Knowledge Discovery techniques.

The aim of Knowledge Discovery is to reveal new, novel and useful information hidden in the unstructured data and not apparent upon first sight. This information can be discovered by deploying Knowledge Discovery techniques into the best practices of the Enterprise 2.0. the so-called Knowledge Discovery Process (KDP) can be implemented into the business processes and by its definition is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [5]. Although the process emphasizes on databases as the main

source of data, it also can be applied to the nondatabase source of data. The process consists of many steps, one of which is Data Mining, and each attempts to complete a particular discovery task. In order to put the KDP into practice, a Knowledge Discovery Model is realized. There are three types of KDPM academic, industrial and hybrid. The latter represents a compilation of the first two it provides more general, research-oriented description of the steps like in the academic model, while at the same time focuses on the understanding of objectives and requirements from the business perspective, like the industrial model. The models have iterative nature and introduce the Data Mining step instead of the modelling one. The six-step hybrid model by Cios et al. [6] is based on the CRISP-DM model and its steps are:

1. Understanding of the problem domain it involves working closely with the domain experts in order to gain better insight of the problem and thus easily defining it and determining the project goals, identifying the key people and learning about current solutions to this problem;

2. Understanding of the data firstly, a collection of sample data is gathered, checked for completeness, redundancy, missing values, the plausibility of attribute values, etc.;

3. Preparation of the data it is defined by sampling, running correlation and significance tests, data cleaning;

4. Data Mining data mining methods are used by the data miners to derive patterns and thus knowledge from the processed data;

5. Evaluation of the discovered knowledge in this step the goal is the understanding the results, checking whether the discovered knowledge is novel and interesting, interpretation of the results by domain experts;

6. Use of the discovered knowledge.

The KDPM puts emphasis on the iterative aspect of the process drawing from the experience of users of previous models. Several explicit feedback loops are identified in this model, e.g. [6]: (1)from the preparation of the data to the understanding of the data in order to guide the choice of specific data processing algorithms additional or more specific information about the data is demanded; (2)from data mining to the understanding of the data poor understanding of the data causes this loop, as it leads also to the incorrect selection of a data mining method and therefore to failure, etc.

The hybrid model with its characteristics is the most suitable KDPM for deployment into the business process of the Enterprise 2.0, as it reflects the need of the business while providing a sufficiently detailed technical description with respect to data analysis. Therefore, the fielding of a hybrid model into the business process can facilitate the processing of the fast amount of generated data within an Enterprise 2.0. A good example would be an enterprise, which besides its many services and products also provides support services via a variety of channels. Such example will be reviewed closed later in this paper.

III. Data Mining using Classification Trees

Data Mining is the fourth step of the KDPM [6]. It is the step of processing the produced data in order to discover data patterns and the deriving from them knowledge. There are four basic techniques for Data Mining Regression, Association Rule Discovery, Classification and Clustering [7]. All of them can be implemented to serve the needs of support-oriented services, depending on the expected end results. This paper focuses on the usage of Classification Trees and provides an example for classifying the generated information from support-centers for the purpose of further use in the Enterprise 2.0 for product development, improvement of the provided services, decision making etc.

Decision Trees (DT), a.k.a. Classification Trees, is one of the most commonly used data mining technique, because of its model, which is easy to understand [8]. It has a flow-chart-like structure, where each internal node denotes a test on an attribute, each branch represents the outcome of this test and leaf nodes that represent classes or class distribution [9]. DTs most common task is to build models for prediction of a class of an object based on attributes. An object can be a companys customer, a patient, email, transactions, even single character. The attributes used in DT are such that describe the object. For example, in textual data like an email, the combination of the characters *failed part* or *not syncing*, can be used as attributes and refer to an object of a class. These attribute that can be used to classify the data, based on if the particular combination of characters can be found or not in the textual data. Based on the attributes, the class of an object can be either positive or negative, depending on, if the object fits certain criteria or not [10]. For better classification results of unstructured raw data, a combination of attributes can be used to describe a class. Thus, improving the classification algorithms implemented in the KDP for business process optimization in the Enterprise 2.0.

IV. Example

When fielding a new process, technology even a workflow, there are plenty of factors that should be taken into consideration. For example the very structure and functionality, the employee training, the time for introduction, security etc. Regarding the last one security of the employees as well of the information, no compromises could be tolerated. The security measures at the Enterprise 2.0 can be kept while implementing and fielding a KDP for the purpose of process optimization as it could be realized by using already available software tools of the working environment. The simplified example is for a customer support team of an Enterprise 2.0, which uses could-based platform for documentation of the support service and standard set of office work tools include at least the Microsoft Office Packet as well as other specifically designed tools to fits the need of the work process. The provided service is technical support for cars. The aim of the team to process the gathered information so that new, valid, useful and novel knowledge is discovered and a corresponding report about the discoveries is issued. As the example is simplified, it is assumed that there is only one channel for communication with the owners of the products/ customers via email.

The induced data by the communication flow is stored in a cloud-based platform like Salesforce e.g. The theoretical and practical implementation can be carried out by a single individual, who is skilled in programming. The solution is realized with Macros for Microsoft Excel using VBA (Visual Basic for Application) and it consists of the following:

1. Microsoft Excel documents serving as dictionaries for the DT algorithm;
2. VBA Macros where the classification algorithms (decision trees in particular) is implemented.

The integrated KDP in the Enterprise 2.0 is based on decision tree algorithm. For the purpose, VBA macro in Excel is implemented in order to semi-automate the processing of the generated data. It sorts out data to predefined classes based on pre-set attributes in the decision three on which the solution is. The DT algorithm is described in Fig. 1 below. Firstly the algorithm checks, if any replacement parts regarding the product are marked as used in the report that is pulled out of the could enterprise platform. If it is so, this means that the car had a hardware defect that has been repaired. Then the algorithm continues to further process the data in order to determine how was the hardware issue fixed by sending a replacement part to the customer (CSR) or by sending a technician to replace the broken part (CE onsite). In order to do this, the algorithm relies on a dictionary, where all parts that are currently on the market and suitable for the model are enlisted. So the algorithm determines, if the part is in the dictionaries and use this check to spit the classes into a total of four classes of hardware issued, to which a case can be assigned to: CSR for known part, CSR for unknown part, CE onsite for known part, CE onsite for unknown part. Other possible failures of the product are software ones. In this case, no parts have been replaced in the car, so the data are processed further using pattern matching to determine the issue in the filed case against the warranty. For the pattern matching, a dictionary is required for reference. The second Excel dictionary consists of patterns combinations defining the most commonly occurring issued. It will be used as a database of references to the different detailed representation of the attributes. Examples for patterns variations are *sync*, *Synchronization*, *lost* + *signal* and *c* + *mobile* + *fail*. By using such dictionaries, the quality of the classification can be increased in at least two different ways. Either the dictionaries are developed further by adding new possible combination the dictionary, or/ and adding different threshold values for registering a hit (classification of an object). Thus, the simple VBA implementation can guarantee the aspect of continuous improvement in the business process of the Enterprise 2.0, which is basic feature of the business process management [11, 12].

By fielding knowledge discovery into the business process can significantly reduce the average handling time for the tasks and improve the quality. The increased quality of the service is assured by unifying the process flow and introducing standards and basic rules, which can further secure the execution and the quality of the results. Thus, the end product is not only more coherent, but also the all employees are encouraged to engage in the improvement of the business processes in the Enterprise 2.0.

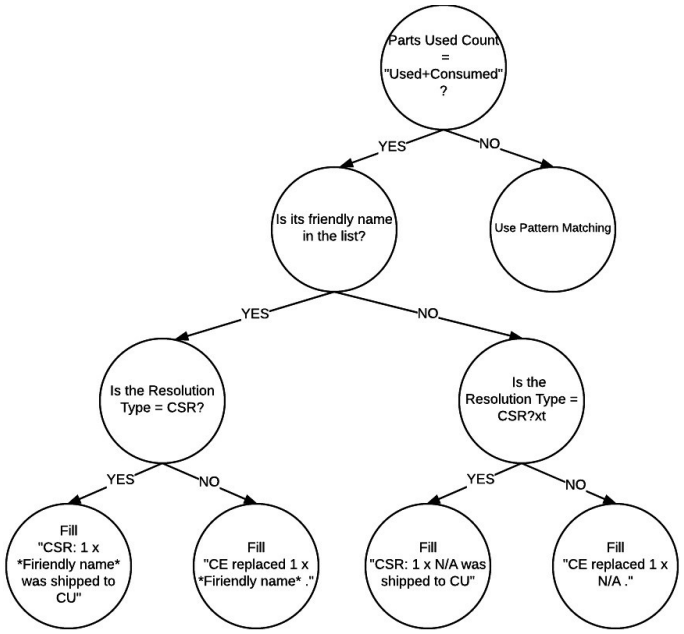


Fig. 1. Decision tree of the logic implemented in the solution

```
while Case != 0 DO
      IF Spare_Part == unused
            SearchIn(Dict_1)
            IF Spare_Part is in the Dict_1
                  IF CE_onsite==1
                        Classify(Case,Class_1)
                  ELSE
                        Classify(Case,Class_2)
            ELSE
                  IF CE_onsite==1
                        Classify(Case,Class_3)
                  ELSE
                        Classify(Case,Class_4)
      ELSE
            Use_Pattern_Matching()
            Classify(Case,Class_X)
```

Fig. 2. Exemplary implementation of the solution with pseudo code

V. Conclusion

Optimizing a business process by deploying Knowledge Discovery can result in different benefits for the Enterprise 2.0. It can help process the exponentially growing volumes of data, so that they can be transformed into valid, novel, potentially useful knowledge. Using simple automation solutions like shown in the example can not only significantly reduce the average time for processing unstructured raw data at the support center, the quality of the service, the business process, but also stimulate the employees to engage in the development process, to seek innovation and perfection.

References

[1] CISCO. (2017, June) The Zettabyte Era: Trends and Analysis. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

[2] T. O'Reilly. (2009, September) What Is Web 2.0. Design Patterns and Business Models for

the Next Generation of Software. [Online]. Available: http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1

[3] AIIM. (2009) Collaboration and Enterprise 2.0. Work-meets-play or the future of business? [Online]. Available: http://www.aiim.org/PDFDocuments/36789.pdf

[4] A. McAfee. (2006, Spring) Enterprise 2.0: The Dawn of Emergent Collaboration. [Online]. Available: http://www.wikiservice.at/upload/ChristopheDucamp/McAfeeEntrepriseDeux.pdf

[5] F. Gullo. (2015) From Patterns in Data to Knowledge Discovery: What Data Mining Can Do, Vol. 4, No. 1. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187538921500036X

[6] K. Cios, W. Pedtycz, R. Swiniaerski, and L. Kurgan, *Data Mining. A Knowledge Discovery Approach.* Germany: Springer, 2007.

[7] M. Brown. (2012, December) Data mining techniques. [Online]. Available: https://www.ibm.com/developerworks/library/ba-data-mining-techniques/

[8] ZenTut. (2017) Data Mining Techniques. [Online]. Available: http://www.zentut.com/data-mining/data-mining-techniques/

[9] A. Ashari, A. M. Tjoa, and I. Paryudi, "Performance Comparison between Naive Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 4, no. 11, pp. 33 – 39, 2013.

[10] P. Ozer, "Data Mining Algorithms for Classification," January 2008, bSc Thesis Artificial Intelligence.

[11] J. Xiao. (2017) life cycle of business process management. [Online]. Available: http://bigthink.com/articles/life-cycle-of-business-process-management

[12] S. Grier. (2008-2017) The 7 Stages of Business Process Management. [Online]. Available: http://itmanagersinbox.com/846/the-7-stages-of-business-process-management/

Automated Genome Indexing for Metagenomic Analysis

Automatisierte Genom-Indexierung für metagenomische Analyse

Kristian Nikolov*, Boris Kirov*, †, Vassil Galabov*, †, ‡

* BioInfoTech, RDIC, Sofia, Bulgaria, bioinfotechlab@gmail.com
† Faculty of Automation, TU – Sofia, Sofia, Bulgaria
‡ FDIBA, TU – Sofia, Sofia, Bulgaria

Abstract — Metagenomics analysis is a novel technique with key importance for the development of modern ecology, biomedicine and food safety. This method is tightly connected to the advance of express genome sequencing, namely the "shotgun" techniques, and depends directly on the efficiency of automated data analysis methods. In the latter years some serious breakthroughs in the field have been made both in reduction of the size of the sample DNA needed for analysis (as low as 1 ng), as well as in the efficiency of the analytic software and its practically complete independence from the way reference libraries are built. In this paper we present the next step towards complete automation of the whole metagenomics analysis, namely the development of automated genome indexing software tool that greatly reduces the amount of labor required in the analysis process.

Zusammenfassung — Metagenomikanalyse ist eine neuartige Technik, die für die Entwicklung der modernen Ökologie, Biomedizin und Lebensmittelsicherheit von zentraler Bedeutung ist. Diese Methode ist eng mit dem Fortschritt der Express-Genom-Sequenzierung verbunden, nämlich der "Shotgun"-Techniken und hängt direkt von der Effizienz der automatisierten Datenanalyse ab. In den letzten Jahren wurden einige ernsthafte Durchbrüche auf diesem Gebiet sowohl in der Verringerung der Größe der für die Analyse benötigten DNA (so niedrig wie 1 ng) sowie in der Effizienz der analytischen Software und ihrer praktisch vollständigen Unabhängigkeit gemacht die Art und Weise, wie Referenzbibliotheken gebaut werden. In diesem Beitrag stellen wir den nächsten Schritt zur vollständigen Automatisierung der gesamten Metagenomik-Analyse vor, nämlich die Entwicklung eines automatisierten Genom-Indexing-Software-Tools, das den Arbeitsaufwand im Analyseprozess stark reduziert.

I. Introduction

Microorganisms are omnipresent in the natural human environment. There are numerous species everywhere, including air, water, soil, even our food. Unfortunately, many of those microbes are pathogenic. Therefore, it is imperative for us to be able to efficiently identify the microbial composition of different samples.

The complete species composition in a foodstuff or soil material, etc. samples is studied by Metagenomics and could be detected using a DNA-Based method, which involves (Fig. 1.):

- Extraction – procedure used to isolate DNA from the nucleus of cells. The purification uses a combination of physical and chemical methods;
- fragmentation – separation of DNA strands into pieces via a variety of methods involving mechanical breakage;
- sequencing – a process of determining the primary structure of DNA (the order of guanine, adenine, thymine and cytosine nucleotides within a DNA molecule).

Genomics is the science which studies the complete set of genes belonging to each organism. It uses a variety of methods, which are commonly referred to as "standard sequencing". The latter is very efficient for sequencing the full genomes of unknown organisms. However, for the detection of species in a sample, traces of standard known organisms are looked for, not their complete genomes. In order to achieve that, a more resource-saving and innovative approach is used [1]. Instead of the whole genome, small parts (i.e. "reads") of it are sequenced and then compared to a database of known reference genomes. Read classification puts the separated and yet read pieces of DNA strands back together using the data obtained through the application of random DNA shotgun sequencing method [2]. Thus, applying a software pipeline for quantitative measurement and taxonomic profiling, we are able to identify organisms found in the sample [3].

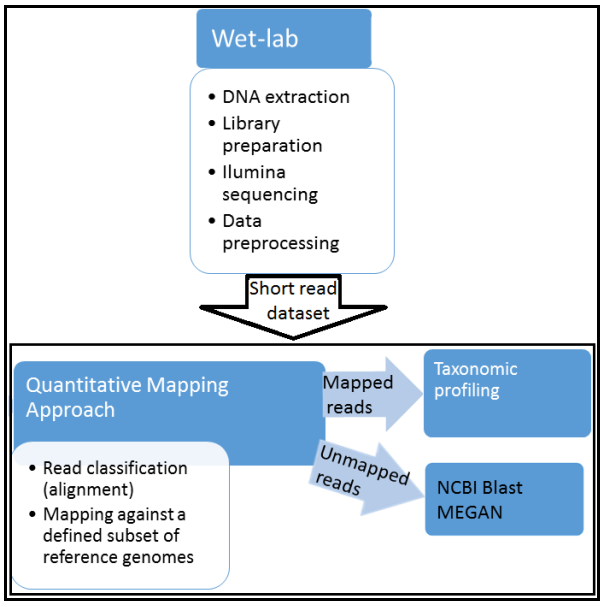


Fig. 1. Pipeline of the metagenomic analysis process.

The software, by itself, does the aligning and profiling of the reads comparing them to a database of full genomes. In order to detect a larger number of species – more genomes need to be added to the database. The latter is a time-consuming process, which could take a significant amount of time, attention, effort and energy, and also depends on internet connection speed, computer performance and size of the needed genome.

II. Genome Indexing

A. *Pre-indexing*

Full genomes are readily available in the NCBI (National Center for Biotechnology Information) database. They have to be downloaded in the FASTA file format (.fasta/.fna), which is a text-based format for representing nucleotide or peptide sequences using single-letter codes. Full genomes are commonly compressed and their size usually varies between 1MB and 1GB. The decompressed file size might expand up to 4GB.

B. *Indexing*

For the sake of time saving, every genome has to be indexed using a compressing algorithm, known as Burrows-Wheeler transform [4]. This algorithm scans for patterns, allowing mismatches and idles (which are consequence of random spontaneous mutations occurring naturally), which makes searching through genomes faster. A single indexing may require up to 50 minutes on a system with an Intel i5-4590 Quad-Core 3.30GHz CPU with 16GB of RAM and operating 64-bit Linux Mint 18.1. The number of iterations depends on the length of the indexed genome, often exceeding 100000.

C. *Post-indexing*

The sequential part is simple, but requires some patience and time, for the process is effort-consuming and labor-intensive. It involves many steps, all of which require one's supervision and attention. Sorted-pointer libraries are transferred to the analyzing software's database after a minor identifier and format conversion. The indexed genome's name needs to be included in a list for recognition. Usually, the process takes about 15 minutes to be completed, making a total of 25-65 minutes for a single organism. Given that, for proper analysis to be performed at least 20 indexed genomes are needed, the total amount of time required might exceed 10 hours.

III. Results

We have developed a script that (in the following order) (Fig. 2.):

- automates the process of full genome downloading – reaches the NCBI server and downloads all given paths freeing the user from any interactions;
- prepares the input data – decompresses the downloaded genome packs, using Gzip (GNU General Public License) and qualifies them for the next step;
- indexes the decompressed genome – runs the genome indexing software for every given genome;
- prepares the output data for the usage of the analyzing software – converts indexed genome files into the required from the metagenomic software format;
- loads formatted output data into the database – creates a directory for every organism and transfers the files into it, putting their names on the reference genome list;
- prevents future errors in case of failures;
- gives an exact count of the successfully indexed genomes and the total count of all attempted genomes;
- exerts control over the successful completion of every step and terminates the execution of the script for the failed genomes only.

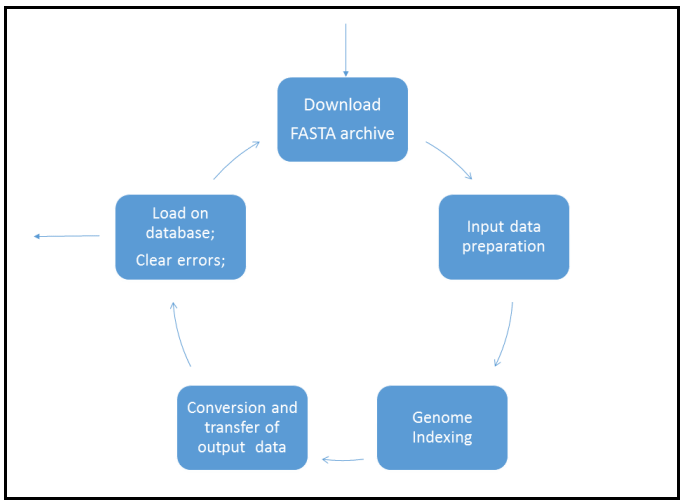


Fig. 2. Workflow of the developed script.

Thus, the user is left with the only responsibility to prepare a list of download URLs of the needed genome's FASTA files, to add their names next to each URL, and finally to run the script. In total, those actions would require only few minutes of labor overall for any number of organisms in comparison with the 15 minutes for each as described above. Therefore, we conclude that the proposed script is an efficient and useful addition to the existing software tools for automated metagenomic analysis, which could save a considerable amount of labor and time if used.

Acknowledgments

The work reported here has been conducted as part of the scientific program of BioInfoTech lab – RDIC, Sofia Tech Park. The authors thank everyone, who helped overcoming any difficulties and staying focused on the goal.

References

[1] Thomas et al., "Metagenomics - a guide from sampling to data analysis", *Microbial Informatics and Experimentation 2012*, 2:3

[2] Ripp et al., "All-Food-Seq (AFS): a quantifiable screen for species in biological samples by deep DNA sequencing", *BMC Genomics 2014*, 15:639

[3] Liu et al., "AFS: identification and quantification of species composition by metagenomic sequencing", *Bioinformatics*, 2017, 1–3

[4] Burrows, Michael; Wheeler, David J., "A block sorting lossless data compression algorithm", 1994, *Technical Report 124*, Digital Equipment Corporation, Palo Alto, CA

Passive voice quality analysis with melodic structure of previously recorded telephone conversations

Passive Voice-Qualitätsanalyse mit melodischer Struktur der vorher aufgezeichnete Telefonkonversation

Angel Garabitov
FDIBA an Technische Universität Sofia, Bulgarien

***Zusammenfassung* —** Der Artikel diskutiert Probleme bezüglich Werkzeuge für nicht-intrusive Bewertung der VoIP-Qualität durch eine melodische Strukturanalyse. Dies erfolgt durch eine objektive Sprachqualitätsmessung. Ziel dieser Untersuchung, die in diesem Artikel dargestellt ist, sei eine neue objektive Messmethode für VoIP-Anwendungen vorzuschlagen und zu untersuchen. Dies ist eine nicht-aufdringliche Sprachqualitäts- und Sprachverständlichkeitsmessmethode, basierend auf der Sprache zur MIDI-Umwandlung, gefolgt von einer klassischen, melodischen Strukturanalyse. Es wird ein Inhaltsanalyse-Schema beschrieben, das Aspekte der melodischen Struktur in großen Proben beurteilen kann. Der Artikel veranschaulicht die Anwendung des neuen Ansatzes für die VoIP-Anwendungen.

***Abstract* —** The paper discusses problems connected with tools to non-intrusively evaluate VoIP quality by melodic structure analysis. This is an objective voice quality measurement. The aims of the study, reported in the paper, are to investigate new objective measurement methods for VoIP applications. This is a non-intrusive speech quality and speech intelligibility measurement method, based on the speech to MIDI conversion sequitur classic melodic structure analysis. A content analytical scheme is described that can assess aspects of melodic structure in large samples. The paper illustrates the application of the new approach to the for VoIP applications.

I. EINFÜHRUNG

Die Notwendigkeit, die Sprachqualität in VoIP (Voice over IP) Anwendungen zu bewerten, ist eine wichtige Voraussetzung für technische und kommerzielle Gründe. Die vorhandenen Methoden sind möglicherweise nicht immer für VoIP-Anwendungen geeignet. Kürzlich ist die objektive Sprachqualitätsbewertung zu einem sehr aktiven Forschungsgebiet geworden. Dies ist ein Versuch, die Einschränkungen der subjektiven Tests zu umgehen, indem sie die Meinungen der menschlichen Tester algorithmisch simulieren. Diese nicht-aufdringliche Methode ist in Umgebungen wirksam, in denen das Referenz-Sprachsignal nicht zugänglich ist. Intrusive-Modelle sind zuverlässiger als die non-intrusive, da der erste Zugang zu einem Referenz-Sprachsignal hat, um das verzerrte Sprachsignal mit zu vergleichen.

Der Artikel diskutiert Probleme, die mit Werkzeugen verbunden sind, um die VoIP-Qualität durch MIDI-Analyse nicht intrusiv zu bewerten. Die Methode erkennt Beeinträchtigungen der Audioqualität für die menschliche Wahrnehmung [9]. Es ermöglicht die Qualität der VoIP-Verbindung auf einen Blick zu sehen und warnt, wenn sich die Qualität verschlechtert. Dies gibt die Möglichkeit das VoIP-Netzwerk zu beheben, noch bevor Benutzer von VoIP-spezifischen Verbindungsproblemen (Echo, Lärm oder Pausen im Gespräch) betroffen sind.

MIDI aus echtem Klang zu machen ist der Prozess, die Verwendung des Frequenzspektrums eines Klangs zu beschreiben und aufzuzeichnen. Der MIDI-Motor besteht aus einer Reihe von Bandpassfiltern, die nur die, in der Modulator-Quelle erkannten Frequenzen, erlauben. So, zum Beispiel, wenn eine menschliche Stimme verwendet wird, um einen Synthesizer-Akkord zu modulieren, wird es so klingen, als spreche ein Synthesizer - die klassische Roboterstimme aus vielen Sci-Fi-Filmen. MIDI-Datei zu erstellen bedeutet grundsätzlich die Realisierung eines Vocoders. Es ist ein Audioprozessor, der die charakteristischen Elemente eines Audiosignals erfasst und dann dieses charakteristische Signal verwendet, um andere Audiosignale zu beeinflussen. Die Technologie hinter dem Vocoder-Effekt wurde zunächst bei Versuchen zur Synthese von Sprache verwendet. Der Effekt, der Vocoding genannt wird, kann auf Datensätzen als "talking synthesizer" erkannt werden, der von Künstlern wie Stevie Wonder populär gemacht wird. Die bei der Vocoder-Analyse extrahierte Grundkomponente wird als Formant bezeichnet [10]. Der Formant beschreibt die Grundfrequenz eines Schalls und seiner zugehörigen Rauschkomponenten. Der gleiche Vorgang kann zur Analyse und Bewertung des VOIP-Signals verwendet werden. Dieser Vorgang kann die Defekte und Verzerrungen während der Übertragung des Signals erkennen. Wenn wir die menschliche Stimme als Melodie betrachten, dann werden alle Seitenverzerrungen ihre Melodie beeinflussen. Die Veränderung kann auf verschiedene Weise beurteilt werden. Eine von ihnen ist die Auswertung der melodischen Erwartung. Wenn jemand eine Melodie hört, ist es üblich eine Erwartung zu haben, was als nächstes kommen wird. Das alles gilt für die Rede. Eine wichtige Studie zu diesem Thema ist das Implikation-Realisierungsmodell von Narmour [7], das die melodische Erwartung vernünftigerweise vorhersagen kann. Hier werden einige Hauptpunkte erwähnt. Das erste Prinzip des Modells ist eine Registernummer und es bedeutet, dass kleine Intervalle dazu neigen, eine Fortsetzung der Tonhöhenrichtung zu verfolgen, während große Intervalle eine Richtungsänderung erwarten. Intervalldifferenz ist das Prinzip, dass kleine Intervalle andere Intervalle von ähnlicher

Größe implizieren, während große Intervalle kleinere Intervalle implizieren. Wenn ein zweites Intervall zu einer Tonhöhe führt, die der ursprünglichen Tonhöhe nahekommt, haben wir das Prinzip der Registrierungsrendite. So erwarten die Zuhörer, dass die Sprünge auf eine ähnliche Tonhöhe zurückkehren. Das Prinzip der Nähe definiert einfach, dass kleine Intervalle eher erwartet werden als große Intervalle. Das letzte Prinzip ist die Schließung und es tritt auf, wenn die Melodie die Richtung ändert oder wenn ein großes Intervall von einem kleineren Intervall gefolgt wird. Das Modell wurde von Frau Carol L. Krumhansl [1] quantifiziert. In ihrer Arbeit schlägt sie auch ein neues Prinzip namens Konsonanz vor. Dieses Prinzip besagt, dass Unisionen, perfekte Viertel, Fünftel und Oktaven bevorzugte Intervalle sind. Dies ist jedoch eine sehr kontextspezifische Behauptung, da Melodien je nach Genre viele weitere chromatische und dissonante Intervalle haben können.

Sprachauswertungssysteme sind Systeme, die ein akustisches Signal durch Algorithmen analysieren. Diese Algorithmen basieren auf einer Anzahl von Theorien, die vorschlagen, welche Merkmale des Sprachsignals in einer gegebenen Sprache einen Klang erzeugen. Mathematische Methoden bestimmen die wichtigen Parameter des akustischen Signals und verwandeln sie in eine unterschiedliche Vollständigkeit in der gewünschten Form [6].

Es gibt Möglichkeiten, diese Erwartungen in einer Evaluierungsfunktion zu nutzen. Die Modelle sind rechenintensiv, da sie ihre Ergebnisse auf die Zeit- und / oder Frequenzbereichsanalyse des zu prüfenden Sprachsignals stützen. Sie verlangen auch, dass der Testaufruf für eine beträchtliche Dauer aufgezeichnet wird, bevor er analysiert werden kann. Daher eignen sie sich nicht für die Echtzeit- und kontinuierliche Überwachung der Sprachqualität, sind aber für die anschließende Analyse anwendbar.

II. Vorgeschlagenes Algorithmus für VoIP Qualität durch Midianalyse

Die Aufgabe besteht darin, das Sprachsignal zu studieren, Parameter zu identifizieren und zu visualisieren, die Klänge zu erzeugen (Phoneme), diese Parameter messen und klassifizieren.

Die Studie untersuchte folgende Aspekte:

- Signalverarbeitung;
- Hörmuster;
- Artikulationsmodelle;
- Aussprachemuster;
- Suchalgorithmen;
- Lernalgorithmen;

Die möglichen Optionen, Metadaten aus dem Sprachsignal abzurufen, können mit denen verglichen werden, die die Qualität einer musikalischen Phrase auswerten [2]. Ähnliche musikalische Analyse-Aufgaben wurden weitgehend entwickelt und hier ist ein Versuch, sie auf die Klanganalyse anzuwenden. Die musikalische Analyse selbst impliziert die erste Umwandlung in ein entsprechendes Format.

Umwandlung in MIDI

Ein Sprachcodierungssystem zum Codieren eines digitalisierten Sprachsignals in ein Standard-Digitalformat, wie beispielsweise MIDI. Das Pitch-Detection-System ist gut, aber nicht gut genug bei komplexen oder harmonisch reichen Inhalt. Was es tut, ist die Analyse und Annäherung der Häufigkeit dessen, was er summt und in eine Folge von MIDI-Noten umwandelt.

Die stimmhaften Töne resultieren aus den Vokal-Akkorden, die vibrieren und so den Luftstrom aus den Lungen unterbrechen und einen Frequenzbereich von Klängen von etwa 50 bis zu 500 Hz erzeugen [5].

Plosivgeräusche sind plötzliche Luftstöße, die zum Beispiel ausgeschaltet werden, wenn der Vokaltrakt geschlossen ist und plötzlich freigegeben wird oder der Mund plötzlich geöffnet wird. Alle diese Töne werden durch die Sünden und Nasenhöhlen der Person beeinflusst und alle machen das, was wir als normale menschliche Sprache verstehen.

1) Grundlagen Klangbeispiele und Entschlüsselung der MIDI-Datei

Die verwendete Probe ist von Open Speech Repository [11] heruntergeladen. Die Aufnahmebedingungen sind Paketverlust 50% und Korrelation 15%. Für die Erstellung des Fehlersignals wird der Befehl Linux tc verwendet. In Abbildung 1 ist der Vergleich des guten und schlechten Signals dargestellt.

Abb. 1. Gutes und schlechtes Signal

2) MIDI-Extraktion mit dem Signal

In Tabelle I ist der Vergleich des guten und schlechten Signals in MIDI Format dargestellt.

TABELLE I. MIDI Format

Gutes Signal	***Schlechtes Signal***
MFile 1 2 120 MTrk 0 Meta Copyright "Created by TS-AudioToMIDI 3.30" 0 TimeSig 4/4 24 8 0 KeySig 0 major 0 Tempo 500000 0 Meta TrkEnd TrkEnd MTrk 22 PrCh ch=1 p=0 22 On ch=1 n=39 v=23 58 Off ch=1 n=39 v=0 75 On ch=1 n=38 v=18 … 876 Meta TrkEnd TrkEnd	MFile 1 2 120 MTrk 0 Meta Copyright "Created by TS-AudioToMIDI 3.30" 0 TimeSig 4/4 24 8 0 KeySig 0 major 0 Tempo 500000 0 Meta TrkEnd TrkEnd MTrk 73 PrCh ch=1 p=0 73 On ch=1 n=24 v=20 85 Off ch=1 n=24 v=0 86 On ch=1 n=34 v=24 … 813 Meta TrkEnd TrkEnd

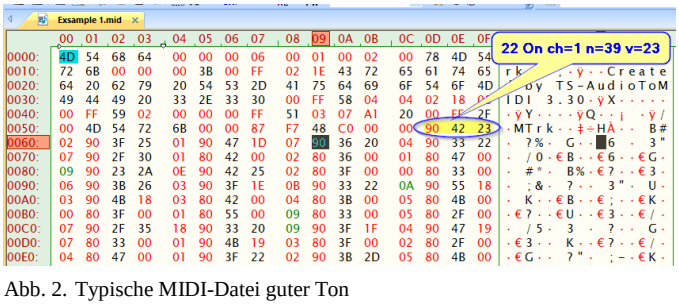


Abb. 2. Typische MIDI-Datei guter Ton

In Abbildung 2 ist ein MIDI Dateien in einen Hexformat dargestellt

3) Konvertieren von MIDI-Dateien in Matlab

Die grundlegenden Funktionen in MIDI Toolbox lesen und manipulieren Typ 0 und Typ 1 MIDI-Dateien. Nicht Ematrix (oder cmat) bezieht sich auf eine Matrixdarstellung von Notenereignissen in einer MIDI-Datei.

TABELLE II. MIDI Kodierung

ONSET *(BEATS)*	***DURAT.*** *(BEATS)*	***MIDI*** *channel*	***MIDI*** *pitch*	***VELO-CITY***	***ONSET*** *(SEC)*	***DURAT.*** *(SEC)*
0.6083	0.1000	0	24.0	20.0	0.3042	0.0500
0.7167	0.1417	0	34.0	24.0	0.3583	0.0708
0.8583	0.1583	0	32.0	21.0	0.4292	0.0792
1.1583	0.2000	0	34.0	42.0	0.5792	0.1000
1.6750	0.1750	0	48.0	49.0	0.8375	0.0875
1.8417	0.1000	0	36.0	26.0	0.9208	0.0500

Die erste Spalte zeigt den Beginn der Noten in Beats an (auf der Grundlage von Zecken pro Viertelnote) und der zweiten Spalte die Dauer der Noten in denselben Beat-Werten. Die dritte Spalte bezeichnet den MIDI-Kanal (1-16) und die vierte MIDI-Tonhöhe, wobei der mittlere C (C4) 60 ist. Die fünfte Spalte ist die Geschwindigkeit, die beschreibt, wie schnell die Taste der Note gedrückt wird, mit anderen Worten, wie laut die Note gespielt wird (0-127). Die letzten beiden Spalten entsprechen den ersten beiden (Beginn in Beats, Dauer in Beats), außer dass Sekunden anstelle von Beats verwendet werden.

4) Statistische Auswertung der Dateien in Matlab

Zuerst können wir die Notenverteilung der Klangbeispiele untersuchen.

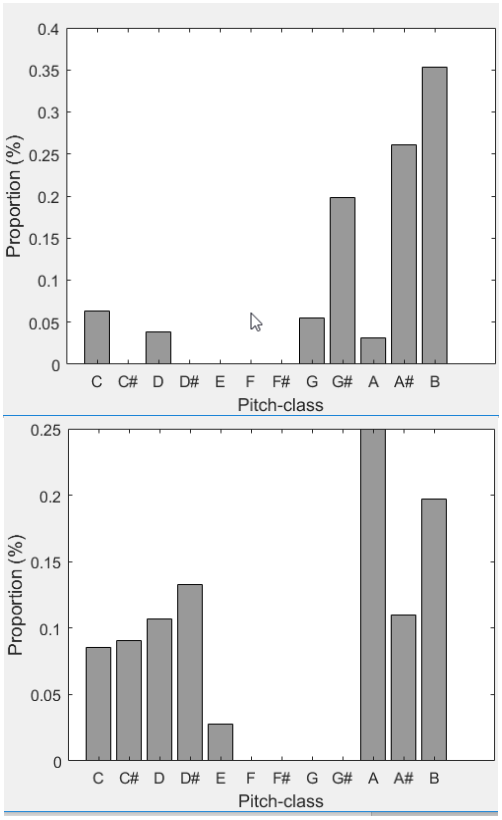


Abb. 3. Notizverteilung der Klangbeispiele

Der Anteil der Zwei-Noten-Fortsetzungen - Anteil der Tonübergänge

Die melodische Kontur beschreibt die Gesamtform der Klangprobe. Die durchgezogene Linie stellt eine gröbere melodische Kontur dar, die nützlicher sein könnte, um die Gesamtstruktur der Tonprobe herauszufinden.

Diese Visualisierung der Tonalität kann verwendet werden, um die Schwankungen des Schlüsselzentrums und der Schlüsselstärke im Laufe der Zeit zu zeigen.

Diese Visualisierung der Tonalität kann verwendet werden, um die Schwankungen des Schlüsselzentrums und der Schlüsselstärke im Laufe der Zeit zu zeigen.

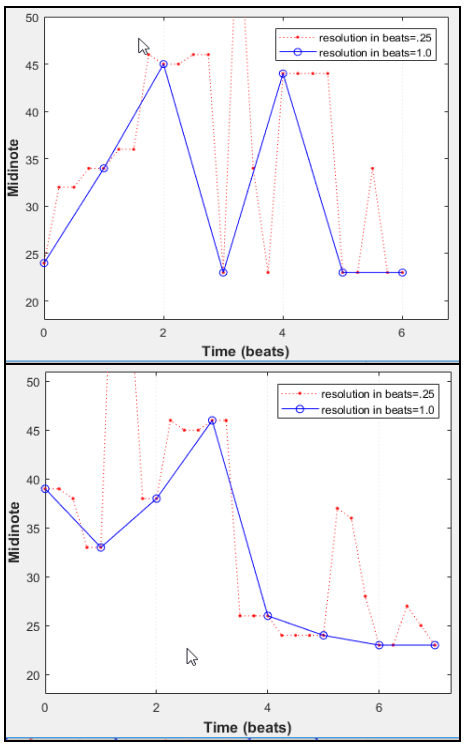


Abb. 4. *Gesamtstruktur der Tonprobe*

5) Melodische Erwartungen und Sprachverständlichkeit [3].

Die früheren Arbeiten zur melodischen Erwartung haben gezeigt, wie sich die Musik auf die üblichen psychologischen Erwartungsprinzipien stützt, die in kognitiv orientiertem musiktheoretischen Modell von Narmours [7] gefangen wurden. Das Modell stützt sich auf die Gestalt-basierten Prinzipien der Nähe, der Ähnlichkeit und der guten Fortsetzung und wurde gefunden, um die melodischen Erwartungen der Zuhörer recht gut vorhersagen zu können. Das Modell arbeitet mit Blick auf implizite Intervalle und realisierte Intervalle. Der ehemalige schafft Implikationen für die Fortsetzung der Melodie und das nächste Intervall führt seine Implikationen aus.

6) Melodische Komplexität und Sprach-verständlichkeit [3].

Gelegentlich ist es interessant zu wissen, wie kompliziert, schwierig oder "originell" eine Melodie der Rede ist. Diese Beziehung ist in Form einer invertierten U-Funktion, wo die populärsten Themen von mittlerer Originalität sind. Infolgedessen sind die einfachsten Themen nicht populär (sie können als "banal" angesehen werden) und auch nicht die komplexesten. Es gibt auch andere Verwendungen für eine melodische Komplexitätsmaßnahme, wie sie es als Hilfe bei der Klassifizierung von melodischem Material zu verwenden (Toiviainen & Eerola, [3]). Simontons Modell der melodischen Originalität basiert auf Tonübergangswahrscheinlichkeiten. Die Ausgabe dieses Modells (compltrans) erzeugt eine Umkehrung der gemittelten Wahrscheinlichkeit, skaliert zwischen 0 und 10, wo ein höherer Wert eine höhere melodische Originalität anzeigt.

Eine weitere Möglichkeit, die melodische Komplexität zu beurteilen, besteht darin, sich auf die Klang- und Akzentkohärenz zu konzentrieren, und auf die Höhe der Pitch-Skips und die Kontur-Selbstähnlichkeit zeigt sich die Melodie [4]. Eine alternative Maßnahme der melodischen Komplexität ist in der kontinuierlichen Messung der Notenereignisverteilung (Pitch-Class, Interval) Entropie (Use Mov-Fenster und Entropie und verschiedene Verteilungsfunktionen) verankert. Diese Maßnahme erzeugt für jeden Punkt in der Melodie melodische Vorhersagungswerte (daher der Begriff kontinuierlich). Es wurde festgestellt, dass diese Werte den Vorhersehbarkeitsangaben entsprechen, die von den Zuhörern in Experimenten gegeben wurden. Diese Maßnahme bietet die Möglichkeit, die Moment-in-Moment-Schwankungen der melodischen Vorhersagbarkeit zu beobachten.

Viele Software- und Hardwarelösungen wurden implementiert.

III. Berechnen der melodischen Funktionen

Es gibt viele Möglichkeiten, um die Melodie zu bestimmen, die zu unserem Zweck anwendbar ist. Im Wesentlichen sind sie sehr komplex zu beschreiben. Die einfachste Methode zur Anwendung ist die Methode von Euler (Grad der Melodie [8]).

Euler schlug vor, dass der Grad der Melodie von den Berechnungen des Geistes abhängt: je weniger die Berechnungen, desto angenehmer die Erfahrung.

Eine geringe Anzahl von Berechnungen führt zu einem hohen Wert für die Melodie, während eine hohe Anzahl von Berechnungen einen niedrigen Wert ergibt.

Dieses Prinzip wird durch eine numerische Technik umgesetzt, die auf der Zerlegung der natürlichen Zahlen in ein Produkt von Kräften verschiedener Primzahlen basiert.

Wenn p_i ..., p_n sind verschiedene Primzahlen und e_1, ..., e_n sind verschiedene Kräfte, dann kann jede natürliche Zahl a ausgedrückt werden als

$$a = p_1^{e_1} p_1^{e_2} \cdots p_1^{e_n} = \prod_{i=1}^{n} p_i^{e_i} \qquad (1)$$

$$G(a) = 1 + \sum_{k=1}^{n} e_k (p_k - 1) \qquad (2)$$

```
function y = gradus(nmat)
if isempty(nmat), return; end
if ~ismonophonic(nmat);
return; end
int=abs(diff(pitch(nmat)));
int=mod(int,12)+1;
frq1=[1,16,9,6,5,4,45,3,8,5,16,15];
frq2=[1,15,8,5,4,3,32,2,5,3,9,8];
for k=1:length(int)
        n(k)=frq1(int(k));
        d(k)=frq2(int(k));
end
for s=1:length(n)
        gs(s)=suavitatis(n(s),d(s));
end
y=mean(gs);
        g=n*d;
        g2=factor(g);
        for i=1:length(g2)
                g3(i)=1*g2(i)-1;
        end
        g=sum(g3)+1;
```

Abb. 5. Algorithm in Matlab

Euler benutzt Primzahlen in der Bemühung, die Melodie, den "Grad der Süße" zu quantifizieren - oder wie er es die "gradus suavitatis" - der Töne nannte. Der Gradus suavitatis einer einzigen Note wurde als 1 und darüber hinaus genommen, wenn das Frequenzverhältnis von zwei Noten ***m: n*** ist und wenn das kleinste gemeinsame Vielfache von ***m*** und ***n*** ***L*** ist, hat er die Definition gemacht.

$$G(m,n) = 1 + \prod_{\substack{p\ prime \\ p\ divides L}} (p-1) \qquad (3)$$

Das Algorithm von Euler ist in Matlab in realisier. (Abbildung 5)

IV. Schlussfolgerungen und zukünftige Arbeiten

Das Problem der Echtzeit-Schätzung von VoIP ist von erheblichem Interesse. Dieses Papier hat einen Ansatz zur Lösung dieses Problems gezeigt, indem er die Melodien des Signals einsetzt. One of the main objectives of this research was to estimate the effect of fragmentation on speech quality. Eines der Hauptziele dieser Forschung war es, die Wirkung der Fragmentierung auf die Sprachqualität abzuschätzen.

Der Fokus der aktuellen Forschung liegt auf der Schätzung der Wirkung aller VoIP-Verkehrsparameter, die die Hörqualität eines Telefonanrufs im Mähdrescher beeinflussen. Ein zukünftiges Ziel wäre es, ein neuronales Netzwerkmodell für die Abbrsation-Qualitätsschätzung eines Aufrufs abzuleiten. Die Konversationsqualität leidet unter der Zunahme der End-to-End-Verzögerung eines Anrufs. Klar wäre das nächste Ziel, die besondere Wirkung von VoIP-Verkehrsparametern und deren Auswirkungen auf die Signalqualität abzuschätzen.

Die digitale Signalmelodie-Bewertung wird verwendet, um ein neuronales Netzwerk zu trainieren, um die Klangqualität von VoIP zu beurteilen.

Danksagung

Diese Publikation wurde finanziert durch Vertrag zur Doktorandenunterstützung mit Nummer 172PD0010-07 der TU – Sofia.

Literaturverzeichniss

[1] Carol L. Krumhansl, *Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press, 1990, 307pp, ISBN 0-19-505475-X.

[2] David Huron and Matthew Royalp "What Is Melodic Accent? Converging Evidence from Musical Practice", Music Perception: An Interdisciplinary Journal, Vol. 13, No. 4 (Summer, 1996), pp. 489-516 Published by: University of California Press Stable, URL: http://www.jstor.org/stable/40285700

[3] Eerola, T. & Toiviainen, P. *MIDI Toolbox: MATLAB Tools for Music Research*. University of Jyväskylä: Kopijyvä, Jyväskylä, Finland, (2004).

[4] Marc Lema, "Music and Schema Theory: Cognitive Foundations of Systematic Musicology", Springer Science & Business Media, Dec 6, 2012 - Science - 234 pages

[5] Marcel Wältermann, *Dimension-based Quality Modeling of Transmitted Speech*, Springer-Verlag Berlin Heidelberg 2013.

[6] Moller, S., "Assessment and Prediction of Speech Quality in Telecommunications", Kluwer Academic Publishers, Boston/Dordrecht/London, 2000, 116-117.

[7] Narmour, E. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago: University of Chicago Press, 1990, ISBN 0-226-56845-8.

[8] Patrice Bailhache, *Euler and music*, Nantes, 1997

[9] Patrik N. Juslin, John Sloboda, *Handbook of Music and Emotion: Theory, Research, Applications*, Oxford University Press, Mar 17, 2011 - Psychology - 992 pages

[10] Trevor A. Harley, *Speech Perception and Spoken Word Recognition, Current Issues in the Psychology of Language Series*, first published 2017 by Routledge.

[11] http://www.voiptroubleshooter.com/open_speech/american.html.

Framework for implementing, evaluating and analyzing some algorithms for non-photorealistic image processing

Framework zur Implementierung, Auswertung und Analyse von Algorithmen für nichtfotorealistische Bildverarbeitung

Virginia Todorova Dimitrova

Faculty of Computer Systems and Technologies, Technical University of Sofia
Sofia, Bulgaria, vergy@tu-sofia.bg

***Abstract* —** The article presents a desktop application that supplies a framework for designing, implementing, testing and evaluating different techniques for non-photorealistic image processing. The framework supplies a common template and an effective interface for programmers who want to implement and analyze a lot of non-photorealistic image processing algorithms. The goal was to allow users to investigate the applicability, functionality, effectiveness of already implemented algorithms for non-photorealistic image processing as well as to implement new versions of the same algorithms and new algorithms in the target area. The main focus of the realization was to develop an extendable system of closely interconnected modules that give the user full control over the total process of loading image, performing needed preprocessing steps, applying selected image processing technique, viewing and storing the resultant images, obtaining information concerning effectiveness of the applied algorithm and manage the history of the performed actions. In this stage the framework can be used for pixel-based non-photorealistic algorithms including such popular techniques as digital halftoning and screening.

***Zusammenfassung* —** Der Artikel präsentiert eine Desktop-Anwendung, die ein Framework für das Entwerfen, Implementieren und Vergleichen verschiedener Techniken für die nicht-fotorealistische Bildverarbeitung liefert. Das Framework bietet eine gemeinsame Vorlage und eine effektive Schnittstelle für Anwender, die viele nicht-fotorealistische Bildverarbeitungsalgorithmen implementieren und analysieren möchten. Ziel ist es, den Anwendern die Möglichkeit zu geben, die Anwendbarkeit, Funktionalität und Effektivität bereits implementierter Algorithmen für die nicht-fotorealistische Bildverarbeitung zu erforschen sowie neue Versionen derselben Algorithmen und neue Algorithmen im Zielbereich zu implementieren. Der Schwerpunkt der Realisierung liegt in der Entwicklung eines ausziehbaren Systems von eng miteinander verbundenen Modulen, die dem Anwender die volle Kontrolle über den gesamten Prozess des Ladens eines Bildes, die Durchführung der benötigten Vorverarbeitungsschritte, die Anwendung der ausgewählten Bildverarbeitungstechnik, das Betrachten und Speichern der Ausgabebilder, um Informationen über die Wirksamkeit des angewandten Algorithmus zu erhalten und die Geschichte der durchgeführten Aktionen zu verwalten. In diesem Stadium können das Framework für pixelbasierte nicht-fotorealistische Algorithmen sowie solcher populären Techniken, wie digitales Raster und Screening, verwendet werden.

I. Introduction

Images are fast becoming a generic data type for general-purpose computer systems and thus a need to process a huge amount of digital image data appears. One of the goals of this type of image processing can be formulated as creating imagery that is effective at conveying information, expressive and beautiful. Image processing algorithms that lay in the core of non-photorealistic rendering (NPR) have the potential to achieve this goal through generating by computer and by user demand images that are effective, attractive and appropriate in the context of the human-computer dialog. These algorithms once implemented will allow to generate meaningful imagery for end users thus enabling them to observe information in graphical form. An overview of NPR techniques can be found in the textbooks by Strothotte and Schlechtweg [1], Gooch and Gooch [2].

A lot of algorithms in image processing are well known from decades and effectively implemented as a build-in functionality of popular program packages like Photoshop, Photoscape, GIMP. Groups of internally connected algorithms are implemented as plug-ins for such program packages. A huge variety of image-processing algorithms are exposed as library functions that can be invoked from applications developed in different programming languages, for different platforms as well as for a lot of devices (desktop, laptop, tablet, mobile devices). The goal of the present work is completely different in at least 2 aspects. First, most of image processing algorithms are implemented in such a way to reduce the artifacts in images induced as a result of the applied image-processing techniques. The present work adopts the same algorithms but slightly modified to induce image artifacts allowing the achievement of various visual effects typical for NPR Second, the above mentioned implementations of algorithms (as build-in functionality, plug-ins, libraries of

functions) allow users to see the results of different image-processing techniques applied to a lot of images. However, these implementations do not allow the user to modify a particular algorithm, to easy add new implementations of the existing or new algorithms, to analyze and evaluate the effectiveness of different implementations.

So, the goal of the presented work is to develop a program solution that serves as a framework for implementation, evaluation and analyzing of selected algorithms in NPR. The user has the ability to use a known template and a common user interface in the process of implementing, evaluating and analyzing the algorithms.

The present work covers a useful area of non-photorealistic image processing called halftoning as well as a special kind of halftoning called screening. There are different techniques for halftoning like patterning, ordered dithering, error diffusion and for each technique there is a variety of algorithms for its realization. The common feature of these algorithms is that they encode the "subjective" artifacts into the resulting image that clearly do not exist in the original image.

All implemented NPR techniques work on the basis of altering images that encode intensities on a per-pixel basis. Intensities are created either by a photorealistic renderer or in a real photographic process. The image is interpreted simply as a pixel matrix, where each entry is either a gray level or a color value (RGB). The pixels are manipulated to produce various different visual effects.

All discussed algorithms require an input image in Bitmap format. Some of the algorithms need no more information than that directly accessible from an input image. Others require a user input in the form of numerical parameters, short text, or reference images. These additional parameters are specific for each class of algorithms and are very important, because the obtained visual effects strongly depend on the values of these parameters. The resultant non-photorealistic images with desired visual effects are generated semi-automatically by analyzing input images, extracting certain information, and accepting a lot of user input.

The main contributions of this paper are two: support a complete programming tool for implementing, evaluating and analyzing some NPR techniques; and demonstrate with a set of already implemented algorithms the possibilities of these techniques. During the program realization are created both: 1) a library of already implemented algorithms; and 2) a couple of modules that allow to implement additional algorithms with no need to think about some common steps like loading images, presenting images, logging the performed actions.

The specific contributions of this paper are:

- An extendable library of algorithms implemented in C#;
- A common template and a common user interface;
- Interfaces for an interactive design of patterns, dither matrices and artistic screens.

II. Related Work

Advances in the field of NPR show the usefulness of image-space approaches. An important class of NPR rendering styles operates directly in image space, such as halftoning, screening, stippling, or hatching methods. In the field of NPR there are two branches: techniques for producing NPRendering from 3D models [3, 4] and techniques for applying NPR styles to 2D images. Typically graphics packages for 3D modeling like 3DS MAX supply tools that allow to use image space techniques to post-process renderings of 3D scenes in order to obtain a lot of non-photorealistic styles. For example, 3DS MAX exposes a special kind of material named Ink'n Paint that can be applied to one or more objects of a 3D scene to render them in a cartoon style. Additionally, the user has the possibility to view the created 3D scene in a lot of non-photorealistic styles. NPR algorithms implemented in 3ds max usually use 2 ½ D (G-buffers) and 3D data as an information source. Some image-processing algorithms in 3ds Max are applied to the content of a variety of G-buffers generated as a co-product of the obvious rendering process [5].

The second branch of NPR encompasses algorithms for artistic drawing styles that work directly in image space. A plenty of such algorithms are implemented in commercial programs for image processing like Photoshop. Photoshop offers a lot of filters that allow turning a photo into a painting. These filters are grouped in different categories (Artistic Filters, Texture filters, Stylized filters and many others) and replicate natural or traditional media effects or simulate the effect of halftone screen. Some NPR techniques for image halftoning such as ordered dithering and error diffusion are widespread in programs for image processing, which are desktop and web applications like MATLAB, Ximagic GrayDither for Photoshop, Meemo. MATLAB provides Bayer dithering [7, 8] by supplying dither() function which converts the grayscale input image into the binary image. The web application "Meemo" supplies Akiston, Bayer and Floyd-Steinberg algorithms for automatically transforming an imported image into a binary image [9]. Ximagic GrayDither is a Photoshop plugin that allows to reduce color/grayscale images to an n-level grayscale using dithering [10]. For Windows platform a lot of image filters some of which non-photorealistic (like Image Distortion Blur Filter) are implemented as separate applications and included in MSDN [11]. All of them supply a lot of parameters that the user can set to change slightly the obtained visual effects.

In the area of NPR there are also a lot of software tools supplied in a form of software development kits like NPR Kit for MODO and REDsdk. The NPR Kit for MODO supports the creation of a variety of familiar artistic effects in MODO including 'toon' shading, stippling, halftones and edge rendering. REDsdk is a software tool that supplies two styles of rendering: Cartoon rendering and Scetch rendering, and can be used by programmers to generate NPR images [12]. Some image-space algorithms that can be classified as non photo-realistic are implemented as classes that make them easy accessible for programmers (EffectFactory class in Android is such a class) [13].

III. Structure and Scope

Overall structure of the presented work is shown in fig. 1. To give a sense of the system from users' viewpoint the modules structure and the organization of the user interface are briefly described.

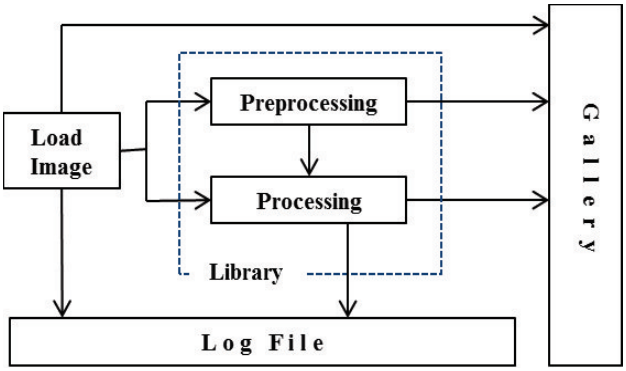


Fig. 1. The main structure of the presented work.

One of the main modules is a library of already implemented NPR techniques (Library). The Library was developed as a separate project using ClassLibrary template of Visual Studio .NET. A reference to the Library was included into the application and all implemented algorithms are accessible through the Preprocessing and Processing top-level menu commands. Each non-photorealistic image-processing algorithm is implemented as a separate method. Methods that perform similar types of NPR techniques pertain to a separate class. This is the approach applied by the author to categorize the included NPR techniques. The name of the class and the name of the method form a fully qualified name used to specify and subsequently identify the corresponding NPR technique. When a new NPR technique has to be implemented the selection of the appropriate class which includes the definition of the method is the responsibility of the programmer.

Another module is the module for loading input images for processing from disk files in one of the supported formats (BMP, GIF, EXIG, JPG, PNG and TIFF). During the loading of the input image metadata form the image file is extracted, a new object of ImageProperties class is created and its fields are partially populated on the base of the extracted metadata. The object of ImageProperties class is stored in the collection of such objects and a corresponding record in the Log file is added. Another operations performed by this module include: creating a Bitmap object that consists of the pixel data for a graphics image and its attributes; and resizing the picture box control in order to prevent interpolation of the input image.

Since one of the main goals of the present work is to allow the user to evaluate and analyze different NPR techniques, the framework supplies two important modules. The first one is the Log file, created and maintained automatically by the application. This Log file is created and opened for writing each time a new input image is loaded for processing. Then an information about the loaded input image and about all actions performed over it is regularly recorded into the file. The collected information is stored as a text file with separate paragraphs each beginning with a predefined keyword (Image, Technique, Elapsed Time). The content and the structure of the Log file are designed to be used for analytical evaluation as well as for subsequent analyzing of the implemented and applied NPR techniques. The user can compare the parameters of the input image and the output images; trace the applied preprocessing and processing steps and see the elapsed time for each applied NPR technique by reading the paragraphs in a log file. It is possible to view the content of the log file at all times during the current image processing session.

The last module is the so called Gallery dedicated mainly for visual evaluation of the obtained results in a particular image processing session. The Gallery is a separate window that shows the input image as well as all output images obtained as a result of applied preprocessing and image processing techniques. Each image is represented in the Gallery as a thumbnail image in other to obtain a common view over the results of the particular image processing session. The user however has the possibility to see each included image in real dimensions. There are two options for each represented in the gallery output image: to obtain detail information concerning the image and to delete the image. Additionally, because the zoomed part of each processed image is of great importance for evaluation, the user has the possibility to interactively select a pixels' region and zoom it thus making the induced artifacts more visible and available for examining.

The main focus in the realization is on achieving the high level of interactivity. So, all user actions from loading an input image to presenting the output images are performed interactively. Conventional menus and dialog boxes are used when needed. The main challenge is on preparing the additional source data consisting either of simple numerical and textual input, or of more complex pixel patterns, dither matrixes, threshold matrixes and dither screens. First, this source data has to be prepared in such a way to induce artifacts in original images to achieve various visual effects. Second, the user must have the possibility to dynamically and easily change this source data. In order to achieve the needed level of interactivity a separate custom dialog boxes are supplied for preparing pixel patterns, for preparing dither matrices and for histogram equalization.

IV. Details of Realization

Most image files contain metadata that allows to determine features of the image such as the number of bits per pixel, number of pixels in each row, and number of rows in the array, a color table. This metadata can be read programmatically, maintained in the memory during the image processing and saved in the output file. One of the ways to save information about the applied image processing techniques over the content of the loaded input image is to use the metadata stored in an image file using the predefined PropertyItem class and the Image.PropertyItems property. GDI+ stores an individual piece of metadata in a PropertyItem object. This object can be used both to read existing metadata (GetPropertyItem method) and to write new metadata to image files (SetPropertyItem). But a PropertyItem object has only four properties: Id, Value, Len, and Type that are not enough for the goals of the present work. Additionally, it is difficult to set property items using these predefined classes. So, the author adopts another approach to manage and store metadata – create a custom class ImageProperties. All objects of ImageProperties class are stored in a separate collection in the memory. This information is the corner element of the Log file and is used to explain details about the selected image in the Gallery.

One of the focuses in the realization is on storing, managing and representing the results. During each session of image processing a lot of output images are created. For each input image as well as for all generated output images a corresponding thumbnail image is created. Both the input image and the set of output images are stored in memory for the time of the image processing session in two separate collections: one collection of images with original size and the other – with their thumbnail versions. This approach allows the user to see by demand the Gallery of all images he/she works with in the current session represented as thumbnails as well as to see each selected image of the Gallery in real dimensions.

V. Functionality

A. Preprocessing

Some of the implemented NPR techniques require one or more preprocessing steps. These preprocessing steps include: conversion of a colored image into a grayscale image; conversion of a grayscale image into a binary image; creating an image histogram; performing histogram equalization. There are a lot of algorithms for conversion of a colored image to a grayscale image (grayscale filters like "single color channel", "averaging", "luminance", "desaturation", "de-composition" and so on [14]). The presented work supplies 4 such algorithms among which the user can make his/her choice. The simplest method for converting a grayscale image to a binary image is by thresholding, i.e. a two-level quantization. The implemented algorithm allows simple quantization using a fixed threshold in the quantization process. Preprocessing is used also for preparing a dither screen with needed qualities for the

implemented screening NPR technique. For this purpose the framework exposes two methods: to compute a histogram of an input image that will be used as a threshold matrix; and to perform a histogram equalization in order to achieve a uniform distribution of the gray values.

B. Processing

The first implemented halftoning technique is a patterning technique. The patterning algorithm is implemented for a bilevel system and a rectangular grid of n by n pixels, allowing to produce n+1 intensity levels. The user has the possibility to specify the size n of the rectangular grid of pixels (pixel-grid size). Pixel patterns for any pixel-grid size are specified with a "mask" of pixel position numbers. The user has the possibility to populate the mask interactively and to see the generated pixel patterns.

Next implemented halftoning technique is the error diffusion technique that utilizes the result of a simple threshold quantization and redistributes the resulting error to the surrounding region of a pixel. Generally, there is a wide range of algorithms for error diffusion. Only the basic and most popular error diffusion algorithm proposed by Floyd and Steinberg [15] is implemented in the Library. Implementations of other algorithms for error diffusion are left to the users of the framework.

The slightly modified Floyd-Steinberg algorithm is used for implementation of a specific halftoning technique - halftoning using lines. This separate NPR technique makes use of error diffusion to get an initial pixel distribution which is then changed to introduce image artifacts explicitly. Artifacts are introduced by drawing short lines that cover several pixels instead of single dots in the output image.

The next implemented halftoning technique is ordered dithering technique. The emphasis of the implementation is on the designing of dither matrices in which pixels are grouped in such a way to form visual patterns to be introduced in the image for achieving interesting effects. A dither matrix was defined by a so called index matrix that is a special case of a family of dither matrices first defined by Bayer [16]. The user has the possibility to both interactively enter the size of the dither matrix and populate its elements and thus to investigate how the size and the content of the dither matrix influence the quantity and the quality of the induced artifacts.

The second implemented NPR technique is screening interpreted as a separate halftoning technique with which two images - an original image and a dither screen, are combined using a special algorithm. The user has the possibility to use an arbitrary image as a dither screen but because there are special requirements for uniform distribution of threshold values and for homogeneous spatial distribution of threshold values the chosen image must be preprocessed.

At last, a special screening technique called screening with text is included. In this NPR technique letter shapes or texts are used as dither screens. The resultant images look like normal digital halftoning from a distance, but when viewed from close up they reveal the text being brought into the rendition by the dither screen used.

VI. Conclusion and Future Work

The results of the present work can be used for studying different NPR techniques with coding, for evaluation of the obtained non-photorealistic effects and for analyzing the program implementations. The main focus of the realization is to develop an extendable system of closely interconnected modules that give the programmer full control over the total process of loading image, performing needed preprocessing steps, applying selected image processing technique, viewing and storing the resultant images (visual evaluation), obtaining information concerning effectiveness of the applied algorithm (formal evaluation) and manage the history of the performed actions (analyze the results). The framework also provides a limited control over the already implemented algorithms through parameters and additional source data specified with a high level of interactivity.

In this stage the framework can be used for pixel-based non-photorealistic algorithms including such popular techniques as digital halftoning and screening. The implemented NPR techniques work over the content of images obtained through scanning, digital photography, or photo-realistic renderer. All algorithms operate on the level of pixels and have one single goal - to transform the intensity of the corresponding pixels in such a way to achieve various visual effects.

There are three main aspects for future work. The first is to extend already implemented screening with text technique in order to allow constructing dither screens from arbitrary contours (contour-based screening). The second is to implement such popular image-space NPR techniques as stippling and image mosaics. The third aspect is to extend the scope of the implemented NPR techniques (working in 2D image space) by adding algorithms that operate over the content of G-buffers (2 ½ D sources of data).

References

[1] T. Strothotte,S. Schlechtweg, "*Non-Photorealistic Computer Graphics: Modeling, Rendering, and Animation*" Morgan Kaufmann, 2002.

[2] B. Gooch and A. Gooch, "*Non-Photorealistic Rendering*" AK Peters Ltd, 2001.

[3] M. Eißele, D.Weiskopf, T.Ert, , "The G2-Buffer Framework" Institute of Visualization and Interactive Systems, University of Stuttgart unpublished.

[4] T. Saito and T. Takahashi, "Comprehensible Rendering of 3-D Shapes" Computer Graphics (Proceedings of ACM SIGGRAPH 90),volume 24, p. 197–206, 1990

[5] V. Dimitrova, "Obtaining G-buffers for Non-Photorealistic Rendering: Software Tools And Procedure", *Computer& Communications Engineering*, 2015, vol. 9, Technical University of Sofia

[6] A.Hertzmann, "*Algorithms for Rendering in Artistic Styles*" New York University, May 2011

[7] Alec Jacobson, "Dithering in MATLAB", http://www.alecjacobson.com/weblog/?p=1624

[8] MathWorks, MATLAB official web site, https://www.mathworks.com/products/matlab.html

[9] Meemoo application official web site, https://app.meemoo.org/#gist/3721129

[10] Ximagic Photoshop plugins web site, http://www.ximagic.com

[11] https://code.msdn.microsoft.com/windowsapps/Image-Distortion-Blur-673b1bdd

[12] Non photo-realistic rendering with REDsdk, http://www.redway3d.com/downloads/public/documentation/bk_re_non_photo_realistic_rendering.html

[13] Android: Image Processing, https://developer.android.com/reference/android/media/effect/EffectFactory.html

[14] Seven grayscale conversion algorithms (with pseudocode and VB6 source code), Tanner Helland, Oct 1, 2011, http://www.tannerhelland.com/3643/grayscale-image-algorithm-vb6/

[15] R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial greyscale*", Journal of the Society for Information Display*, vol. 17, no. 2, pp. 75–77, 1976

[16] B. E. Bayer, "An optimum method for two-level rendition of continuous-tone pictures" IEEE International Conference on Communications, vol. 1, June 11-13 1973, pp. 11–15.

Approach for knowledge transfer for building autonomous agent behaviour

Ansatz fur den Wissenstransfer fur den Aufbau eines autonomen Agentenverhaltens

Vanya Markova* and Ventseslav Shopov†

* Institute of Robotics with Bulgarian Academy of Sciences,
139 Rouski blvd, Plovdiv, 4000, Bulgaria, e-mail: markovavanya@yahoo.com
†139 Rouski blvd, Plovdiv, 4000, Bulgaria, e-mail:vkshopov@yahoo.com

Abstract — In this paper building of autonomous agent behaviour is discussed. We study an application of three different algorithms: greedy algorithm, reinforcement learning and reinforcement learning with knowledge transfer. We present two case studies that demonstrates the effectiveness of knowledge transfer approach. Finally we present that transfer learning is an essential mean of agent control learning task.

Zusammenfassung — In diesem Artikel wird der Aufbau des Verhaltens eines autonomen Agenten betrachtet. Wir erforschen die Anwendung von drei verschiedenen Algorithmen: Greedy-Algorithmus, bestärkendes Lernen und bestärkendes Lernen mit Wissenstransfer. Wir stellen zwei Fallstudien vor, die die Effektivität des Wissensaustauschansatzes demonstrieren. Schließlich stellen wir vor, dass Transferlernen ein wesentliches Mittel der Lernaufgabe der Agentenkontrolle ist.

I. Introduction

In this paper we present a new approach for transfer learning trough reinforcement learning.

An essential quality of a cognitive being is its ability to learn, that is, to gain new knowledge or skills as well as to improve existing knowledge or skills based on experience. Cognitive beings are able to cope with situations they have been previously confronted with as well as they are able to adapt to new situations sensibly. Thus, when designing an artificial agent that shall exhibit cognitive qualities—which we refer to in short as a cognitive agent—one central task is to develop computational means that enable the agent to learn. In this paper, we subscribe to one of the most influential paradigms of machine learning, reinforcement learning (RL) [1]. Reinforcement learning is very valuable when the characteristics of the underlying system are not known and/or difficult to describe or when the environment of an acting agent is only partially known or completely unknown.

Agents situated in the real world can perceive a great variety of information. Two situations are unlikely to lead to the same perception even if the situations are very similar. Learning requires the agent to acknowledge important details in a perception, to recognise commonalities across situations, and, most importantly, to disregard irrelevant information. In other words, the ability to learn involves the ability to abstract. Abstraction enables us to conceptualise the surrounding world, to build categories, and to derive reactions from these categories that can adapt to different situations. Complex and overly detailed circumstances can be reduced to much simpler concepts and not until then it becomes feasible to deliberate about conclusions to draw and actions to take. Abstract concepts explicate commonalities in different scenes and, thus, can cover new situations[2] [3] [4].

Transfer learning(TL) or inductive transfer is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.[5].

The main goal of this paper is to apply transfer learning trough reinforcement learning in area of building autonomous agent behaviour.

In second part of this paper we will describe the underlying theory of methods in our study: the RL/QL and MDP , Autonomous Agent behaviour and TL. In addition we will describe the implementation of our new approach for using TL in building of Autonomous Agent behaviour. In third part we describe the experiments and Gathers evidence to support our hypothesis.

II. Methods and materials

A. Autonomous Agent behaviour

The information about past and current states of the agent and environment allow the agents to estimate its own progress. Moreover this information allow the agent to make the corrections in existing pans if any needed and even to ma make new plans if it is necessary.

However if the agent makes corrections in the existing plans too often then this could lead to poor overall performance. Moreover the frequently dropping and building plans could make the things even worse. Hence it is desirable to reduce (or completely avoid) situations in which

the agent should changes its mind. There are two main approaches to do that: the first is to make the changing of the plans less recourse consuming task and the e second is to make such plans that are able to deal with volatile environment behaviour. This material is concerning the second approach.

As a key issue in building more efficient plans in rapidly changing environment we can point the ability of the agent to makes its plans in accordance not only with past and current states of the environment but also bearing in mind the future. To do that the agent needs to predict or forecast the future st states of the environment. So if we describe the stat es of the agent and environment as a time series then the task of making efficient plans will be significantly aided if the agent could forecasts the future with desirable accuracy.

An n-tipple (vector) is a result of one cycle of a work of the agent. It consists of the parameters of the behaviour of the agent: $b(b1, b2, \ldots, bn)$. The data from environment are collected and transformed into time series in the knowledge base of the agent.

B. Markov Decision Process

We formulate the transfer learning problem in sequential decision making domains using the following framework of Markov Decision Process (MDP) We use the following definition of MDP as a 5-tuple

$$< S, A, P, R, \gamma > \tag{1}$$

where the set of states, set of actions, transition function and reward function are described. And

$$P : S \times A \to \Pi(S) \tag{2}$$

is a transition function that maps the probability of moving to a new state given an action and the current state,

$$R : S \times A \to R \tag{3}$$

is a reward function. that gives the immediate reward of taking an action in a state.

And

$$\gamma \in [0, 1) \tag{4}$$

is the discount factor.

So the MDP of the agent is described in (1), where S is the set of states, A is the set of actions, P is transition function and R is a reward function. The transition function P maps the the probability of moving to a new state given an action and the current states is shown in (2). The reward functions R that gives the immediate reward of taking an action is described in (3). An the discount factor γ is bounded as is shown in (4).

C. Reinforcement learning

Reinforcement learning [1] is a popular and effective method to solve an MDP.

In our work we reference the reinforcement learning algorithm Q-learning as is described in [6]. At each moment of time, the agent is in a given state $s \in S$, and the agent's view is represented by a feature vector. Upon this information the agent makes the decision which action a from a set of all possible actions A to take in order to reach its goal. The outcome of Q-learning is a Q-function

$$Q(s, a) \tag{5}$$

that attaches to any state-action tuple (s, a) the expected reward over time. We discuss here the overall reward when starting in s and executing action a. From that Q-function, one can derive the policy π by always choosing the action with the highest Q-value:

$$\pi(s) = \underset{a \in A}{\operatorname{argmax}}(Q(s, a)) \tag{6}$$

Under these conditions, Q-learning should converge to an optimal Q-function

$$Q* = \pi(s) = \underset{a \in A}{\operatorname{argmax}}(Q(s; a)) \tag{7}$$

that returns the highest reward for any state-action tuple (s, a). Hence in this way we establish an optimal policy $\pi*$.

D. Transfer learning

Machine learning and data mining techniques have been used in numerous real-world applications. An assumption of traditional machine learning methodologies is the training data and testing data are taken from the same domain, such that the input feature space and data distribution characteristics are the same. However, in some real-world machine learning scenarios, this assumption does not hold. There are cases where training data is expensive or difficult to collect. Therefore, there is a need to create high-performance learners trained with more easily obtained data from different domains. This methodology is referred to as transfer learning.[7]

There is a hierarchical Bayesian framework for transfer in sequential decision making tasks of transferring two basic kinds of knowledge [8] [9]

In our paper we uses meta-data (e.g., attribute-value pairs) associated with each task to learn the expected benefit of transfer given a source-target task pair. An example of such a metadata is given in [10].

E. Implementation

In our implementation of QL, we claim that if the values of the P matrix at the beginning of the training are zero, then we will reach an optimal policy for a final number of epochs (steps). In order to speed up the training, it is good that the coefficients of the P matrix are somewhat closer to the desired policy. This can be achieved through a TL in a simpler environment (or just a part of the environment). The classic reinforcement learning consists of finding an optimal policy for the whole area with high details.

Our approach is based on [10]:

- loading the whole map and scraping all details but geometric obstacles
- find a reinforcement learning solution for this plain map
- transfer the Q matrix as predefined knowledge
- load full map and using predefined knowledge enhance the learning process

Notation and transfer learning: Let G be the set of all possible tasks. Let $G_{source} \subset G$ be a set of source tasks for which the agent has already learned a policy and let $G_{target} \subset G$ be another set of target tasks to be learned by the agent. For each task

$$g_i \in G, letD_i inR^n \tag{8}$$

is a descriptor of features for the given task.We assume that g_i and D_i that are known to the agent.

We define a target task $g_j \in G_{target}$, as the goal of the agent. Thereafter to achieve g_j we have to select a task $g_i \in G_{source}$ such that g_i serves as an effective source for learning g_j . For task pair, g_i and g_j , let

$$f_u(g_i, g_j) \in R^n \tag{9}$$

denote the function of usefulness. This is the function of transferring the policy learned in g_i to the task g_j , where $f_u(g_i, g_j) > 0$ indicates positive benefit of transfer, while $f_u(g_i, g_j) < 0$ indicates negative benefit of transfer.

We assume that for each pair of tasks (g_i, g_j) such that $g_i, g_j \in G_{source}$, the agent could reliable estimate $f_u(g_i, g_j)$. So the agent can use these estimates to predict the expected transfer benefit between tasks in G_{source} and tasks in G_{target}.

III. Experiments and results

We Gathers evidence to support the hypothesis that building of Autonomous Agent behaviour trough transfer learning will significantly speed up the process of constructing Autonomous Agent behaviour. We claim that building of Autonomous Agent behaviour trough transfer learning is more efficient that applying the RL direct approach.

We perform the following experiment: For a given map we should find an optimal autonomous agent behaviour. The map is described by its size nxn and complexity rate Rc. And we chose a start point and final target in this map. The goal is to find the fastest route from start point to final target.

We have three methods Greedy Approach, Reinforcement learning without knowledge transfer and Reinforcement learning whit knowledge transfer. And two cases: Plain map and rough terrain map. Plain map have all terrain with same squares that cost 1 clique and obstacles have inf cliques. In the rough terrain map the squares could be 1,2,4 or inf cliques each.

We study following algorithms:

- case I - a Greedy Approach(GA)
- case II - Reinforcement Learning without knowledge transfer(RL)
- case III -Reinforcement Learning with Transfer Learning(RLTL)

We do the following task: for a given map we need to find optimal behaviour of an autonomous agent. The agent's task is to travel on a closed route for a minimum of time. The environment is represented as a two-dimensional obstacle map. The map is described by its size nxn and the

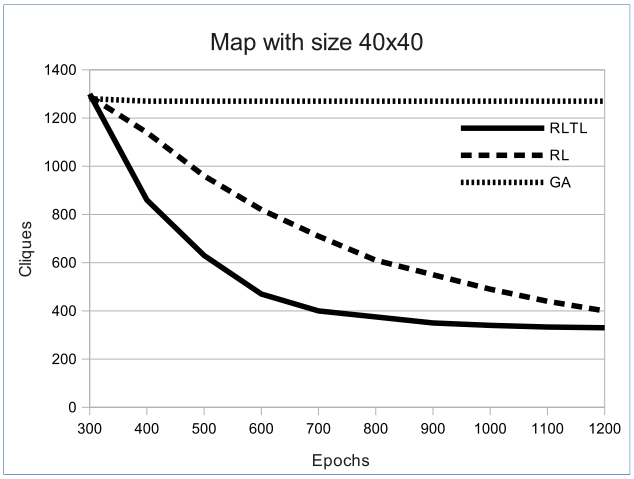


Fig. 1. Comparison of three approaches: Greedy Algorithm(GA), Reinforcement Learning without Transfer Learning(RL) and Reinforcement learning with transfer learning(RLTL) om map with size 40x40.

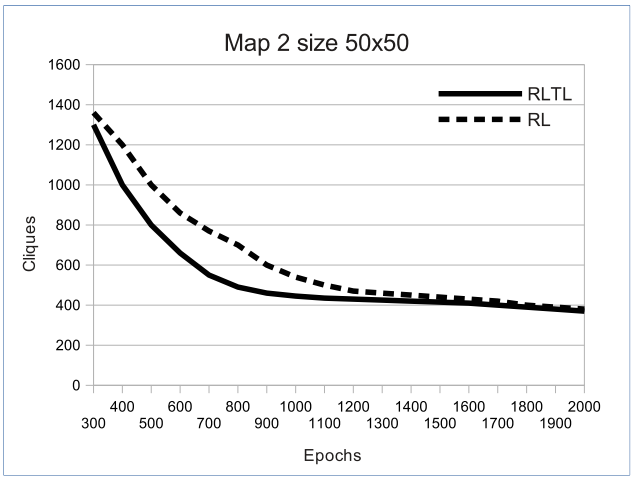


Fig. 2. Compare two algorithms: Reinforcement Learning without Transfer Learning(RL) and Reinforcement learning with transfer learning(RLTL) on map 2 with size 50x50.

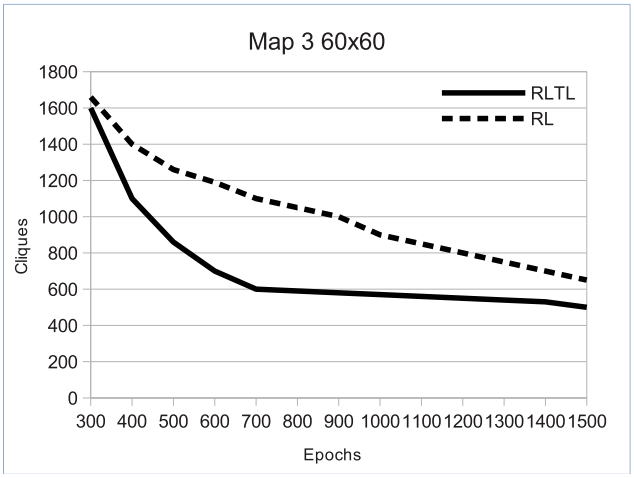


Fig. 3. Compare two algorithms: Reinforcement Learning without Transfer Learning(RL) and Reinforcement learning with transfer learning(RLTL) on map 3 with size 60x60.

rate of complexity Rc. And we chose the starting and ending points of the route on this map. The goal is to find the fastest way from the starting point to the end goal.

We discuss two cases: A simple regular 2D map and a rough 2D map of the pitch. The regular map has all the terrains with the same squares that cost 1 click and the obstacles have inf clicks. In the rough map of the terrain the squares can take 1,2,4 or inf cliques respectively. The fastest its the route with minimum cliques. Agents do not know in advance how many cliques each point of the map takes. So the environment is partially observable. We have three algorithms: Greedy Approach , Reinforcement learning without knowledge transfer(RL) and Reinforcement learning with transfer learning(RLTL) The goal is to achieve optimal policy: the route which takes minimum cliques from the starting point to the end goal. And we perform up to 2000 epochs to finding an optimal policy. As an assessment measure, we use the number of clicks to reach the goal of the final policy. We use three 2D maps: map 1 - 40x40, map 2 - 50x50 and map 3 - 60x60. All maps have random placed obstacles. Moreover all maps have random placed rough points. The distribution of the rough points however for all maps is the same.

In the first experiment, we compare the three approaches on the first map : GA, RL, RLTL. GA does not improve with time because it does not take into account the unobserved properties of the terrain. The second and third experiments show that the RLTL starts much faster and converges to the optimal policy earlier.

Three different algorithms are presented: Greedy Approach, Reinforcement Learning and Reinforcement Learning with Transfer Learning. In two case studies we demonstrate that knowledge transfer could make more effective the learning part of building the behaviour of autonomous agent. From 1 one can see that Greedy Approach does not improve over time. Figures 2 and 3 shows that Reinforcement learning with transfer learning converge much faster to optimal policy than Reinforcement Learning without Transfer Learning. The RLTL and RL are compared on three 2D maps: map 1 - 40x40, map 2 - 50x50 and map 3 - 60x60.

IV. Conclusion

The impact of different approaches for building of autonomous agent behaviour is discussed in this paper. Three different algorithms are presented. greedy algorithm, reinforcement learning and reinforcement learning with knowledge transfer. In two case studies we demonstrate that transfer learning could make more effective the learning part of building the behaviour of autonomous agent. The agent is able to adapt to new scenarios trough transferring knowledge from simplified environments to more complex maps with rough terrain. Finally we present that transfer learning is an essential mean of agent control learning task.

References

[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.

[2] V. Markova, "Adaptive behaviour approach for autonomous mobile sensor agent," in *Proc. Int'l Conference InfoTech-2012*, 2012, pp. 1314–1023.

[3] ——, "Autonomous agent design based on jade framework," in *Proceedings of the International Conference on Information Technologies (InfoTech-2013), no*, 2013, pp. 19–20.

[4] V. Markova and V. Shopov, "An approach to build fuzzy cognitive map for time series," in *Proc. of the Int'l Conference InfoTech*, 2014, pp. 207–211.

[5] D. Ventura and S. Warnick, "A theoretical foundation for inductive transfer," *Brigham Young University, College of Physical and Mathematical Sciences*, 2007.

[6] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Cambridge, 1989.

[7] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, p. 9, 2016.

[8] A. Wilson, A. Fern, and P. Tadepalli, "Transfer learning in sequential decision problems: A hierarchical bayesian approach," in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, 2012, pp. 217–227.

[9] A. Wilson, A. Fern, S. Ray, and P. Tadepalli, "Multi-task reinforcement learning: a hierarchical bayesian approach," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 1015–1022.

[10] J. Sinapov, S. Narvekar, M. Leonetti, and P. Stone, "Learning inter-task transferability in the absence of target task samples," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 725–733.

Sensor data fusion for topological localization of rail vehicles

Sensordatenfusion zur topologischen Ortung von Schienenfahrzeugen

Stefan Hensel*, Marin B. Marinov‡

* University of Applied Sciences Offenburg, Department for Electrical Engineering,
Badstraße 24, D-77652 Offenburg, Germany, stefan.hensel@hs-offenburg.de
‡ Technical University of Sofia, Faculty of Electronic Engineering and Technologies, Department of Electronics,
8, Kliment Ohridski Blvd., BG-1756 Sofia, Bulgaria, mbm@tu-sofia.bg

***Abstract* —**This contribution presents a novel onboard localization system, based on an eddy current sensor system (ECS) that is capable of a precise train localization when combined with a simple topological map enriched with ECS features. To determine the position upon the track, we propose an approach based on counting in contrast to commonly applied travel distance determination by integrating the estimated velocity. ECS features are derived from the signal by means of specially defined virtual sensors and subsequently processed within a Bayesian formalism. This virtual sensor based approach allows parallelization of sensor signal processing and flexible integration of new sensors into the positioning system. The possibility to detect turnouts with a hit rate of 99% allows the vehicle position to be tracked over the entire track, using the measurement data supplied by the ECS alone.

***Zusammenfassung* —** Die in dieser Arbeit vorgestellte Vorgehensweise erlaubt die Ortung von Schienenfahrzeugen in topologischen Karten allein mit Hilfe eines Wirbelstromsensorsystems (WSS). Zur Ortung primär erforderlich ist die Identifizierung des befahrenen Gleises selbst, wofür unterschiedliche in einer Karte gespeicherte Merkmale herangezogen werden sowie der zurückgelegte Weg, der durch Zählen der passierten Schwellen ermittelt wird. Diese Merkmale werden mittels eigens definierter, virtueller Sensoren aus dem Signal des WSS gewonnen und mittels einem Bayes'schen Formalismus mit den Referenzdaten aus der vorliegenden topologischen Karte abgeglichen. Diese auf virtuellen Sensoren basierende Vorgehensweise erlaubt eine Parallelisierung der Sensorsignalverarbeitung und eine flexible Einbindung von Sensoren in das Ortungssystem. Die Möglichkeit, Weichen mit einer Trefferquote von 99% zu detektieren, erlaubt die Verfolgung der Fahrzeugposition über die gesamte Fahrstrecke hinweg, unter alleiniger Verwendung der vom WSS gelieferten Messdaten.

I. Einführung

Die Ortung von Schienenfahrzeugen spielt eine wichtige Rolle im modernen Bahnbetrieb. Sie ist unerlässlich für Fahrgastinformationssysteme und die Betriebsleittechnik. Die dafür eingesetzten Verfahren stützen sich meist auf das Auslesen ortsfester Markierungen (Baken und Balisen), auf bordeigene Sensorik oder auf den Einsatz von Satellitennavigation (GNSS). Ortsfeste Markierungen müssen jedoch kostspielig auf der Strecke installiert werden und benötigen die entsprechenden Sensoren am Fahrzeug. Die Ortung mittels GNSS wird oftmals durch Verdeckungen und Störungen beeinträchtigt und bietet nicht die für eine *gleisselektive* Ortung erforderliche Genauigkeit und Zuverlässigkeit. Auch klassische Fahrzeugsensoren, wie beispielsweise der Radumdrehungszähler oder Mikrowellenradare, unterliegen in der Regel den durch Schlupf oder Drift verursachten Fehlern, besonders wenn sie den zurückgelegten Weg integrativ aus Geschwindigkeitswerten berechnen [1].

Der hier vorgestellte neue Ansatz nutzt ein auf dem Wirbelstromprinzip basierendes Sensorsystem, das metallische Bauteile in der Nähe der Schiene berührungslos detektiert. Mittels geeigneter Signalverarbeitung kann damit schlupffrei und witterungsunabhängig die Geschwindigkeit des Fahrzeugs präzise ermittelt werden. Im Rahmen früherer Arbeiten wurde bereits gezeigt, dass sich dieses Sensorsystem auch hervorragend zur bordautonomen Ortung eignet. Dabei wurde ausgenutzt, dass das Sensorsignal beim Befahren von Weichen eine spezielle Charakteristik besitzt, mit der Weichen nicht nur erkannt (detektiert), sondern auch eindeutig identifiziert (klassifiziert) werden können. Da durch die Weiche und ihre Befahrungsrichtung die Position im Schienennetz jederzeit bekannt ist, ermöglicht dies die Lokalisierung des Fahrzeuges ohne zusätzliche Sensorik oder anderweitige ortsfeste Markierungen.

Im Rahmen dieser Arbeit soll nun ausgenutzt werden, dass auch die Befestigungen der Schienen an den Schwellen im Signal des Wirbelstromsensors sichtbar sind. Da sie im gesamten Gleisverlauf in sehr regelmäßigen Abständen verbaut sind, lässt sich die vom Fahrzeug zurückgelegte Strecke auch durch das Zählen dieser Schienenbefestigungen ermitteln. Diese Art der Wegmessung verspricht eine hohe Genauigkeit im Vergleich zu geschwindigkeitsbasierten Methoden, in denen immer die Kumulierung von Messfehlern durch Integration auftritt. Im weiteren Verlauf werden die Grundlagen der Ortung mit Hilfe der Weichenerkennung auf komplette Gleisabschnitte übertragen, so dass nicht nur die seit der letzten Weiche zurückgelegte Strecke gemessen, sondern auch das befahrene Gleis genau identifiziert wird [2].

Für die Weichenerkennung, das Zählen der Schienenbefestigungen und für weitere Merkmale des Gleises, welche im komplexen Sensorsignal enthalten sind (wie z.B. Signalleistung oder mittlerer Schwellenabstand) werden *virtuelle Sensoren* definiert. Dieser Ansatz erlaubt eine Parallelisierung und Modularisierung der Sensorsignalverarbeitung des WSS und erlaubt das einfache Einbinden weiterer, virtueller oder realer, Sensoren. Das gesamte Ortungssystem wird dadurch flexibel und robust gegenüber Ausfällen und Schwankungen der Bandbreite.

Grundlage für den Ortungsprozess ist die Weichenerkennung (Virtueller Sensor 1 – VS_1). Für die Erhöhung der Ortungsgenauigkeit können nun zusätzliche virtuelle Sensoren

definiert und ergänzt werden $(VS_2, \ldots, VS_n)$. Das Gesamtkonzept ist in Abb. 1 dargestellt.

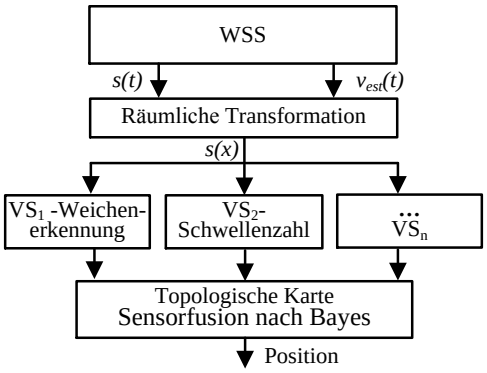


Abb. 1. Gesamtkonzept der Positionsbestimmung mit virtuellen Sensoren.

II. Ortung mit Wirbelstromsensoren

Beim Einsatz des WSS zur Positionsbestimmung und Geschwindigkeitsmessung werden die Eigenschaften des Gleises, auf dem sich das Fahrzeug befindet, genutzt. Die konventionelle Anwendung des Sensorsystems liegt zunächst im Messen der Geschwindigkeit, aus der sich durch Integration bereits der zurückgelegte Weg bestimmen lässt. Zusätzlich lassen sich ortstypische Merkmale des befahrenen Gleises bestimmen und direkt zur Ortung heranziehen.

A. Signalverarbeitung

Der Wirbelstromsensor registriert metallische Inhomogenitäten im Schienenbereich. Aufgrund des relativ großen Abstandes zur Schiene $(100\,mm)$ sind dies keine Unregelmäßigkeiten im Material der Schiene selbst, vielmehr aber metallische Bauteile, die entlang der Schiene verbaut sind.

Das real vom Sensor gelieferte Signal hat die in Abb. 2 gezeigte Form. Der sinusförmige Signalanteil aus den Schienenbefestigungen ist gut zu erkennen, allerdings ist er verzerrt und von höherfrequenten Störungen überlagert.

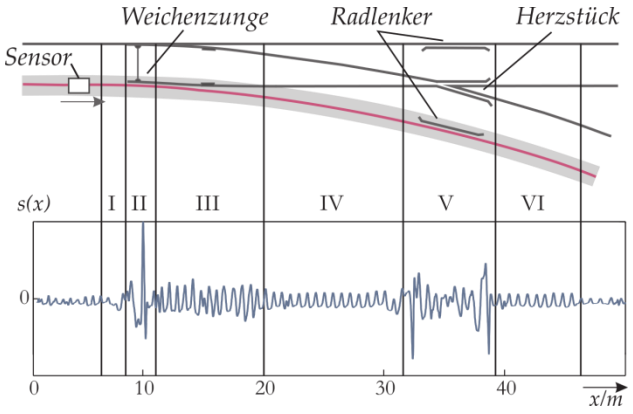


Abb. 2. Sensorsignal $s(x)$ beim Befahren von Weichenkomponenten (Bereiche II und V) und von Gleisabschnitt nur mit Schienenbefestigungen (Bereich III und IV).

Gänzlich anders sieht der Signalverlauf beim Überfahren einer Weiche aus. Hier dominieren nichtperiodische Signale höherer Signalenergie, die von den massiven und näher an den Sensor heranreichenden Weichenbauteile her rühren, wie ebenfalls in Abb. 2 zu erkennen.

Für die Bearbeitung der Signale erfolgt die Filterung mit einem adaptiven Bandpassfilter, dessen Eckfrequenzen von der aktuellen Geschwindigkeit abhängig sind, für die Unterdrückung hochfrequenter Störungen sowie potentiell vorhandener Gleichanteile.

Für die weitere Verarbeitung des Signals ist es hilfreich, wenn es im Ortsbereich als $s(x)$ vorliegt. Für die Transformation wird der Schätzwert der momentanen Geschwindigkeit genutzt [2]. Der Zusammenhang zwischen $s(t)$ und $s(x)$ besteht dann über ein Zeit-Weg-Diagramm, wobei jedem Abtastwert in $s(x)$ ein Wert in $s(t)$ zugeordnet werden kann. Eine der Sensorbandbreite angemessene Auflösung im Ortsbereich liegt bei 50 Abtastwerten pro Meter, die im weiteren Verlauf dieser Arbeit betrachteten Signale haben eine Auflösung von 2 cm pro Sample.

B. Topologische Karten

Da das WSS keinerlei Informationen über die Geometrie oder globale Position des Gleises liefert, findet die Ortung in einer topologischen Karte statt. Diese stellt sämtliche Gleise und Weichen des Schienennetzes in ihrer gegenseitigen Lage und Verknüpfung dar, enthält jedoch keine Informationen über den geometrischen Verlauf der einzelnen Gleise. Die Karte kann daher durch ein einfaches Knoten-Kanten-Modell in einen Graphen überführt werden, in welchem die Weichen als Verzweigungspunkte durch die Knoten und die Gleise als Kanten dargestellt werden.

Zur praktischen Durchführung einer Ortung werden sinnvolle Referenzpunkte benötigt. Weichen bieten sich hier besonders an, da sie als Knoten schon in der topologischen Karte vorhanden sind und als Verzweigungspunkte die Freiheitsgrade der Karte repräsentieren. In Abb. 3 sind schematisch die Elemente der hier verwendeten topologischen Karten sowie die Abstraktion der Karte als Graph aus vorhandenen Gleisplänen gezeigt.

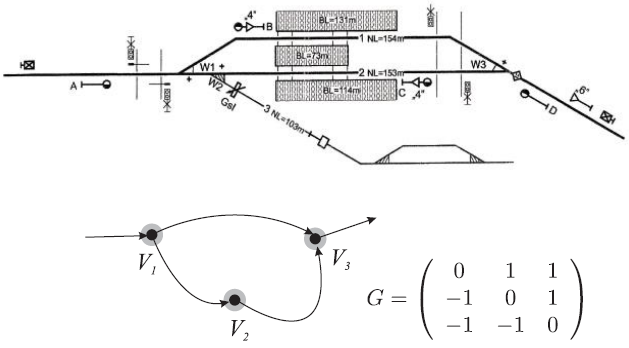


Abb. 3. Grundelemente der topologischen Karte (Knoten V_i und Adjazenzmatrix G) und Gewinnung der Karte aus gegebenen Gleisplänen.

Da die Weichen eine räumliche Ausdehnung besitzen, aber auf einen Punkt abgebildet werden müssen, wird als Knotenpunkt stets der Weichenanfang genutzt. Die Länge der Weichen wird dem Gleis, das die Weichenanfänge verbindet, zugeschlagen.

Die Gleise werden in der Karte als Kanten dargestellt. Da in der topologischen Karte keine geometrischen Informationen über den Gleisverlauf abgelegt werden, ist das charakteristische Merkmal eines Gleises zunächst seine Länge. Da die Wegmessung in Metern über die Integration der Geschwindigkeit jedoch fehleranfällig ist, dient als Abstandsmaß die Anzahl der Schwellen beziehungsweise Schienenbefestigungen, die jedem Gleis als konstante Größe zugewiesen werden.

III. Virtuelle Sensoren

Bei dem hier betrachteten Verfahren liefert das über der Schiene angebrachte WSS die Daten. Das gewonnene Signal beinhaltet eine Anzahl von Merkmalen, die jeweils gesondert untersucht werden können. Eines davon ist die Anzahl der

Schienenbefestigungen, die für die Bestimmung der aktuellen Position und der Länge eines Gleises genutzt werden. Um nun alleine durch Betrachten des Wirbelstromsensorsignals zu bestimmen welches Gleis gerade befahren wird, bietet es sich an, zusätzliche Eigenschaften des Signals heranzuziehen. Es wird nach Merkmalen gesucht, die für jedes Gleis möglichst charakteristisch sind. Für jedes Merkmal wird nun ein virtueller Sensor definiert, dessen Messungen E^k das jeweilige Merkmal repräsentieren.

A. Leistungssignatur

Ein sehr mächtiges Werkzeug zur Erkennung des aktuellen Gleises ist die Betrachtung der Signalleistung

$$S(x) = \frac{1}{T}\int_{x-T/2}^{x+T/2} |s(t)|^2 \tag{1}$$

des vom Wirbelstromsensor aufgezeichneten Signals. Ihr Verlauf ist für jedes Gleis charakteristisch und behält auch bei verschiedenen Befahrungen stets dieselbe Form. Dieser Verlauf kann als Signatur für jedes Gleis angesehen werden. Zur weiteren Verarbeitung wird der Mittelwert der Leistung im Ortsbereich über 25 Samples ($\cong 50cm$) gebildet. Anschließend erfolgt ein Vergleich mit der in der Karte als Referenz hinterlegten Leistungssignatur über die Bildung des euklidischen Abstandes

$$D_{sr} = \|s - r\| = \sqrt{\sum_m^{25} |s_m - r_m|^2}. \tag{2}$$

B. Schwellenabstand

Ein weiteres charakteristisches Merkmal ist der Abstand der Schwellen. Im Mittel ist dieser bei automatisiert verlegten Schwellen $64\ cm$, unterliegt jedoch Schwankungen, die unter anderem durch die Signalentzerrung mit der geschätzten Geschwindigkeit beeinflusst wird, so dass sich Geschwindigkeitsfehler auf die Genauigkeit der ermittelten Abstände auswirkt. Aus diesem Grund wird der Schwellenabstand als Mittelwert über einen Abschnitt von 25 Schwellen angegeben, was auch entsprechend in der Karte hinterlegt ist.

Der virtuelle Sensor ermittelt aus dem aktuellen Signal des Wirbelstromsensors nun den aktuellen Mittelwert der Schwellenabstände und vergleicht ihn mit den abgelegten Referenzen für jedes Gleis. Diese Messung ist mit relativ großen Unsicherheiten behaftet. Daher wurde die resultierende Likelihood auf einen Wertebereich von 0,6 bis 0,8 begrenzt.

Die Abweichung wird mit

$$D_{sr,i} = |\bar{x}_n| - |\bar{x}_{nr,i}| \tag{3}$$

berechnet, wobei $\bar{x}_n$ der gemessene mittlere Schwellenabstand und $\bar{x}_{nr,i}$ die zur Gleishypothese H_i gehörige Referenz ist. Für die Likelihood $P(E^k|H_i)$ gilt somit

$$P(E^k|H_i) = \begin{cases} -0{,}04 D_{sr,i} + 0{,}8 & D_{sr,i} < 5 \\ 0{,}6 & D_{sr,i} \geq 5 \end{cases}. \tag{4}$$

C. Schwellenzahl

Die Schwellenzahl eines Gleises ist proportional zu dessen Länge, so dass das durch Zählen der Schienenbefestigungen ein sehr starkes Merkmal für einen Gleisabschnitt gewonnen werden kann. Durch Schwankungen in der Detektion des Weichenanfanges ergibt sich für mehrere Bewährungen desselben Gleises eine jeweils unterschiedliche Anzahl an gezählten Schwellen. Aus diesen Messwerten wird als Referenzwert für jedes Gleis der Mittelwert μ und die Standardabweichung σ als Referenzwert in der Karte abgelegt.

Der virtuelle Sensor für die Bewertung der Gleishypothesen wird durch eine geeignete Umrechnung der aktuellen Position in Schwellen in eine zugehörige Likelihood erzeugt. Die Kenntnis von μ_i und σ_i für jedes H_i genügt, um zu erkennen, ob die seit der letzten Weiche zurückgelegte Distanz für den Aufenthalt auf einem Gleis *i* spricht. Die aus diesem Ansatz abgeleitete Likelihood-Funktion des virtuellen Sensors wird daher durch

$$P(E^k|H_i) = \begin{cases} 1 & n_i < \mu_i \\ \exp\left(-\frac{1}{2}\left(\frac{n_i - \mu_i}{\sigma_i}\right)^2\right) & n_i \geq \mu_i \end{cases} \tag{5}$$

modelliert. Diese Funktion beschreibt eine konstante Wahrscheinlichkeit bis zum Erreichen von μ_i, danach erfolgt ein Abklingen in Abhängigkeit σ_i.

IV. Sensorfusion nach Bayes

Die Bayes'sche Sensorfusion bietet eine elegante Möglichkeit, Informationen von Sensoren zusammenzufassen, die nach unterschiedlichen Prinzipien arbeiten [3]. Mit ihr lässt sich angeben, mit welcher Wahrscheinlichkeit $P(H_i|E)$ eine Hypothese H_i erfüllt ist, wenn das Ereignis E aufgetreten ist. In diesem Falle also die Wahrscheinlichkeit, dass sich das Fahrzeug auf einem Gleis *i* befindet, wenn die Beobachtung E eines virtuellen Sensors vorliegt.

Wie in Kapitel III beschrieben, liefert jeder Sensor für dasselbe Signal eine Likelihood $P(E|H_i)$ dafür, wie die beobachtete Messung E einer bestimmten Hypothese H_i zugeordnet werden kann. Die a-posteriori-Wahrscheinlichkeit $P(H_i|E)$ berechnet sich wie folgt:

$$P(H_i|E) = \frac{P(E|H_i)}{P(E)} = \frac{P(E|H_i)P(H_i)}{\sum_j P(E|H_j)\,P(H_j)} \propto P(E|H_i)P(H_i) \tag{6}$$

Hierbei gibt die a-priori-Wahrscheinlichkeit $P(H_i)$ an, mit welchen Wahrscheinlichkeiten die Hypothesen H_i im Vorfeld erfüllt sind. Ohne Vorwissen kann hier eine Gleichverteilung angesetzt werden, die allen Hypothesen dieselbe Wahrscheinlichkeit zuweist.

A. Rekursives Bayes'sches Update

Zur fortlaufenden Berechnung der a-posteriori-Wahrscheinlichkeit wird als a-priori-Wahrscheinlichkeit die a-posteriori-Wahrscheinlichkeit des letzten Berechnungsschrittes verwendet. In (7) beschreibt H_i wie gehabt die Hypothesen, der Vektor E^N stellt vergangene Messergebnisse des Sensors dar, E ist die aktuelle Messung. Sobald $P(H_i|E^N)$ berechnet ist, wird die Auswirkung der aktuellen Messung mit

$$P(H_i|E^N, E) = \frac{P(E|E^N, H_i)P(H_i|E^N)}{P(E|E^N)} \tag{7}$$

beschrieben. Die Wahrscheinlichkeit $P(H_i|E^N)$ übernimmt hier die Rolle der a-priori-Wahrscheinlichkeit. Sie repräsentiert die vergangenen Messungen Wird die Messung E als unabhängig von den vorangegangenen Messungen E^N angenommen werden, vereinfacht sich der Likelihood-Term in (7) weiter zu

$$P(E|E^N, H_i) = P(E|H_i). \tag{8}$$

B. Berechnung der A-Posteriori-Wahrscheinlichkeit

Angenommen es werden q Gleise verfolgt, wovon jedes als eine Hypothese H_i dargestellt wird: $H = \{H_1, H_2, \ldots, H_q\}$. Dann gilt für den Likelihood-Vektor λ^kdes virtuellen Sensors *k*

$$\lambda^k = \left(\lambda_1^k,\ \lambda_2^k, \ldots, \lambda_q^k\right), \tag{9}$$

und $\lambda_1^k = P(E^k|H_1)$. Unter der Voraussetzung, dass die Hypothesen H_i unabhängig voneinander sind, werden die Likelihood-Vektoren der einzelnen Sensoren nun elementweise miteinander multipliziert, um den Likelihood-Vektor für H_i über alle *K* Sensoren $\Lambda_i = \lambda_i^1, , \dots, \lambda_i^K$ wie folgt anzugeben:

$$\Lambda_i = \prod_{k=1}^{K} P(E^k|H_i). \quad (10)$$

Für die a-posteriori-Wahrscheinlichkeit, gilt dann

$$P(H_i|E^1, \dots, E^K) = \propto P(H_i) \cdot \Lambda_i. \quad (11)$$

C. Aufstellen und Verfolgen der Hypothesen

Neben den virtuellen Sensoren, die für jeden Berechnungsschritt der Bayes'schen Sensorfusion die Likelihood-Vektoren liefern, ist die initiale Bestimmung der a-priori-Wahrscheinlichkeit und dadurch die Auswahl der zu verfolgenden Gleishypothesen von zentraler Bedeutung. Die Hypothesen bilden die Auswahl an Gleisen, auf denen die momentane Position des Fahrzeuges vermutet wird.

V. Ergebnisse

Mit den vorgestellten virtuellen Sensoren und der Bayes'schen Sensorfusion wurden Berechnungen anhand der aufgezeichneten Messdaten von insgesamt 20 Befahrungen einer realen Teststrecke durchgeführt. Diese sollen die unter realen Bedingungen erzielbaren Ergebnisse bei der Ortung von Schienenfahrzeugen mit dem vorgestellten Verfahren aufzeigen. Dabei werden zunächst die Ergebnisse der Sensoren einzeln vorgestellt, indem die bei der Fahrt auf einem bestimmten Gleis erzeugten Likelihood-Vektoren dargestellt werden. Anschließend werden die Ergebnisse der Sensoren mit der Sensorfusion kombiniert, um dann die Verfolgung mehrerer Hypothesen entlang einer kurzen Strecke durch einen kleinen Bahnhof zu zeigen.

Betrachtung der Schwellenabstände

Aufgrund annähernd gleichen Abständen im gesamten Netz liefert der virtuelle Sensor für den Schwellenabstand den geringsten Beitrag zur Gleiserkennung.

Betrachtung der Schwellenzahl

Ein gutes Kriterium, um Hypothesen zu Gleisen auszuschließen, ist die Anzahl der gezählten Schwellen. Wurden im Verlauf des aktuellen Gleises mehr Schwellen gezählt als zu einer aktiven Hypothese gehören, so kann diese mit einer sehr kleinen Wahrscheinlichkeit versehen werden.

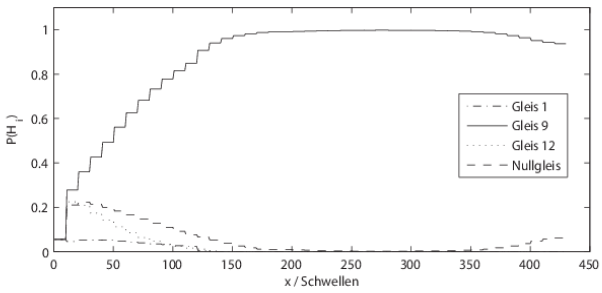


Abb. 4. Verlauf der a-posteriori-Wahrscheinlichkeit $P(H_i)$ unter ausschließlicher Betrachtung des Signaturabstands.

Signaturvergleich

Da die Weichen sich in den Leistungssignaturen wiederfinden, entspricht dieser virtuelle Sensor einen Korrelationsansatz und ermöglicht aufgrund des Vorhandenseins einer Weiche bereits wenige Meter hinter dem Kartenknoten eine verlässliche Aussage über das befahrene Gleis. In Abb. 4 ist der Verlauf von $P(H_i)$ bei Fahrt über Gleis 9 mit Hypothesen für zwei weitere, hier lokal nahe, Gleise (1 und 12) dargestellt.

Verfolgung der Position mit Hilfe der Hypothesen

Die Auswertung der virtuellen Sensoren und die Fusion ihrer Messwerte mit der Bayes'schen Sensorfusion erlaubt ebenfalls das Verfolgen der Fahrzeugposition entlang eines Fahrweges. Hierzu wird zunächst eine bekannte Startposition an einer Weiche vorausgesetzt. Nun werden die in Fahrtrichtung an die Weiche anschließenden Gleise als Hypothesen aufgestellt und mit den von nun an gemessenen Sensordaten aktualisiert. Abb.5 zeigt, wie bereits nach wenigen Metern die korrekte Hypothese gefunden und beibehalten wird.

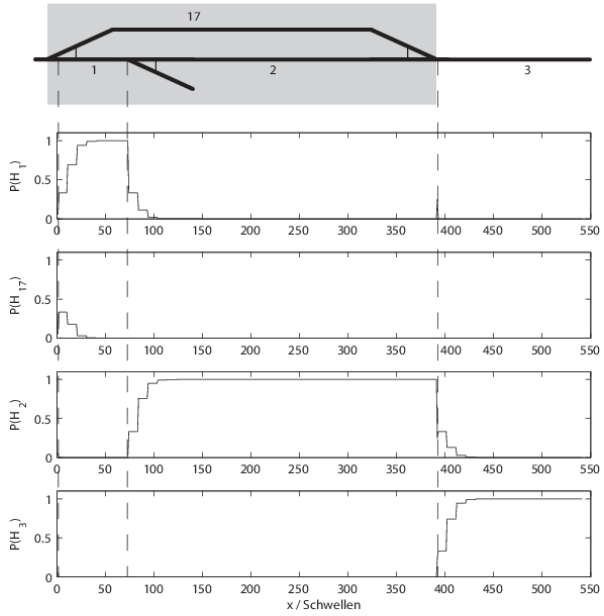


Abb. 5. Verlauf der A-posteriori-Wahrscheinlichkeit $P(H_i)$ bei der Befahrung eines Bahnhofes.

Die zurückgelegte Strecke wird ebenfalls über die gezählten Schwellen und somit driftfrei bestimmt.

VI. Zusammenfassung

Der Aufsatz zeigt, wie allein mit Hilfe des WSS Signals die zurückgelegte Distanz sowie das aktuell befahrene Gleis innerhalb einer topologischen Karte bestimmt werden können. Die Erweiterung der Karte mit gleisspezifischen Signalcharakteristika erlaubt den Einsatz virtueller Sensoren, die beliebig mit realen Sensoren kombiniert werden können. Der rekursive Ansatz der Bayes'schen Sensorfusion erlaubt zudem eine Positionsbestimmung zur Laufzeit, deren einzelne Hypothesen mit quantifizierbaren Wahrscheinlichkeiten weiteren Verarbeitungsmodulen zur Verfügung stehen. Auf Basis einer Weichendetektion von 99 % ist eine robuste und präzise Positionsbestimmung möglich, was anhand realer Messdaten aufgezeigt wurde.

Literaturverzeichnis

[1] Hensel, S., Hasberg, C., Stiller, C., „Probabilistic Rail Vehicle Localization with Eddy Current Sensors in Topological Maps," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, Nr. 4, p. 1–13, 2011.

[2] Hensel, S., Strauss, T., Marinov, M., „Eddy current sensor based velocity and distance estimation in rail vehicles," *IET Science, Measurement & Technology*, Bd. 9, Nr. 7, p. 875–888, 2015.

[3] L. Klein, Sensor and Data Fusion, Washington: SPIE-International Society for Optical Engineering,, 2004.

[4] G. Denoix, „Topologische Ortung von Schienenfahrzeugen, Diplomarbeit," Universität Karlsruhe, 2009.

German as a foreign language and the mastery of relevant technical language as prerequisites for a successful engineering education at FDIBA

Die deutsche Sprache als Fremdsprache und die fachbezogene Beherrschung als Voraussetzung für eine erfolgreiche Ingenieurausbildung an der FDIBA

Julia Popowa*, Lucas Pfeiffer†
*FDIBA, TU Sofia, Bulgarien, Allgemeiner Maschinenbau
†FDIBA, TU Sofia,DAAD-Sprachassistent im Studienjahr 2016/17, Deutsch als Fremdsprache

Abstract **—** This presentation gives an account of three lessons from the summer semester of 2017, which successfully integrate technical language into the general German language course of FDIBA, in an attempt to address difficulties that students might have with technical terminology.

Zusammenfassung **—** In diesem Vortrag werden drei Unterrichtsversuche aus dem Sommersemester 2017 vorgestellt, in denen Fachsprachenunterricht gewinnbringend in den allgemeinen Deutschunterricht an der FDIBA integriert werden konnte, und so möglichen Sprachhemmnisse von Studierenden im Umgang mit Fachterminologie begegnet werden konnte.

Seit der Gründung der Fakultät für deutsche Ingenieur- und Betriebswirtschaftsausbildung (FDIBA) an der TU-Sofia sind 27 Jahren verstrichen. Dieses ist eine lange Zeitspanne, während welcher diese Ausbildungseinheit ständig als „vorbildlich funktionierende“ und mit hoher Anerkennung der erreichten Ergebnisse und studentischer Leistungen sowohl in Bulgarien als auch im Ausland begutachtet wurde. Weltweit gilt die FDIBA ist das „größte und älteste deutsche Ausbildungsprojekt in Osteuropa“.

Während dieser Zeit hat man auf Grund der zeitbedingten Änderungen verschiedene Verfahren und Methoden bei der Ausbildung angewandt. Die technische Entwicklung auf allen Niveaus und die Praxis (Industrie) stellte an die Ausbildung immer neue und hohe Forderungen. Die einzige und immer aktuelle dabei war die Forderung an höchste Qualität der Studierenden zum Abschluss aller Fachgebiete und -stufen (Bachelor, Master, Promotion).

An der FDIBA erfolgt der Unterricht der Studierenden in den Fachsprachen der verschiedenen Disziplinen der Studiengänge „Allgemeiner Maschinenbau“, „Informatik“ und „Mechatronik“ sowie „Business Administration“ und „Industrial Management“ im eigentlichen Fachunterricht. Da der Unterricht in der gesamten Fakultät auf Deutsch erfolgt, erhalten die Studierenden, die meist bulgarischer Erstsprache sind, ihre Fachsprachenausbildung also in einer Fremdsprache.

Die Fachsprache bedient sich einer sehr präzisen und hochspezialisierten Terminologie, die den Erfordernissen des jeweiligen Faches angepasst ist und dank des technischen Fortschritts einer ständigen Innovation unterliegt. Aufgabe des Fachunterrichts ist es, die Studierenden zu einer selbstständigen, wissenschaftlich fundierten Auseinandersetzung mit der Fachsprache ihres Faches zu befähigen, die es ihnen ermöglicht, sie in vielen verschiedenen beruflichen und akademischen Kontexten kompetent anzuwenden.

Fachsprache zeichnet sich durch bestimmte Merkmale aus, die auch an Erstsprachler des Deutschen besondere Anforderungen stellen. Diese Anforderungen sind bei Fremd- bzw. Zweitsprachen lernen natürlich noch erhöht.

Beispiele für Fachwortschatz sind etwa „kondensieren“, „diffundieren“, „Volumen“, „Kohlenwasserstoff“, die gerade aufgrund ihrer Nähe zu Begriffen in anderen Sprachen als dem Deutschen besondere Schwierigkeiten machen können.

Ein weiteres Beispiel sind die mehrgliedrigen Komposita wie Essigsäure, Luftzufuhr, Brennerflamme, Lohnsteuerjahresausgleich.

Die Komposita folgen dabei besonderen Regeln (z.B. determiniert das letzte Substantiv Genus und damit Artikel des gesamten Kompositums), andererseits sind sie auch bemerkenswert „anarchisch“, da sie manchmal einen Fugenbuchstaben benötigen, manchmal nicht.

Formeln oder Abkürzungen können für die Deutschlernenden verwirrend sein, darüber hinaus gibt es Mehrwortverbindungen wie „Kälte-Wärme- Verhältnis“ sowie die sogenannte „Terminalisierung umgangssprachlicher Wörter“, bei der Wörter, die in der Alltagssprache in anderer Weise benutzt werden, in der Fachterminologie neu verwendet werden.

Beispiele sind „aufheben“, was normalerweise das Anheben von Gegenständen bezeichnet, in der Mathematik aber bedeutet, dass beim Multiplizieren und Dividieren mit der gleichen Zahl diese einander aufheben. Ebenso „umkippen“, das in der Ökologie mit dem Kollaps eines Gewässers („Der See kippt um“) metaphorisch gebraucht wird.

Abgesehen von den spezifischen Erfordernissen der Fachgebiete gibt es auch bestimmte überdauernde Merkmale von Fachsprache, die an der FDIBA in der allgemeinsprachlichen Kompetenz fördernden Deutschunterricht behandelt werden. Die Struktur von Fachsprache zeichnet sich gegenüber der Alltags- und Standardsprache durch bestimmte Besonderheiten, z.B. morphologischer oder syntaktischer Art aus. So überwiegen etwa substantivierte Infinitive (das Fräsen, das Schleifen), Adjektive auf -bar, -los, -reich, -arm, - frei, -fest usw. (informationsreich, brennbar,

nahtlos, kostenlos, sauerstoffarm, rostfrei, säurefest usw.), Funktionsverbgefüge (in Angriff nehmen, Anwendung finden, in Betrieb nehmen) oder Nominalisierungsgruppen (die Instandsetzung der Maschine, der Überführungsvorgang). Eine Besonderheit stellt auch den Unterschied zwischen Vorgangs- und Zustandspassiv („die Flamme wird entzündet“ vs. „die Flamme ist entzündet“) dar.

Diese strukturellen Besonderheiten, die Fachsprache – in unterschiedlichem Maße – von der Alltags- oder Standardsprache unterscheiden, können im Deutschunterricht von den Dozierenden, deren Aufgabe und Kompetenz in der Förderung allgemeinsprachlicher Kompetenz in den vier Teilbereichen schriftlicher und mündlicher Ausdruck, Hör- sowie Leseverstehen liegt, sehr gewinnbringend vermittelt werden.

Es ist offensichtlich, dass sowohl die angemessene Beherrschung der deutschen Sprache als auch der Fachsprache von größter Wichtigkeit sowohl für die Studienergebnisse als auch – für die Zufriedenheit der Studierenden mit ihrem Studium sind. Von Studenten wird immer wieder erwähnt, dass sie das Fachliche in Fächern wie Mechanik, MKL, Werkstoffkunde usw. auf Grund der Sprache nicht verstehen. Wenn man in einer Sprechstunde langsamer und mit anderen Worte die Probleme und Lösungen diskutiert, können die Unklarheiten jedoch in der Regel behoben werden.

Diese Tatsache brachte uns auf die Idee in die Pflichtveranstaltungen des Deutschunterrichts fachliche Diskussionen zu integrieren. Dabei ist es wichtig, das Hörverstehen und den mündlichen Ausdruck spezifischer Fachterminologie zu verbessern. Praxisnahe Simulationen und konkrete Beispiele aus der Praxis können die Schwelle zur selbstständigen Anwendung von Fachterminologie senken und den Studierenden erste Routinen ermöglichen. Dafür haben wir in Unterrichtsversuchen für Studenten der Fachrichtungen „Allgemeiner Maschinenbau“ und „Informatik“ verschiedene Varianten ausprobiert:

- Ein schon im Plan der Veranstaltungen festgelegtes Thema wird mit Vorbereitung von den Teilnehmern diskutiert. Eine zusätzliche und auf die Technik orientierte Diskussion wird durch die Expertise einer Fachkraft erweitert. Dabei nahmen die Studenten diese „Erweiterung“ gern auf und improvisierten „nach Herzenslust“, wobei sie ihrer Phantasie freien Lauf ließen.
- Nach einer von einem Fachmann gehaltenen Vortrag über ein Erzeugnis aus der Technik /Automobilgetriebe/ Informationstechnologie, welches als Begriff bekannt war, durften die Studierenden ein Rollenspiel betreiben. Ganz frei erfolgte dabei eine Gruppenteilung. Die Studierenden haben Rollen und Themen selber gewählt. Eine Gruppe hat die Preis-Leistungsverhältnisse unter den Gesichtspunkt „Händler – Käufer“, eine andere die Energieprobleme mit Umweltdiskussion, eine weitere die Möglichkeit der Erweiterung der Technologie in Hinblick auf Elektromobilität und selbstfahrende Autos erörtert.
- Anhand eines Modells wurden Begriffe benutzt, die bei dem danach erfolgten Rollenspiel verwendet wurden. Die Fachausdrücke wurden erläutert ohne die rechnerische und wissenschaftliche Seite anzusprechen. Es wurden z.B. keine Festigkeitsberechnungen und Funktionsvarianten besprochen und durchgeführt. Es blieb bei dieser Form des Unterrichts als einziger Anhaltspunkt die Verwendung der Fachausdrücke und später im Rollenspiel–die neue und richtige Anwendung dieser Fachausdrücke. So konnte man feststellen wie weit die Fachausdrücke richtig verstanden wurden und ob die angebrachte Anwendung erfolgte.

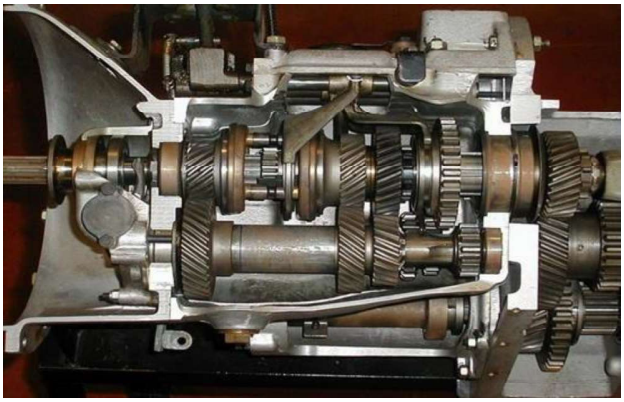


Abb. 1. Das Getriebemodell für die Begriffserklärung aus dem Unterrichtsversuch.

Die verschiedenen Varianten wurden unter Beteiligung einer Sprachlehrkraft ohne technische Ausbildung und einer technischen Fachfrau mit guten Deutschkenntnissen, aber ohne didaktische Ausbildung zum Sprachlehren durchgeführt. Unterrichtsversuche wie diese können bei guter Konzeption und Durchführung zum Nutzen der Studierenden „das Beste beider Welten“ vereinen. Wichtig ist aber, schon bei der Planung darauf zu achten, dass die Grenzen und Schwerpunkte der unterschiedlichen Kompetenzen der beteiligten.

Lehrpersonen geachtet und möglichst vorteilhaft kombiniert werden. Die „Simulation von Praxis“ bzw. Beispiele aus der Praxis selbst ermöglichen es den Studierenden, nach ihrer akademischen Ausbildung besser auf die Erfordernisse der Berufswelt vorbereitet zu sein.

Ist die Gefahr einer „Kompetenzüberdehnung“ gebannt, kann den Studierenden ein abwechslungsreiches Programm geboten werden, dass sie sowohl in allgemein- als auch fachsprachlicher Hinsicht den Ausbildungszielen der FDIBA näherbringt.

Zur Weiterentwicklung des Curriculums in den verschiedenen an der FDIBA angebotenen Studiengängen ist es möglich, weitere Unterrichtsexperimente wie dieses durchzuführen. Nach gründlicher quantitativer und qualitativer Evaluation und sich daraus ergebender Bewährung kann diese Form des interdisziplinären Fachsprachenunterrichts in die Lehrpläne aufgenommen werden. Die Ergebnisse der Auswertungen und können auch in den Entwurf und die Verbesserung softwareunterstützter Fachsprachenausbildung einfließen.

LITERATUR

[1] Rosemarie Buhlmann, Anneliese Fearns, *Handbuch des Fachsprachenunterrichts*, Gunter Narr Verlag, Deutschland, 2000.

[2] Fandrych, C., Thurmair, M., „Plädoyer für eine textsortenbezogene Sprachdidaktik,“ *DaF 2*, 2011. S. 84-93.

[3] Heidi Rösch, *Kompetenzen im Deutschunterricht*, Lang Verlag, Frankfurt a.M., 2005.

[4] Maria Steinmetz, Heiner Dintera, *Deutsch für Ingenieure. Ein DaF-Lehrwerk für Studierende ingenieurwissenschaftlicher Fächer,* Springer Verlag, Heidelberg u.a., 2014.

[5] Ulrich Steinmüller, „Deutsch als Fremdsprache: Didaktische Überlegungen zum Fachsprachenunterricht,“ *Zielsprache Deutsch*, 21.Jg. 1990, Heft 2, S. 16 –23.

[6] J. Popowa, E. Lazarou, M. Kovachev, “New Form of student Education on the German Engineering Education and Industrial management department at the Technical University of Sofia,” *Third international conference on Challenges in higher education and research in the 21st century*, June 1–3 2005, Sozopol, Bulgaria.

Ways to identify influence factors of virtual reality-systems on the acceptance in teaching - differences in generations X, Y and Z

Wege zur Identifikation von Einflussfaktoren auf die Akzeptanz von Virtual Reality-Systemen in der Lehre – Unterschiede in den Generationen X, Y und Z

Thomas Scheicher[*], Prof. Dr. Thomas Heupel[†]

[*] FOM Siegen, PhD-Candidate FDIBA Sofia, thomas.scheicher@web.de
[†] FOM Siegen, Vice-Rector Research, thomas.heupel@fom.de

***Abstract* —** Virtual reality-systems experience a boom in the public awareness and the availability not only in niches raises due to falling prices. Furthermore they are available as entertainment devices for a broad level of the population. Aside from big and stationary systems there are new possibilities to use virtual reality-systems in education. Nevertheless, how great is the acceptance of virtual reality-systems in the classical education in schools, vocational training and in the academic sector? And which factors influence the acceptance and willingness to use the systems? This contribution shall be seen as a way to identify and examine potential influence factors. In the framework of a preliminary study the author derived, that the knowledge of possible fields of application can raise the acceptance. Therefore, the author illustrates the aspired research design for the doctoral program of the FDIBA to identify additional influencing factors.

***Zusammenfassung* —** Virtual Reality-Systeme erleben einen Boom in der öffentlichen Wahrnehmung und sind durch gesunkene Preise nicht mehr nur in Nischen verfügbar, sondern stehen als Entertainment-Devices breiten Bevölkerungsschichten zur Verfügung. Auch in der Bildung ergeben sich dadurch neue Möglichkeiten abseits von großen und stationären Systemen. Doch wie hoch ist die Akzeptanz für einen Einsatz von Virtual Reality-Systemen in der klassischen Schul-, Berufs- und akademischen Bildung in Deutschland? Und welche Faktoren beeinflussen die Akzeptanz und Bereitschaft für den Einsatz dieser Systeme? Der eingereichte Beitrag soll einen Weg zur Identifikation und Untersuchung möglicher Einflussfaktoren aufzeigen. Der Autor hat im Rahmen eine Vorstudie abgeleitet, dass die Kenntnisse über mögliche Einsatzszenarien die Akzeptanz steigern und möchte innerhalb der FDIBA-Tagung das Forschungsdesign für die angestrebte Promotion vorstellen, mit der weitere Einflussfaktoren identifiziert werden können.

I. Einflussfaktoren auf die Akzeptanz für den Einsatz von Virtual Reality-Systemen in der Bildung

Virtual Reality-Systeme, also Systeme in denen ein Nutzer in Echtzeit mit einer virtuellen Welt interagieren kann, erleben in den 1990er Jahren einen ersten Boom in der öffentlichen Wahrnehmung [1]. Kommerzielle Anwendungen entwickeln sich in der Folge aber nur in sehr begrenztem Umfang und in einzelnen Branchen wie Militärtechnologie oder der Medizin [2]. Seit 2012 bieten viele Hersteller, wie zum Beispiel Facebook, Google, Sony, Hardware für Virtual Reality-Systeme für private Nutzer an [3]. Dies kann als Indiz dafür gewertet werden, dass die Systeme einen erneuten Boom erleben. Gestützt wird diese Einschätzung durch die Umsatzprognose des IT-Branchenverbandes BITKOM für den deutschen Markt, die einen jährlichen Umsatz von einer Milliarde Euro in 2020 prognostiziert. Dieser Trend stützt sich vor allem auf die steigende Bereitschaft Virtual Reality-Systeme privat einzusetzen [3]. Weltweit werden ebenfalls stark steigende Umsätze erwartet [4] [5] [6].

Diverse Studien zeigen, dass ein Einsatz von elektronischen Lernplattformen die Effektivität des Lernens positiv beeinflusst [7]. So überrascht es nicht, dass an Universitäten und Hochschulen aktuell eine Vielzahl von Virtual Reality-Lernsystemen entwickelt und eingesetzt werden [8] [9].

Die Einbindung von Virtual Reality-Systemen stellt jedoch dann eine sinnvolle Ergänzung von traditionellen Lernsituationen dar, wenn sich positive Effekte für den Lernprozess ergeben. Verschiedene Studien zeigen, dass unter anderem die räumliche Präsenz, die Authentizität der Lernumgebung sowie das Flow-Empfinden wichtige Einflussfaktoren für das Interesse am Lerngegenstand und damit den Erfolg einer Lernumgebung sind und dass sich diese Faktoren auf virtuelle Lernumgebungen übertragen lassen [10] [11] [12].

Neben aktueller Forschung zu Einsatzmöglichkeiten von Virtual Reality in der Bildung, kann der Stand der Forschung in Bezug auf die Akzeptanz von Virtual Reality in zwei weitere Bereiche unterteilt werden. Es existieren Untersuchungen zur prinzipiellen Akzeptanz von Informationstechnologien, dieser Bereich wird bereits seit längerer Zeit erforscht. Zusätzlich wird die spezifische Akzeptanz in Bezug auf Virtual Reality in besonderen Einsatzsituationen untersucht. Diese beziehen sich meist auf Gebiete, wie zum Beispiel die medizinische Ausbildung, in denen Virtual Reality-Technologien bereits im Einsatz sind. Tabelle 1 stellt eine Übersicht der entsprechenden Studien zur Verfügung.

TABELLE I. STUDIEN ZUR AKZEPTANZ

Autor	Studie
Davis, F. D. (1989)	Perceived Usefulness, Perceived Ease Of Use And User Acceptance Of Information Technology, MIS Quarterly, Ausgabe 13-3, S. 319-340.
Davis, F. D., Bagozzi, R. P., Warshaw, P. R. (1989)	User Acceptance of Computer Technology: A Compariosn Of Two Theoretical Models, in: Management Science, Ausgabe 35-8, S. 982-1003.
Legris, P., Ingham, J., Collerette, P. (2003)	Why Do Peapole Use Information Technology? A Critical Review Of The Technology Acceptance Model, in: Information & Management, Ausgabe 40-3, S. 191-204.
Venkatesh, V., Morris, M.G., Davis, G. B., Davis, F. D. (2003)	User Acceptance Of Information Technology: Toward A Unified View, in: MIS Quarterly, Ausgabe 27-3, S. 425-478.
Turner, M., Kitchenham, B., Brereton, P., Charter, S., Budgen D. (2010)	Does The Technology Acceptance Model Predict Actual Use? A Systematic Literature Review, in: Information And Software Technology, Ausgabe 52-5, S. 463-479.
Betrand, M., Bouchard, S. (2008)	Applying The Technology Acceptance Model To VR with people Who Are Favorable To Ist Use, in: Journal Of Cyber Therapy & Rehabilitation, Ausgabe1-2.
Rasimah, C. M., Ahmad, A., Zaman, H. B. (2011)	Evaluation Of User Acceptance Of Mixed Reality Technology, in: Australian Journal Of Educational Technology, Ausgabe 27-8, S. 1369-1387.
Disztinger, P., Schlögl, S., Groth, A. (2017)	Technology Acceptance of Virtual Reality for Travel Planning, in: Information and Communication Technologies in Tourism 2017 - Proceedings of the International Conference in Rome, Italy, January 24-26, 2017, S. 255-268.
Stein, J.-P., Ohler, P. (2017)	Venturing into the uncanny valley of mind – The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting, in: Cognition, Ausgabe 160, S. 43-50.
Mütterlein, J., Hess, T. (2017)	Immersion, Presence, Interactivity: Towards a Joint Understanding of Factors Influencing Virtual Reality Acceptance and Use, in: Adoption and Diffusion of Information Technology (SIGADIT) – Proceedings of Americas Conference on Information Systems (AMCIS), Boston, 2017.
Newbutt, N., Sung, C., Kuo, H. J., Leahy, M. J. (2017)	The Acceptance, Challenges, and Future Applications of Wearable Technology and Virtual Reality to Support People with Autism Spectrum Disorders, in: Recent Advances in Technologies for Inclusive Well-Being, in: Intelligent Systems Reference Library, Ausgabe 119, S. 221-241, Berlin, 2017.
Huang, H.-M., Rauch, U., Liaw, S. S. (2010)	Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist, in: Computers & Education, Ausgabe 55-3, S. 1171-1182.
Zhang, X., Jiang, S., Ordóñez de Pablos, P., Miltiadis, D. L., Sun, Y. (2017)	How virtual reality affects perceived learning effectiveness: a task–technology fit perspective, in: Behaviour & Information Technology, Ausgabe 36-5, S. 548-556.
Shen, C., Ho, J., Kuo, T., Luong, T. H. (2017)	Behavioral Intention of Using Virtual Reality in Learning, in: Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion), Genf, S. 129-137.
Pan, Z., Cheok, A. D., Yang, H., Zhu, J., Shi, J., (2006)	Virtual reality and mixed reality for virtual learning environments, in: Computers & Graphics, Ausgabe 30-1, S. 20-28.
Innocenti, A. (2017)	Virtual reality experiments in economics, in: Journal of Behavioral and Experimental Economics, In Press, Accepted Manuscript.

In der Folge ergibt sich eine Lücke in der Erforschung der Einflussfaktoren auf die Akzeptanz von Virtual Reality-Systemen und damit die Bereitschaft diese Systeme einzusetzen, insbesondere im Rahmen der schulischen, beruflichen und akademischen Bildung in Deutschland. Im Rahmen einer empirischen Vorstudie mit Studierenden konnte bestätigt werden, dass die Bereitschaft entsprechende Systeme einzusetzen sich durch die Kenntnis möglicher Einsatzszenarien steigern lässt: Der Anteil der Studienteilnehmer die keine Erfahrung mit Virtual Reality-Systemen hatten sich aber einen Einsatz als Lehrmedium im Studium vorstellen konnten lag bei knapp über 50%. Nach Beschreibung eines realistischen Einsatzszenarios lag dieser Prozentsatz bei über 90%. Ebenso zeigt ein erstes Experiment bei dem Lehrende ohne Vorkenntnisse in der Nutzung von Virtual Reality-Systemen, einfache Virtual Reality-Systeme einsetzen (Google Expeditionen) das sich der Prozentsatz der Teilnehmer, die sich den Einsatz in der Lehre vorstellen können von 30% auf über 80% steigt. Diese Erkenntnisse müssen aber weiter spezifiziert werden um sicherzustellen, dass die aktuell entwickelten Konzepte für den Einsatz von Virtual Reality-Systemen in der Bildung auch auf die Bereitschaft treffen, diese einzusetzen. Die bisherigen Untersuchungen berücksichtigen dabei den Einfluss der unterschiedlich geprägten Generationen X, Y und Z nicht. Als Generation wird dabei eine Gruppe von Menschen verstanden, die eine gemeinsame soziale Identität aufgebaut haben, da sie sich in Bezug auf ihre Eigenschaften und Wertvorstellungen ähneln [13]. Unterschieden wird dabei aktuell zwischen den in Tabelle 2 dargestellten Generationen [14]. Für den weiteren Verlauf der Untersuchung werden nur die Generationen X, Y und Z betrachtet, da die Mehrzahl der Lehrenden und Lernenden in deutschen Bildungseinrichtungen zu diesen Generationen gehören.

TABELLE II. ÜBERSICHT GENERATIONEN

Generation	Geboren von	Geboren bis
Tradionalisten	1929	1945
Baby Boomer	1946	1964
Generation X	1965	1979
Generation Y	1980	1995
Generation Z	1996	

II. IDENTIFIKATION VON WEITEREN EINFLUSSFAKTOREN UND FORSCHUNGSDESIGN ZUR VALIDIERUNG

Zur Untersuchung der Einflussfaktoren auf die Akzeptanz und die Rolle der Generationen wird im Rahmen der Promotion ein dreistufiges Vorgehen, schematisch in Abbildung 1 dargestellt, gewählt.

A. 1. Stufe: Qualitative Vorstudie: Welche Einflussfaktoren lassen sich identifizieren?

Durch Befragung von Experten und Beobachtung von Testgruppen die Virtual Reality-Systeme praktisch einsetzen, sollen, neben den aus der Literatur bekannten Faktoren wie Qualität der Darstellung, Benutzerfreundlichkeit, Kosten, verfügbarer Content und Präsenzerleben, Einflussfaktoren auf die Akzeptanz identifiziert werden. Zusätzlich soll die Befragung der Experten eine Einschätzung der Technologieentwicklung in den nächsten Jahren liefern und so eine Abschätzungsmöglichkeit für zukünftige Einflussfaktoren liefern.

Dabei liefert ein Pretest der Vorstudie mit insgesamt 188 Teilnehmern folgende Ergebnisse: Innerhalb der Generationen X (35), Y (113) und Z (40) ist ein unterschiedlicher Kenntnisstand in Bezug auf das Konzept Virtual Reality festzustellen. Während innerhalb der Generation Z 98% der Befragten mit dem Konzept vertraut sind, beträgt die Bekanntheit in den Generationen Y und Z nur jeweils 86%. Deutlicher wird der Unterschied zwischen den Generationen bei der Akzeptanz der Systeme. Einen Lehreinsatz von Virtual Reality-Systemen können sich in Generation X 57%, in Y 61% und innerhalb der Generation Z 80% der Teilnehmer des Pretests

vorstellen. Nach Erhebung dieser Daten werden die Teilnehmer mit einem theoretischen Einsatzszenario konfrontiert und beantworten erneut einen Fragebogen. Dabei erhöhen sich die Akzeptanzwerte für einen Lehreinsatz von Virtual Reality-Systemen in allen Generationen. Am deutlichsten fällt der Zuwachs in Generation X und Y mit einem Anstieg auf 86% beiziehungsweise 91% aus. Auch in Genration Z findet ein Zuwachs auf 90% statt. Abbildung 1 stellt diesen Zuwachs grafisch dar.

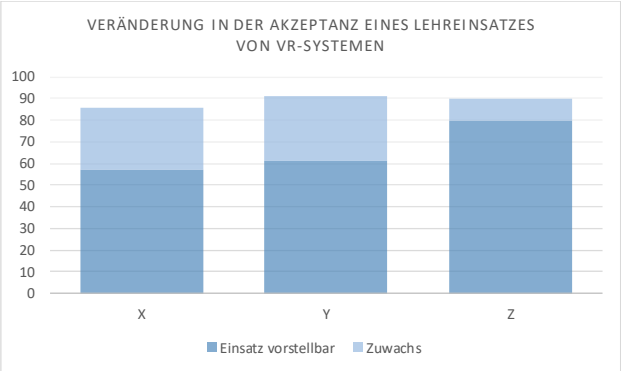


Abb. 1. Veränderung in der Akzeptanz des Lehreinsatzes durch Kenntnis eines theoretischen Einsatzszenarios in den Generationen X, Y und Z.

Die Ergebnisse des Pretests liefern jedoch nur einen ersten Eindruck und erfordern eine weitere Untersuchung mit erhöhter Fallzahl und einer differenzierteren Betrachtung der Einflussfaktoren sowie der praktischen Durchführung des theoretischen Einsatzszenarios.

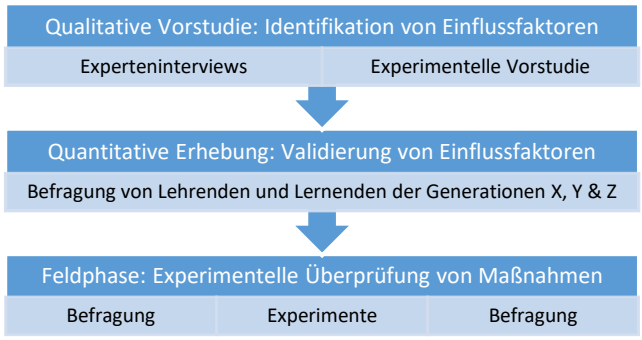


Abb. 2. Darstellung des Forschungsdesigns

B. 2. Stufe: Quantitative empirische Erhebung: Welche Einflussfaktoren können quantitativ validiert werden?

Durch Befragung von Lehrenden und Lernenden aus allen Stufen des deutschen Bildungssystems und aus den drei Generationen X, Y und Z findet eine quantitative Überprüfung der in der ersten Stufe ermittelten Einflussfaktoren statt. Aus den validierten Einflussfaktoren werden dann Maßnahmen abgeleitet, die zu einer Steigerung der Akzeptanz für den Einsatz von Virtual Reality-Systemen in der Bildung führen sollen.

C. 3. Stufe: Feldphase: Welche abgeleiteten Maßnahmen zur Akzeptanzsteigerung können experimentell validiert werden?

In der dritten Stufe werden die aus den Ergebnissen von Stufe zwei abgeleiteten Maßnahmen in Form von Experimenten mit Teilnehmern der Generationen X, Y und Z aus allen Stufen des deutschen Bildungssystems durchgeführt. Dazu findet zunächst eine Befragung der Teilnehmer statt, im Anschluss werden entsprechende Experimente durchgeführt, die wiederum mit einer Befragung der Teilnehmer abschließen. Durch diesen Aufbau können Veränderungen in der Akzeptanz gemessen werden und die Auswirkungen der Maßnahmen quantifiziert werden.

III. Fazit

Durch das gewählte Design kann am Ende die Forschungsfrage beantwortet werden, welche Faktoren aktuell und in Zukunft die Akzeptanz von Virtual Reality-Systemen im Bildungseinsatz in Deutschland beeinflussen und welche Rolle die Zugehörigkeit zu den Generationen X, Y und Z dabei spielt. Im Rahmen der Tagung soll die bereits begonnene Vorstudie kurz dargestellt und das hier skizzierte Forschungsdesign diskutiert werden. Es sollen durch das Feedback des Auditoriums wertvolle Hinweise zur Durchführung der sich anschließenden Stufen 2 und 3 gewonnen werden. Die gewonnenen Ergebnisse können dazu dienen den Einsatz von Virtual Reality-Systemen mit Maßnahmen zur Akzeptanzsteigerung zu begleiten.

References

[1] McLellan, H. (2004): “Virtual Realities”, in: Jonasson, D. H. (Hrsg.), *Handbook Of Research On Educational Communications And Technology*, 2004, 2. Auflage, Mawah, S. 461-498.

[2] Warnecke, H.-J. (1993): „Virtual Reality – Anwendungen und Trends“, 1993, in: Bullinger, H.-J., Warnecke, H.-J., Westkämper, E. (Hrsg.), *IPA-IAO - Forschung und Praxis*, Band 35, Mainz, S. 9-16.

[3] BITKOM (2016): „Zukunft der Consumer Technology“, 2016, URL: www.bitkom.org/noindex/Publikationen/2016/Leitfaden/CT-Studie/160831-CT-Studie-2016-online.pdf, Abruf am 30.09.2016.

[4] Statista (2014): „Prognose zum Umsatz mit Virtual Reality weltweit in den Jahren 2014 bis 2018 (in Millionen US-Dollar)“, 2014, URL: de.statista.com/statistik/daten/studie/318536/umfrage/prognose-zum-umsatz-mit-virtual-reality-weltweit/, Abruf am 17.06.2016.

[5] Statista (2015): „Anteil der Befragten in Deutschland im Jahr 2015, die sich vorstellen können eine Virtual-Reality-Brille (z.B. Oculus Rift) zu nutzen, nach Altersgruppe“, 2015, URL: de.statista.com/statistik/daten/studie/438899/umfrage/umfrage-zum-interesse-an-virtual-reality-brillen-in-deutschland-nach-alter/, Abruf am 10.12.2016.

[6] Statista (2016): „Anteil der deutschen Internetnutzer, die Virtual-Reality-Brillen für digitale Spiele nutzen wollen, nach Altersgruppe“, 2016, URL: de.statista.com/statistik/daten/studie/448749/umfrage/umfrage-zum-interesse-an-virtual-reality-brillen-fuer-digitale-spiele-in-deutschland-nach-alter/, Abruf am 10.12.2016.

[7] Kapp, K. M. (2012): *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*, 2012, San Francisco.

[8] Zinn, B., Guo, Q., Sari, D. (2016): „Entwicklung und Evaluation der virtuellen Lern- und Arbeitsumgebung VILA“, 2016, in: Zinn, B., Tenberg, R., Pittich, D. (Hrsg.), *Journal of Technical Education*, 4 Jg., Ausgabe 1, S. 89-117.

[9] HarvardUniversity (2016): “CS50”, 2016, URL: cs50.harvard.edu/, Abruf am 31.12.2016.

[10] Schraw, G., Bruning, R., Svoboda, C. (1995): “Sources of situational interest”, 1995, in: *Journal of Reading Behavior*, 27. Jg., Ausgabe 1, Seiten 1-17.

[11] Schraw, G., Flowerday, T., Lehman, S. (2001): “Increasing situational interest in the classroom”, 2001, in: *Educational Psychology Review*, 13. Jg., Ausgabe 3, S. 211-224.

[12] Sun, J. C. Y., Rueda, R. (2012): “Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education”, 2012, in: *British Journal of Educational Technology*, 43. Jg., Ausgabe 2, S. 191-204.

[13] Troger, H. (2016): *7 Erfolgsfaktoren für wirksames Personalmanagement – Antworten auf demographische Entwicklungen und andere Trends*, 2016, Wiesbaden.

[14] Mangelsdorf, M. (2014): *30 Minuten – Generation Y*, 2014, Offenbach.

This page is intentionally left blank

Quality Assurance and Product Liability with Selected Examples from the European Automobile Industry

Qualitätssicherung und Produktverantwortung Ausgewählte Probleme am Beispiel der europäischen Automobilindustrie

Invited Talk

Peter Greulich
Akad. Dir. i. R. Leibniz Universität Hannover,
Bundesrepublik Deutschland – D 30559 Hannover – Kronsberger Str. 27
e-mail: greulich@jura.uni-hannover.de

Abstract — The presentation covers one of the most critical problems in the industrialized world nowadays – especially for the countries of the European Union. With regards to the industry, we are all essentially dependent on production in our individual economic area; we live with and from its output. This is the basis of our economic prosperity. More innovative, effective and precise work ensures that we are able to stay competitive in the global economy. The main prerequisite for our economic activities is the capability to produce high quality, safe and durable products in the European Economic Area – as a global leader in industrial progress. At the same time, the EU's common market requires optimal consumer protection. The talk focuses on the following two general topics:

- the assurance of the product quality; and
- the protection of the "consumer" more broadly.

Zusammenfassung — Der Vortrag behandelt eines der wichtigsten Problemfelder unserer Industriegesellschaft weltweit – insbesondere für die Länder der Europäischen Union. Aus der Sicht der Industrie sind wir alle im Wesentlichen von der Produktion in unserem einheitlichen Wirtschaftsraum abhängig; wir leben mit und von ihr. Sie ist die Basis unseres wirtschaftlichen Wohlstandes. Je innovativer, produktiver und präziser wir arbeiten, umso besser werden wir gegenüber der Konkurrenz im Rahmen der globalen Wirtschaft bestehen können. Grundvoraussetzungen für unser wirtschaftliches Handeln sind: hochqualitative, sichere und haltbare Produkte, die wir in unserem europäischen Wirtschaftsraum herstellen – als weltweite Spitze im industriellen Fortschritt. Auf der anderen Seite verlangt unser gemeinsamer Markt innerhalb der Europäischen Union einen optimalen Schutz der Verbraucher. Mit beiden Themenfeldern will sich der Vortrag in Grundzügen auseinandersetzen:

- mit einer Sicherung der Qualität unserer Produkte und
- mit dem Schutz der „Verbraucher" im weitesten Sinne.

This page is intentionally left blank

The Requirement to Deceive – a Serious Shortcoming of the Bulgarian Computer Fraud Regulation

Die Täuschungsanforderung - ein ernster Mangel der bulgarischen Computer-Betrugsverordnung

Ralitza Dimitrova

Faculty of Management, Technical University of Sofia
Sofia, Bulgaria, rvd@tu-sofia.bg

***Abstract* —** Digital transformation is impossible without effective protection against cyber crime. Different forms of unlawful conduct committed through new technologies pose a serious threat against information society. The present article deals with computer-related fraud – one of the most widely spread cyber crime offences which causes significant losses to individuals, businesses and whole countries. The computer fraud provisions in the Bulgarian Criminal Code are analyzed in relation to a brief comparative analysis of the computer fraud regulation in another six Member States of the European Union. Some serious shortcomings in the Bulgarian criminal law regulation of computer fraud are pointed out and proposals for amendments are made.

***Zusammenfassung* —** Digitale Transformation ist ohne wirkungsvollen Schutz gegen Cyberkriminalität unmöglich. Die verschiedenen Formen des rechtswidrigen Verhaltens, die durch neue Technologien begangen werden, stellen eine ernsthafte Bedrohung für die Informationsgesellschaft dar. Der vorliegende Artikel befasst sich mit Computer-Betrug - ein der am weitesten verbreitete Verbrechensverhalten, das erhebliche Verluste für Einzelpersonen, Unternehmen und ganze Länder verursacht. Die Computerbetrugsbestimmungen im bulgarischen Strafgesetzbuch werden in Bezug auf eine kurze Vergleichsanalyse der Computerbetrugsverordnung in weiteren sechs Mitgliedstaaten der Europäischen Union analysiert. Es werden auf einige ernsthafte Mängel in der bulgarischen Strafrechtsregulierung von Computerbetrug hingewiesen und Vorschläge für Änderungen unterbreitet.

I. Introduction

It seems that at every turn, someone is talking about digital transformation [1]. But digital transformation is not just a modern word. It is a phenomenon intended to benefit our lives in almost every aspect. Even to the casual observer, the existence of technology trends such as hyper connectivity, smart devices, and cloud computing is obvious [1]. Digital transformation changes both the way the public sector operates and the business is exercised. Through e-government people are allowed to communicate with state and local institutions in a fast, easy and convenient way. In the private sector more and more individuals and companies use new technologies to buy and sell goods and services, make electronic payments, check their bank accounts, etc. Digital transformation allows companies to improve how they currently operate, create new business models and sources of value, and maintain a competitive advantage [1] and thus benefits also customers of goods and services.

Unfortunately, it is clear that information technologies are used not only for good purposes – to get the mankind's life better, but also by criminals for the commission of various wrongful acts with adverse consequences. Infringement of the inviolability of correspondence, child pornography, fraud, hacking, cracking – all these, and many other crimes are committed through new technologies. Individuals and companies suffer monetary and moral damages unseen before. Therefore, security in the information society and digital transformation are at great risk. Experts point out that the automation of virtually all business processes and the increasing digital connectedness of the entire value chain create agility, but they also significantly raise cyber security risks and threat levels [2]. Cyber security has become a key strategic priority for digital business and is a topic we need to be open about if we want to succeed in digital transformation [3]. Therefore, the reliable cyber security and the effective fight against cybercrime are major priorities for the information society.

The very nature of cybercrime makes it especially dangerous and difficult to tackle. Cybercrime is one of the fastest growing forms of crime, with more than one million people worldwide becoming victims each day. Cybercrimes are high-profit and low-risk, and criminals often exploit the anonymity of website domains. Cybercrime knows no borders - the global reach of the Internet means that law enforcement must adopt a coordinated and collaborative cross-border approach to respond to this growing threat [4,5]. It is very important that the majority of countries worldwide implement the main international instrument in this field – the Council of Europe Convention on Cybercrime, and guarantee an effective criminal law protection against cybercrime. EU Member States should also harmonize their national legislations with legal instruments such as Directive 2013/40/EU on attacks against information systems.

Bulgaria has already made a series of amendments to its Criminal Code (BCC) to meet the requirements of both the international and the EU instruments dedicated to the fight against cybercrime. The different cybercrime offences – against the person, against the inviolability of correspondence, against the property, and those representing a direct attack against computer systems and data, were introduced either by

creation of new provisions or by insertion of amendments to already existing provisions of the BCC.

Because of the limited volume of the present paper, it will focus only to one of the so-called computer-related computer crimes – the computer fraud. Nowadays it is more dangerous than traditional fraud – it causes significant losses to both individuals and businesses, even to whole countries. That's why it is very important that criminal law legislations ensure an effective protection of property against fraud committed through new technologies.

In the first part of the paper a brief analysis of the current criminal law regulation of computer fraud in Bulgaria will be made. In the second part the main features of the relevant criminal law provisions in another six EU Member States will be presented. Some of them are old Member States and the other – new Member States of the EU. Besides, five of the said jurisdictions pertain to the Civil Law system and one to the Common Law system, so that the different approaches influenced by the different legal principles and traditions can be taken into account. The third part of the article is dedicated to a comparative analysis where the positive and negative features of the Bulgarian criminal law regulation of computer fraud are pointed out in relation to what are the main trends in the analyzed foreign jurisdictions. In the end the author draws some conclusions based on the comparative analysis and makes some proposals for amendments in the BCC.

II. Criminal Law Regulation of Computer Fraud in Bulgaria

The cybercrime offences were introduced into the BCC by a series of amendments in 2002, 2007 and 2010. For the offences that represent a direct attack to computer systems and data a new chapter "Cybercrime" was elaborated. The other computer-related offences were placed in different chapters of the Code, close to the respective traditional offences, following the principle to arrange the different types of offences in the code according to the character and the importance of the protected social relations. Computer fraud was introduced as article 212a of the BCC, immediately after the traditional fraud and the document fraud

Article 212a BCC describes the computer fraud offence in two main provisions without any special provisions which provide heavier or lighter sanctions.

Under article 212a, paragraph 1 BCC where an individual, in view of providing a benefit to him-/herself or another, brings or maintains misleading representations in someone through introducing, modifying, deleting, or erasing computer data or through the use of an electronic signature of another causes him/her or another damage, shall be punished for computer fraud by deprivation of liberty from one to six years and a fine of up to BGN six thousand.

Under article 212a, paragraph 2 BCC the same sanction shall be imposed to the individual who, without being entitled thereto, introduces, modifies, or erases computer data in order to unduly obtain something, that should not go to him.

The following components of the offence are worth noting.

The object of computer fraud is complex and includes different groups of social relations: (1) the relations that guarantee the normal exercise of the property rights; (2) the relations that ensure that people make transfer of property rights upon free decision and clear mind; and (3) the relations which protect computer data against any unlawful manipulation.

The prohibited conduct under article 212a, paragraph 1 reveals the first form of the computer fraud as just a "computerized" version of the traditional fraud – bringing or maintaining of misleading representations [6]. In other words the perpetrator should influence upon another person's mind and deceive him or her about significant elements of an act of transfer of property.

What differentiates computer fraud is the specific method of committing the crime - through a manipulation of a computer data or through the use of an electronic signature of another. The perpetrator either interferes with computer data or uses an electronic signature, thus creating the false impression that the electronic statement is made by the real holder of the signature.

Another component from the objective aspect of the offence is the conduct of the deceived person – guided by the misleading representations he or she performs an act of disposition of rights over real or personal property.

The crime is naturally a result one – to be completed damage must be caused to another's property. Computer fraud is an offence against property which means that the damage must have monetary nature. On the other hand, the description of the offence does not specify the nature of the damage – therefore, causing a moral harm will meet the requirements [6].

The subject of computer fraud is every physical person who is not authorized to introduce the mentioned modifications to the computer data or a person who is not the real holder of the digital signature.

The second form of computer fraud (article 212a, paragraph 2 BCC) has a very broad formulation and criminalizes all kinds of unlawful modifying of existing computer data [6].

This form of the offence requires a criminal result to be completed too – the fact of introducing of the said modifications. However, damage to another person's property here is not needed – it is not an element of the offence.

The perpetrator commits the prohibited conduct without authorization, which is another element in the objective aspect.

The subject is a physical person which has not the right to influence the computer data.

The subjective aspect of the computer fraud requires a direct intent as well as a special aim under article 212a, paragraph 1 (providing a benefit for oneself or another) and a special intention under article 212a, paragraph 2 (unduly obtain something).

The analysis shows that article 212a of the BCC criminalizes an illegal conduct committed through new technologies with the intent to cause an illegal transfer of property in compliance with the aim of the international and EU instruments. However, some studies state that the Bulgarian criminal law provisions on computer fraud are not in full compliance with the Convention on Cybercrime [5,6]. In fact there are some serious shortcomings which make the computer fraud regulation ineffective. They will be discussed below in the light of the comparative analysis.

III. Criminal Law Regulation of Computer Fraud Abroad

Computer fraud is usually described as one of the computer-related traditional crimes [7] or computer based or aided crimes [8]. Deceiving somebody else in order to get for oneself (or for another person) a material benefit is in fact a traditional offence for the criminal law systems. What distinguishes the computer fraud is the use of new technologies for committing the crime and, of course, as we shall see below, the very different mechanism of the crime.

Most criminal law legislations provide for a specific offence criminalizing unauthorized manipulation committed during data processing with the intent to cause an illegal transfer of property (all analyzed legislations below). Other jurisdictions rely on the provisions on traditional fraud to cover also the computer-related fraud (for example France).

Further, the comparative analysis shows that computer fraud is usually regarded as one of the crimes against property, alongside with its predecessor – the traditional fraud. The computer fraud provisions are intended to protect the property interests and naturally the offence is most commonly placed in those parts of the criminal laws which are dedicated to the crimes against property (Germany, Italy, Poland, Romania, Latvia, etc.).

The computer fraud provisions, like other cybercrime provisions, make reference to specific terms connected to new technologies. The most frequently used terms are "computer data" (Poland, Romania) or just "data" (Latvia, Germany); "computer system" (Romania) or "information system" (Italy). Not so long ago in the Latvian Criminal Law was inserted the concept of an "automated data processing system" [9] and the German Criminal Code uses another term – "data processing operation". The Irish provision refers just to a "computer". Generally these terms have a legal definition in the respective piece of legislation. The terms used and their definitions in the national legislations are important aspect of the effective legal framework of fight against cybercrime.

As far as the prohibited conduct is concerned, usually it takes a variety of forms under the analyzed legislations. For example, under the German Criminal Code the perpetrator damages another person's property by influencing the result of a data processing operation through incorrect configuration of a program, use of incorrect or incomplete data, unauthorized use of data or other unauthorized influence on the course of the processing (Section 263a of the German Criminal Code). Under article 640 ter of the Italian Criminal Code the unlawful conduct can take two alternative forms - altering in any way the operation of a computer system or intervening in any manner without the right on data, information or programs contained in a computer system [10]. The forbidden conduct in Poland is any inputting, altering or deleting of computer data or other unauthorized interference with its processing (article 287 1(55) of the Polish Criminal Code); in Romania – entering, altering or deleting computer data or restricting access to such data or hindering in any way the operation of a computer system (article 249 of the Romanian Criminal Code). In Ireland there is a really broad statement which deserves to be mentioned - the offender "operates or causes to be operated a computer" (section 9 of the Criminal Justice (Theft and Fraud Offences) Act of 2001).

In some of the analyzed jurisdictions computer fraud is formulated as a result crime (Germany, Italy, Romania). The offence is completed only when the victim's property has suffered some monetary damage [10,11, 12]. Under the Italian legislation the criminal result has two cumulative components - obtaining a benefit for the perpetrator (or for another person) and causing loss to the victim's property. In some studies it is pointed out that the damage can have also moral nature [10]. Other jurisdictions do not require any materialized criminal result from the offender's conduct (conduct crime) (Poland, Ireland, Latvia). In the opinion of some authors the formulation of the offence as a conduct crime in Poland hinders its effective implementation [7].

A very interesting aspect of computer fraud is the mental element of the offence. All analyzed computer fraud provisions require such element. A direct intent is needed accompanied by a special purpose – "the purpose of making a profit or causing harm" (Poland); "in order to obtain a benefit" (Romania); "for the acquisition of property … in order to influence the operation of the resources thereof" (Latvia) or a special intention – "intent of obtaining an unlawful material benefit" (Germany); "with the intention of making a gain for himself or herself or another, or of causing loss to another" (Ireland). As it can easily be seen in some cases the additional mental element contains alternatively or cumulatively the aim or intention to cause damage to another person's property (Poland, Ireland).

In the end, the comparative analysis will be incomplete without mentioning the special provisions of computer fraud. For example, in a "case of lesser importance" the Polish Criminal Code provides for a lighter penalty than in the main provision (article 287, paragraph 2). The Italian Criminal Code contains a number of special provisions created with a view to different qualifying features: if the offence is committed to the detriment of the State or to another public body; if the quality of an operator of the system is abused; if the offence is committed by theft or by illicit use of digital identity to the detriment of one or more persons (article 640 ter, paragraph 2 and 3). In Latvia the computer fraud under article 177, paragraph 2 of the Criminal Act brings a heavier penalty if it has been committed by a group of persons pursuant to prior agreement, and under article 177, paragraph 3 - if it has been committed on a large scale or if it has been committed in an organized group.

It is worth mentioning that some studies give the German and the Romanian provisions as an example for a model of full alignment with article 8 of the Convention on Cybercrime [5]. Italian and Polish provisions should be praised for the formulation of special descriptions which make possible the differentiation of the criminal liability. Besides, the broad statement of the Irish computer fraud provision should be noted as a good example for a formulation tailored to cover any future technical developments.

IV. Conclusions

On the basis of the analysis of computer fraud regulation in Bulgaria and abroad, the following conclusions can be drawn.

First, the Bulgarian computer fraud provisions are placed in the chapter "Crimes against Property" in the BCC which adequately corresponds to the object of the offence. The comparative analysis already showed that this is also the preferred approach in other jurisdictions. However, there is no reasonable explanation why the document fraud regulation was torn into two parts by inserting the computer fraud offence between the special provisions of document fraud providing for heavier penalties and the special provisions providing for lighter penalties.

Second, the Bulgarian computer fraud provisions use adequate special terms which are in compliance with the main international and EU legal instruments in this field and have suitable legal definitions in the Code (article 93 BCC). It's worth mentioning that some foreign experts also retain that the definitions in the BCC correspond to those in the Council of Europe Convention on Cybercrime [5].

Third, but most important, the present formulation of computer fraud in paragraph 1 of article 212a BCC requires that a physical person must be deceived (the perpetrator "brings or maintains misleading representations in someone"). This is in fact the traditional construction used for the traditional fraud. The problem is that the lawmaker obviously do not realize that the two offences have a very different mechanism – while in the case of the classic fraud the perpetrator needs to influence the mind of a physical person and deceive him or her about one or more substantial elements of an act of transfer of property, in the case of the computer fraud the perpetrator uses the possibilities of new technologies and needs not influence a human mind, only a computer system and/or data, in order to commit the crime. As it was shown above, the analyzed foreign legislations naturally do not require such element. Although their formulations vary

from one another, they obviously correspond far more adequately to the mechanism of computer fraud and to the model provisions of international and EU instruments. Moreover, the difference between the two offences is explained both in the literature and the case law [6, 10, 11, 13,14]. It is pointed out that the core element of the traditional fraud is the deception of the victim, while the computer fraud is committed through an unauthorized influence on a computer system.

While discussing the objective elements of the offence, it's worth mentioning also that foreign legislations often give more general and clear descriptions of the offence which encompass the broad variety of forms and ways to illegally influence computer systems and/or data and allow further technological development. As it was mentioned above some of the analyzed legislations are given as good examples in international studies on cybercrime.

Therefore, exactly the current description of the actus reus should be highlighted as the most serious problem of the Bulgarian computer fraud regulation. The present formulation of article 212a BCC does not correspond to the mechanism of performing the computer fraud by means of new technologies. On one hand, it renders the provisions inconsistent with the current development of new technologies, and on the other – with the trends in the development of the criminal law in this field. As a result, the scope of application of the current provisions is too limited.

Fourth, rather uncommon for the Bulgarian criminal legislation is the lack of special provisions of computer fraud in the BCC. First of all, offences against property offer broad possibilities for differentiation of the criminal liability through elaboration of special provisions with a view to different qualifying features, or, on the contrary, features that decrease the social danger of the respective criminal conduct and justify a lighter penalty. Second, as it was demonstrated above, foreign legislations provide for such special provisions of computer fraud. Such special provisions should be introduced at least for the cases of organized crime, conspiracy, large amount of the damages caused, etc.

In conclusion, the present provisions of the computer fraud in the BCC should be amended in the light of the analysis results above.

References

[1] M. Haendly, "5 Tangible benefits of digital transformation. How digital technologies can increase insight and innovation in your business", *SAPinsider*, vol. 17, issue 2, April 2016.

[2] M. Golz, J. Somaini, "Cybersecurity in the Age of Digital Transformation", *MIT Technology Review Custom*, January 2017.

[3] "Cybersecurity: security risks and solutions in the digital transformation age". Online. Available at https://www.i-scoop.eu/cyber-security-cyber-risks-dx/, 2017.

[4] European Commission, Joint Communication to the European Parliament, the Council, the European Economic and Social Council and the Committee of the Regions. Cybersecurity strategy of the European Union: an open, safe and secure cyberspace /* JOIN/2013/01 final */, 07.02.2013.

[5] Council of Europe, Project on Cybercrime, National legislation implementing the Convention on cybercrime – comparative analysis and good practices, March 2008.

[6] A. Stoynov, *Criminal law. Crimes against property*, Sofia: Ciela, 2006, p. 169 and the following.

[7] A. Adamski, "Cybercrime legislation in Poland", Nicolaus Copernicus University, Poland, 2010.

[8] P. Ryan and A. Harbison, "The Law on computer fraud in Ireland - development of the law on dishonesty", Society for Computers and Law, Arthur Cox, June 2010.

[9] R. Dimitrova and E. Saulitis, "Criminal law framework for combating computer crime in Latvia and Bulgaria", XV International Scientific Conference "Management and Engineering'17", Sozopol, Bulgaria, 2017.

[10] S. Logroscino, "La frode informatica quale autonoma figura". Online. Available at http://www.altalex.com/documents/news/2012/01/02/la-frode-informatica-quale-autonoma-figura-di-reato-rispetto-al-delitto-di-truffa, 2011.

[11] K.J. Heller and M.D. Dubbler, Eds. *The handbook of comparative criminal law*, Stanford: Stanford University Press, 2011, p. 279.

[12] M. Bohlander, *Principles of German criminal law*, Oxford: Hart Publishing, 2008, p. 30, 223.

[13] M.R. Pensabene, "La frode informatica, una figura autonoma di reato". Online. Available at http://www.masterlex.it/approfondimenti/la-frode-informatica-figura-autonoma-reato/, 2017.

[14] Sentenza 24 febbraio 2017 n.9191, Corte di Cassazione della Repubblica Italiana.

Digital Transformation and its radical changes for external management accounting: a consideration of small and medium-sized enterprises

Die digitale Transformation wird für kleine und mittlere Unternehmen das externe Controlling radikal verändern

Serap Demiröz*, Thomas Heupel†
*PhD Student at University of Applied Sciences for Economics and Management FOM Essen/UCAM Essen, Germany serap.demiroez@fom-net.de
†Vice rector for research, Professor for management, especially management accounting, FOM Essen Essen, Germany, thomas.heupel@fom.de

***Abstract* —** Small and medium-sized enterprises (SMEs) are decisive for Germany's growth. They provide a large number of jobs, secure a company's long-term existence and place great value on Germany's competitiveness by investments. Despite such strengths, they lack in management accounting. Analyses reveal that 79% of company insolvencies can be traced back to a lack in management accounting. Such situations require the strengthening of external management accounting in SMEs. In the light of digital transformation, SMEs will hence have the chance to expedite external management accounting in contrast to internal management accounting. Thus, SMEs will reach new business models and advantages in competition. The aim of this paper is to answer the question if Digital Transformation can advance external management accounting for SMEs. A SWOT analysis will serve as methodology to answer this question. Firstly, by means of SWOT analyses, strengths and weaknesses of SMEs regarding external management accounting will be outlined and chances and risks of the Digital Transformation will be revealed for the next three to five years. Subsequently, management implications caused by Digital Transformation will be elucidated. Finally, strategies for SME management will be established.

***Zusammenfassung* —** Kleine und mittlere Unternehmen (KMUs) sind entscheidend für Wachstum und Wohlstand in Deutschland. Sie schaffen Arbeitsplätze und prägen durch Investitionen die Wettbewerbsfähigkeit Deutschlands. Trotz dieser Stärken weisen KMUs dennoch Schwächen auf, zu denen ein fehlendes Controlling gehört. Analysen zufolge stellt ein fehlendes Controlling mit einem Anteil von 79% aller Fälle eine der häufigsten Insolvenzursachen dar. Dieser Zustand erfordert, dass das Controlling in KMUs gestärkt werden muss. Vor dem Hintergrund der digitalen Transformation haben KMUs nun die Möglichkeit ein externes Controlling, im Vergleich zu einem internen Controlling, zu forcieren. Dadurch gelangen sie zu neuen Geschäftsmodellen, optimierten Prozessen sowie höherer Tragweite im Wettbewerb. Diese Arbeit versucht die Frage zu beantworten, ob die digitale Transformation für KMUs das externe Controlling vorantreiben kann. Zur Beantwortung dieser Frage wird als Methode eine SWOT Analyse angewandt werden. Zu Beginn werden mit Hilfe der SWOT Analyse Stärken und Schwächen der KMUs bezüglich eines externen Controllings aufgezeigt und anschließend Chancen und Risiken, die die digitale Transformation für ein externes Controlling mit sich bringt, für die nächsten drei bis fünf Jahre dargelegt werden. Zudem werden Managementimplikationen durch den Einsatz der digitalen Transformation aufgezeigt sowie Strategien für das Management abgeleitet werden.

I. Scope and Aim

Small and medium-sized enterprises (SMEs) play an important role in the Germany economy. This can be determined by data analyses of the IFM Bonn (Institute for German Mittelstand Research) from 2015 that reveal the following: Classification of 99.6 per cent of all German enterprises as SMEs; generation of 35.0 per cent of the turnover of all German enterprises; employment of 16.85 million employees, which is 58.5 per cent of all employees who are subject to social insurance contributions, and occupation of 81.8 per cent of apprentices [1]. Although the economic situation is significantly co-determined by SMEs, studies show that there is a lack in management accounting [2]. Management accounting has slightly developed in SMEs. One reason might be that during the past two decades SMEs have had their focus rather on core business such as the production of goods and less on the analysis of data and figures. Another reason might be the comparison of benefits and costs of the implementation of management accounting. Such considerations caused weaknesses in management. The current situation is alarming and data analyses of the German Credit Insurance Hermes reveal that 79 per cent of company crises can be traced back to a lack in management accounting [2].

Thus, in small and medium-sized enterprises, management accounting has to be supported. Such an approach can be forced by dint of external management accounting that is – compared to internal management accounting – by far at a reduced rate.

In the light of Digital Transformation, in small and medium-sized enterprises external management accounting can hence be launched and enhanced.

However, the question arises as to how small and medium-sized enterprises can reach a valuable external management accounting. There is only one solution: Small and medium-sized enterprises have to be conditioned for Digital Transformation. Digital Transformation will enable SMEs amongst others to analyze unprecedented data and to minimize managerial issues.

The aim of this paper is to answer the central question if Digital Transformation can advance for SMEs an external management accounting. By means of this question, we will argue where we see the advantages of Digital Transformation for SMEs in view of an external management accounting.

Initially, we will outline characteristics of SMEs and afterwards explicate the terminology *external management accounting*, which will provide the basis to understand the importance of SMEs in terms of external management accounting. Secondly, with the aid of a SWOT analyses, today's strengths and weaknesses of SMEs in view of an external management accounting will be highlighted. Subsequently, chances and risks for SMEs caused by Digital Transformation in view of external management accounting will be revealed for the next three to five years. Finally, management implications caused by Digital Transformation in SMEs from the perspective of external management accounting will be formulated and dos and don'ts for SMEs will be deduced.

II. Characteristics of Small and Medium-sized Enterprises

SMEs are oftentimes compared to large enterprises in order to make a distinction. As a result, SMEs are in general characterized by qualitative and quantitative features. Characteristics such as turnover, number of employees, total assets and output can be ascribed to the quantitative features of SMEs [4]. The distinction concerns primarily the number of employees and turnover. Besides, German mid-sized enterprises can be narrowed down to 10 to 499 employees and to less than 50 million euros [4; 5]. Comparatively, small-sized enterprises can be characterized by employees up to a number of 9 employees and a turnover up to one million euros [5]. On the other hand, qualitative characteristics can be distinguished, amongst others, by the close link of enterprise and entrepreneur, however, likewise by a lack of leadership abilities, little planning and insufficient information systems [6]. Long-term enterprise success and securing the enterprise's existence can also be ascribed to the qualitative characteristics. Innovative power, close customer orientation and flexibility round out SME's qualitative characteristics [7].

III. Definition of External Management Accounting

In terms of external management accounting, it can be ascertained that it implies the performance of management accounting issues by dint of a person who is located outside an enterprise. Moreover, an external management accountant can be characterized by an independent employer-employee relationship [8]. In most instances, external management accounting consultants are mandated to implement management accounting instruments such as operating figures or control systems that contribute to the development of an enterprise. Thus, know-how concerning corporate planning, capital budgeting, implementation of internal control systems and interpersonal skills play a significant role for external management accountants/consultants [3]. In any case, it should be noted that the function of external management accounting is to raise efficiency and, in conjunction with this, external management accounting can be advantageous in cases where new systems have to be implemented and budgets have to be prepared [9].

IV. Methodology

A SWOT analysis will be applied, which reveals strengths, weaknesses, opportunities and threats of Digital Transformation. The reason behind this is to deduce perceptions obtained by this analysis for SMEs in view of an external management accounting. The SWOT analyses is the planned research design. At this stage of the dissertation project an evaluation of conducted studies will, however, serve as own contribution, respectively as intermediate result. Such a SWOT analysis is extensive, however, this paper allows to present a little cutout and hence some researches of the four analyses fields will be presented subsequently. Digital Transformation entails the following strengths: First, the increase of internal data volume, which enables to structure processes more transparent and to investigate them in view of weak points [10]. Moreover, Digital Transformation will facilitate automatic real-time analyses, which offer enterprises the opportunity to investigate processes in view of efficiency or issues and to remedy shortcomings in short-term [11]. Besides, forecast errors will be reduced, which provides a basis for improved decision-making on the part of leadership and enterprise management [12]. Among other things, items within the enterprise round out the strengths of Digital Transformation. Thus, the enterprise will have the possibility to interchange data and it will be able to re-organize in real-time [13]. Thus, it will be possible to reduce idle times and to increase flexibility. In view of the weaknesses it can be ascertained that there is an existence of inconsistent standards [14]. Such existing standards are merely isolated applications [11]. Suchlike given conditions can restrict projects related to Digital Transformation. Moreover, Digital Transformation will be accompanied by huge data traffic [15]. However, suchlike huge data can result in reservoirs of unused data which in turn can complicate the day-to-day management. Though, Digital Transformation entails chances such as cloud computing and hence enterprises will be able to use the latest infrastructure [16]. In this regard enterprises will have the opportunity to pass investment and maintenance costs on to the cloud maintainer and to work with the latest hardware and software. Besides, by dint of 3-D-Printing enterprises will have the chance to utilize completely new production processes and hence will be able to produce more environmentally friendly and to keep pace with changing trends [17]. However, Digital Transformation bears risks such as standard risks [18]. That is connected with the fact that many enterprises utilize technologies from various time periods and on this account it will hardly be possible to ensure a machine-to-machine communication, whereas suchlike automated communication could contribute to react fast to environmental changes.

Finally, in the context of suchlike researches, the question arises in how far small and medium-sized enterprises can apply them to an external management accounting.

V. intermediate Results

We believe that after presenting the whole SWOT analysis, we will be able to deduce five dos and five don'ts for SMEs regarding external management accounting due to management implications caused by Digital Transformation. Such dos and don'ts should contribute to reconsidering business models and rationalization processes. Subsequently, an outline of three study results with small and medium-sized enterprises will be revealed as initial validity of above-mentioned claims and in addition a brief evaluation will be conducted as own

contribution. The results of the first study are related to a conducted survey from October to December 2014 with 36 medium-sized enterprises. The results reveal significant improvements within internal and external reporting and forecasting [19]. Moreover, there is an increase in speed and quality; manual intervention as well as error rates are minimized and processes can now be investigated in view of weak points or issues [19]. Besides, the preparation time of monthly and quarterly reports could be reduced and finalized reports are available within two days, whereas data collection is already completed after only one day [19]. On the basis of automated and improved data, which besides contributes to reduce forecast errors, management accountants assume a function as consultant and advise leadership and management which in turn provides a basis for well decision-making on the part of the management [19]. In view of risks most of the management accountants of small and medium-sized enterprises ascribe less importance to automated data transfer from previous systems and hence do not consider an upgrade of systems as necessary [19]. The results of the second study are related to a conducted survey from February to May 2017 with 87 management accounting managers and specialists. The study was about the question at which evolution stage enterprises are in view of Digital Transformation. The results of this study show amongst others that only few enterprises use extended Data (Big Data) as basis for decision-making, however, 9% of the enterprises which make use of it, realize a value-added contribution [20]. Moreover, in view of weaknesses the study reveals that half of the interviewed enterprises account their system landscape as being below 40% which seems at present less to meet the requirements of Digital Transformation [20]. Chances in view of cloud computing can be realized by dint of the third study where 317 enterprises have been interviewed in December 2016. The results of this study show that more than 60% of the interviewed enterprises are at the initial stage of cloud computing or started cloud computing as pilot project and thus expedite cloud computing [21]. The results of the conducted studies show that claims made by dint of the SWOT analyses are proved to be true and that small and medium-sized enterprises have to look at Digital Transformation in its entirety when they consider to be successful in Digital Transformation. However, most of the enterprises have not quite decided yet to implement Digital Transformation, although it is not a choice which has to be made. Digital Transformation should be considered as obligatory and there is no way around it. In terms of external management accounting Digital Transformation will have a great impact on this field of functions as the conducted studies show that there are several impediments within internal management accounting. There are for instance insufficient automated processes which require manual input which is extensive, late data reporting within associated enterprises or errors in budgeting. Suchlike barriers will lead to foster an external management accounting which will moreover by dint of Digital Transformation be fast, valuable and efficient. One reason for this claim is that external management accountants are just adapting advanced competences regarding Digital Transformation and that they already have all attributes to implement successfully new business models such as Digital Transformation.

REFERENCES

[1] IfM (Institut für Mittelstandforschung, Bonn) (2015). *Volkswirtschaftliche Bedeutung der KMU*. Retrieved 2017, September 3, from http://www.ifm-bonn.org/statistiken-/mittelstand-im-ueberblick-/#accordion=0&tab=0.

[2] Euler Hermes: Wirtschaft Konkret Nr. 414. *Ursachen von Insolvenzen*, 2006, p. 20.

[3] Pössl, W., *Controlling und externe Beratung*, Vahlen: München, 1991, pp. 53-54.

[4] Mäder, O. B. & Hirsch, B.: Controlling – Strategischer Erfolgsfaktor für die Internationalisierung, in: Keuper, F., Schunck, Hendrik A. (Hrsg.), *Internationalisierung deutscher Unternehmen*. Strategien, Instrumente und Konzepte für den Mittelstand. 2. Auflage, Wiesbaden: Gabler, 2011, pp. 175-176.

[5] Schauf, M., *Unternehmensführung im Mittelstand.* Rollenwandel kleiner und mittlerer Unternehmen in der Globalisierung. 2. Auflage, München, Mering: Hampp, 2009, pp. 4-5.

[6] Schneider, A., *Erfolg durch E-Learning in kleinen und mittleren Unternehmen.* Voraussetzungen für das selbstgesteuerte Lernen, Hamburg: Diplomica, 2014, p. 8.

[7] Broich, D. J.: *Mitarbeiterbindung in KMU.* Analyse von Instrumenten und Maßnahmen, Hamburg: Igel, 2015, p. 18.

[8] Horak, C., *Controlling in Non-Profit Organisationen.* Erfolgsfaktoren und Instrumente. 2. Auflage, Wiesbaden: Gabler, 1995, p. 128.

[9] Becker, J./Winkelmann, A., *Handelscontrolling.* Optimale Informationsversorgung mit Kennzahlen. 3. Auflage, Berlin, Heidelberg: Springer, 2014, p. 439.

[10] Baumöl, U. /Berlitz, P.- D., *Big Data als Entscheidungsunterstützung: Herausforderungen und Potenziale*. Controlling und Big Data. In: Gleich, R./Grönke, K., Kirchmann, M., Leyk, J. (Hrsg.), Controlling und Big Data, Anforderungen, Auswirkungen, Lösungen, Freiburg, München: Haufe, 2014, pp. 159-176.

[11] Bachmann, R./Kemper, G/Gerzer, T., *Big Data - Fluch oder Segen?* Unternehmen im Spiegel gesellschaftlichen Wandels, Heidelberg: mitp, 2014, p. 40 & 187.

[12] Heuberger-Götsch, O., Der Wert von Daten aus juristischer Sicht am Beispiel des Profiling. In: Fasel, D./Meier, A. (Hrsg.), *Big Data*. Grundlagen, Systeme und Nutzungspotentiale, Wiesbaden: Springer Vieweg, 2016, pp. 83 & 86.

[13] Kleinemeier, M., Von der Automatisierungspyramide zu Unternehmens-steuerungsnetzwerken. In: Bauernhansl, T./Ten Hompel, M./Vogel-Heuser, B. (Hrsg.), *Industrie 4.0 in Produktion, Automatisierung und Logistik*, Wiesbaden: Springer Vieweg 2014, pp. 571-580.

[14] Damm, W. /Achatz, R. /Beetz, K. /Broy, M. /Daembkes, H. /Grimm, K. / Liggesmeyer, P., Nationale Roadmap Embedded Systems. In: M. Broy (Hrsg.), *Cyberphysical Systems. Innovation durch Software-intensive eingebettete Systeme.* „acatech diskutiert", Heidelberg: Springer 2010, pp. 67-136.

[15] Fasel, D. /Meier, A., Was versteht man unter Big Data und NoSQL?, In: Fasel, D. /Meier, A. (Hrsg.), *Big Data*, Wiesbaden: Springer Wieweg 2016, pp. 3-16.

[16] Brinkmann, F.-M. /Fetai, I. /Schuldt, H., *SLA - basierte Konfiguration eines modularen Datenbanksystems für die Cloud*. In: Fasel, D. /Meier, A. (Hrsg.), Big Data, Wiesbaden: Springer-Wieweg, 2016, pp. 179-195.

[17] Cole, T., *DIGITALE TRANSFORMATION*, München: Vahlen, 2015, p. 149.

[18] Brühl, V., *Wirtschaft des 21. Jahrhunderts*, Wiesbaden: Springer Gabler 2015, pp. 106 & 107.

[19] PWC (PriceWaterhouseCoopers International) (2015). *Digital Cont-rolling*, Digitale Transformation im Controlling, Retrieved 2017, October 30, from https://www.pwc.de/de/digitale-transformation/assets/pwc-studie-digitale-transformation-im-controlling.pdf.

[20] Camelot (Management Consultants) (2017). *Digitalisierung als Treiber der Controlling Transformation*, Sind Sie bereit für einen Paradigmenwechsel? Retrieved 2017, October 30, from https://www.camelot-mc.com/de/study/digitalisierung-als-treiber-der-controlling-transformation/.

[21] IDC, *Cloud Computing in Deutschland 2017*, Cloud-Transition als Erfolgsfaktor der digitalen Transformation, Retrieved 2017, October 30, from http://idc.de/de/research/multi-client-projekte/cloud-computing-in-deutschland-2017/cloud-computing-in-deutschland-2017-projektergeb-nisse.

This page is intentionally left blank

Accretion strategies referencing start-ups inclined to affiliation

Emanzipationsstrategien für übernahmeaffine Start-Ups

Gero Peter Gunkel

Institut für Marketing, Carl-Friedrich-Gauß-Fakultät, Department Wirtschaftswissenschaften,
Technische Universität Braunschweig, Deutschland

***Abstract* —** Regarding successful technology-oriented start-ups, the implementation of benefiting actions may be economically relevant with respect to the causation of intended takeover bids by established corporations. Hence the evaluation of the respective mechanisms is within the scope of this study. This context involves the disclosure of dimensions that allow for the drafting of specific options for operational emancipation. The focus will be set on start-ups that aspire to a specialisation on clients in the field of public authorities and governmental agencies (B2A – Business to Administration). The contribution concludes with a critical acclaim on the perspectives in the presented spheres of influence.

***Zusammenfassung* —** Für erfolgreiche technologie-orientierte Start-Ups kann die Implementierung von begünstigenden Maßnahmen zur Generierung von Übernahmeangeboten durch größere bzw. etablierte mittelständische Unternehmen von betriebswirtschaftlicher Relevanz sein. Gegenstand der Untersuchung ist daher die Beurteilung entsprechender Mechanismen. In diesem Kontext werden Dimensionen aufgezeigt, anhand derer spezifische Emanzipationsoptionen erarbeitet werden können. Der Schwerpunkt der Betrachtung liegt dabei auf Start-Ups, die eine Spezialisierung auf Auftraggeber aus dem Bereich der öffentlichen Verwaltung (B2A – Business to Administration) verfolgen. Der Beitrag schließt mit einer kritischen Würdigung zur jeweiligen Perspektive in den aufgezeigten Wirkungsfeldern.

I. Einleitung

A. Motivation und Problemstellung

Aufgrund von globalen Trends wie der Digitalisierung von Technologien und der weltweiten Vernetzung werden die Marktstrukturen in der digitalen Wirtschaft immer dynamischer. So gaben 36 % der befragten Start-Ups in einer Studie des deutschen Start-Up-Monitors an, dass das Tempo, mit dem Produkte und Services überholt sind, sehr hoch ist (vgl. [1], S. 71). In diesen Märkten haben Start-Ups großes Entwicklungspotenzial, da sie aufgrund ihrer innovativen Ausrichtung die Nachfrage nach technologischen Neuerungen gut bedienen können. Dabei sind Kooperationen hilfreich, um ein erfolgreiches und wachsendes Unternehmen zu ermöglichen (vgl. ebd., S. 66).

Besonders vor dem Hintergrund öffentlicher Ausschreibungen stellt dabei das Kompetenzmarketing einen wichtigen Aspekt dar. In vielen Ausschreibungen werden Referenzen und Erfahrungen verlangt, die junge Start-Ups nicht bieten können. Da sie allerdings über ein hochinnovatives technisches Know-how verfügen, bietet sich die Kooperation mit einem etablierten Unternehmen an, um eine bessere Ausgangsposition bei der Vergabe der Projekte zu bekommen. Aus diesem Grund sollen die dabei einschlägigen Instrumente und Maßnahmen in diesem Beitrag näher betrachtet und kritisch hinterfragt werden.

B. Forschungsleitfragen

Die konkret formulierten Forschungsleitfragen, die in diesem Kontext explizit beantwortet werden sollen, lauten:

- Welche Maßnahmen eignen sich für die Darstellung und Kommunikation von Kompetenz bei Start-Ups?
- In welcher zeitlichen Dimension können Maßnahmen des Kompetenzmarketing eingeordnet werden?
- Welche Maßnahmen generieren den höchsten Nutzen?
- Wie beeinflusst die B2A-Beziehung die Emanzipationsstrategie?
- Wie können abgeleitete Maßnahmen effektiv implementiert werden?

II. Grundlagen

Mit Blick auf ein einheitliches Verständnis werden im Folgenden die zur thematischen Einordnung erforderlichen Definitionen und Begriffe erläutert. Zudem werden die Fragestellung und der untersuchte Bereich von anderen Forschungsfeldern thematisch abgegrenzt.

A. B2A-Beziehung

Insgesamt wird auf die Marketingstrategien von Technologie-orientierten Start-Ups bezüglich ihrer Kundensegmente näher eingegangen. Um im Folgenden auf eine einheitliche Definition zurückgreifen zu können, wird für die B2A-Beziehung das Verständnis einer Bezeichnung für Leistungsaustauschprozesse zwischen Unternehmen und öffentlichen Institutionen herangezogen (vgl. [2], S. 27).

B. Begriffsklärung

Generell übernehmen Start-Ups zwei wichtige Funktionen in der Erforschung von Technologien und lokalen Netzwerken. Sie tragen wesentlich dazu bei, neue Technologien und Möglichkeiten hervorzubringen, die von größeren Unternehmen nicht betrachtet werden. Kleinen Unternehmen wie Start-Ups gelingt dies besser, da sie engere Verbindungen zu innovativen Institutionen pflegen und sich Forschungswissen innerhalb ihres Netzwerks in der Regel schneller verbreiten kann (vgl. [3], S. 22).

Die Leistungsfähigkeit technologie-orientierter Start-Ups wird durch drei Komponenten beeinflusst; die unternehmerische Ausrichtung, die technologischen Ressourcen bzw. Fähigkeiten, und schließlich die finanziellen Mittel bzw. die Kapitalausstattung. Ein weiterer wichtiger Einfluss auf das Leistungsvermögen des Start-Ups ist außerdem die Interaktion zwischen internen Ressourcen und Netzwerken wie partnerschaftlichen Verbindungen. Besonders Kapitalbeteiligungen anderer Unternehmen helfen dabei dem Start-Up, sein technologisches Können für das Schaffen absatzfähiger Produkte einzusetzen (vgl. [4], S. 633 f.).

C. *Thematische Abgrenzung*

Das Ziel der Marketingmaßnahmen soll eine Vermittlung der eigenen Kompetenz sein, um durch Kooperationen und Übernahmen durch etablierte Unternehmen eine bessere Ausgangsposition bei öffentlichen Ausschreibungen zu erlangen. Daher werden im Folgenden ausschließlich Marketingmaßnahmen betrachtet, die die Kompetenz des Start-Ups vermitteln. Maßnahmen und Instrumente zur Vermarktung von Produkten an Endkunden sind in diesem Kontext nicht relevant.

Mit Kooperationen und Übernahmen durch etablierte Unternehmen sind Unternehmenskonstellationen gemeint, die die Vermarktung und Ausgangsposition des Start-Ups bei öffentlichen Ausschreibungen verbessern. Kooperationen durch Investoren oder reine Kapitalgeber sind hierbei nicht beabsichtigt.

III. Methodik der Emanzipationsentwicklung

Bei dem Vorgang, innovative Geschäftsideen zu verwirklichen, spricht man allgemein von „Entrepreneurship“. Dieser Begriff kann folglich als „Prozess der Identifikation und Umsetzung neuer unternehmerischer Gelegenheiten in marktfähige Produkte und Dienstleistungen“ definiert werden ([5], S. 9).

Insbesondere bei der Akquise von öffentlichen Ausschreibungen sind die hohen Anforderungen oftmals ein Hindernis für junge Unternehmen, da für die Teilnahme an diesen Ausschreibungen in den meisten Fällen Referenzen vergleichbarer Projekte, finanzielle Mittel sowie wirtschaftliche Ressourcen vorausgesetzt werden (vgl. [6]). Vor diesen Anforderungen schrecken die meisten Start-Ups zurück, sodass nur ein geringer Anteil der öffentlichen Aufträge direkt an Start-Ups vergeben wird. Zu berücksichtigen ist dabei, dass die unmittelbare und persönliche Beteiligung der Unternehmen an der Ausführung des jeweiligen Auftrags erforderlich ist (vgl. [7]). Dies verstärkt den Wunsch vieler Start-Ups, von etablierten Unternehmen in deren Unternehmensstruktur integriert zu werden, um so einer möglichen Benachteiligung bei öffentlichen Ausschreibungen zu entgehen. Hieraus ergibt sich die Frage, wie Kompetenz optimal vermittelt werden kann, um andere Unternehmen als Partner zu gewinnen.

IV. Konzeptionelle Dimension und Kriterien

Gegenstand dieses Kapitels ist zunächst die Darstellung der Spezifika in B2A-Märkten. Hieran anknüpfend werden diese in einem gesamtwirtschaftlichen Kontext in Verbindung gebracht.

A. *Spezifika des B2A-Umfeldes*

Der Begriff „Business to Administration“ beschreibt eine Kommunikations- und Geschäftsbeziehung zwischen einem Unternehmen und Verwaltungen, Behörden und Einrichtungen der öffentlichen Hand. Synonym für eine solche Beziehung wird in der Literatur auch die Bezeichnung „Business to Government (B2G)“ verwendet (vgl. hierzu z.B. [8]).

Die Anwendungsbeispiele für B2A-Geschäftsbeziehungen sind sehr weit gefächert. Sie reichen von der Abwicklung von Steuerangelegenheiten von Unternehmen über Subventions- und Hilfeleistungen bis hin zur Abwicklung von Geschäften zwischen öffentlichen Einrichtungen und Unternehmen (vgl. [9]). Der Kaufprozess in B2A-Geschäftsbeziehungen wird dabei typischerweise durch Ausschreibungen mit detaillierten Anforderungen initiiert. Die Informationen und Unterlagen werden in der Anbahnungsphase meist komplett elektronisch bereitgestellt. Dies kann auf Plattformen wie etwa „vergabe-24.de“, „bund.de“ oder „evergabe-online.de“ erfolgen. Generell müssen Unternehmen sich dabei an strikte Vorgaben halten, um nicht den Ausschluss vom Wettbewerb zu riskieren (vgl. [10], S. 14 ff.).

Das Vergabeverfahren erfolgt größtenteils voll elektronisch. Die komplexe Bewertungsphase kann nicht vollständig elektronisch ablaufen und erfordert einen personellen Eingriff in ausgewählten Prozessschritten. Der Preis für Produkte und Dienstleistungen wird in B2A-Märkten hauptsächlich über Ausschreibungsvorgaben ermittelt (vgl. [11]).

B. *Marktwirtschaftliches Geschäftsumfeld*

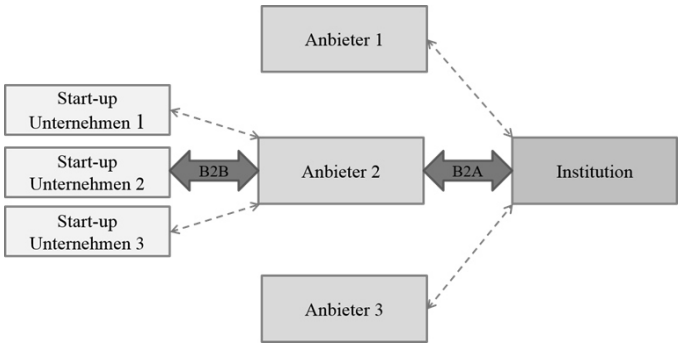


Abb. 1. Marktwirtschaftliches Geschäftsumfeld

In Abbildung 1 werden die Geschäftsbeziehungen in einen marktwirtschaftlichen Gesamtkontext eingebettet. Ausgangspunkt ist dabei eine öffentliche Institution, welche in der Praxis eine Verwaltungseinrichtung mit einem Bedarf an einem bestimmten Produkt oder einer Dienstleistung sein kann. Die Nachfrage wird mittels einer Ausschreibung dem Markt mitgeteilt. Diese Ausschreibungsunterlagen stehen bestimmten Unternehmen zur Verfügung, die potenziell in einen Wettbewerb zur Erbringung der Leistung treten können. Dabei ist zu beachten, dass das Ausmaß des Wettbewerbs stark mit den Anforderungen und Eigenschaften der nachgefragten Leistung korreliert. Handelt es sich etwa um besondere Anforderungen wie kleine Losgrößen oder spezifisches Know-how, so sind Kooperationen für das anbietende Konsortium unabdinglich, um wettbewerbsfähig zu sein und ein Angebot erstellen zu können. In dieser Betrachtung sind dabei besonders Kooperationen mit Start-Up-Unternehmen von Relevanz, welche sich durch technologisches Know-how oder spezifische Produkte auszeichnen. Diese Start-Ups stehen in einer gefestigten Kooperation und B2B-Beziehung zum Anbieter, weswegen der Wettbewerb zwischen den Start-Up-Unternehmen hier nicht betrachtet wird. Ziel des Start-Ups ist die Intensivierung der Beziehung und die Generierung eines Übernahmeangebots.

V. Ableitung von Optionen und Maßnahmen

Kern dieses Kapitels ist die Auseinandersetzung und Ableitung von Erkenntnissen aus den vorangegangenen Betrachtungen sowie die Entwicklung geeigneter Emanzipationsstrategien.

A. *Darstellung möglicher Wirkungsfelder*

Die Ableitung möglicher Wirkungsfelder erfolgt aus einer vornehmlich marktwirtschaftlichen Perspektive. Eine wichtige Rahmenbedingung ergibt sich dabei aus dem Informationsver-

halten zwischen den Unternehmen. Da das potenziell integrierende Unternehmen auf einen Kunden trifft, der informationsbasierte rationale Kaufentscheidungen trifft, muss ein transparenter Informationsfluss gewährleistet werden. Das Unternehmen in der B2A-Beziehung muss daher bei Teilnahme an Ausschreibungen detaillierte Auskünfte über das Produktportfolio geben können und die Verfügbarkeit von Informationen über die gesamte Wertschöpfungskette hinweg gewährleisten können.

Dem Start-Up müssen darüber hinaus auch Informationen über die Ausschreibungsvorgaben und technischen Anforderungen bereitgestellt werden. Gezielte Informationspolitik kann dabei die Bindung der beiden Unternehmen nachhaltig stärken. Das anbietende Unternehmen muss darüber hinaus im Bereich der Informationstransparenz viel Eigeninitiative aufbringen und nach öffentlichen Ausschreibungen recherchieren (vgl. [10], S. 45 f.).

B. Implementierung

Die horizontale Achse der in Abbildung 2 dargestellten Matrix beschreibt die Wirkungsfelder für mögliche Strategien, die aus den Rahmenbedingungen der Geschäftsbeziehung abgeleitet werden können. Die vertikale Achse ermöglicht eine Gruppierung der Strategien nach zeitlicher Perspektive. Es wurde eine Einteilung in kurzfristige, mittelfristige und langfristige Strategien gewählt.

Kurzfristige Strategien sind dabei operativen Planungen zuzuordnen und haben einen zeitlichen Rahmen von sechs Monaten bis zu drei Jahren. Mittelfristige Strategien beziehen sich auf taktische Planungen im Zeithorizont von einem bis fünf Jahren. Strategische Planungen basieren auf langfristigen Strategien, die sich ggf. auf zwei bis zu zwanzig Jahre beziehen, jedoch in der Regel auf etwa fünf Jahre ausgelegt sind (vgl. hierzu [12], S. 52 ff.).

	Wirkungsfeld 1: Kooperation	**Wirkungsfeld 2: Information**	**Wirkungsfeld 3: Kommunikation**
Kurzfristig (ca. 1 Jahr)	Ähnlichkeiten zwischen den Transaktionspartnern auf personeller Ebene nutzen	Abstimmung von Anforderungsniveau und Fachkompetenzen	Aufgeschlossenheit, Einfühlungsvermögen und Offenheit bei Personalauswahl beachten
Mittelfristig (ca. 3 Jahre)	Kundenkontaktpersonal für auftragsbezogene Wahrnehmung des Kunden sensibilisieren	Förderung von Fach-, Methoden- und Personalkompetenzen	Förderung von Fach-, Methoden-, Sozial-, Kommunikations- und Personalkompetenzen
Langfristig (ca. 5 Jahre)	Ähnlichkeiten zwischen den Transaktionspartnern auf Unternehmensebene schaffen	Austausch von Information zur langfristigen Erhöhung und Sicherung des Wissenstands	Unternehmenskultur einer offenen Kommunikation zwischen den Mitarbeitern – Kommunikationspolitik

Abb. 2. Matrix zur Strategieimplementierung

Im Wirkungsfeld des Kooperationsverhaltens können kurzfristig Ähnlichkeiten zwischen den Transaktionspartnern auf personeller Ebene genutzt werden. Hierbei sollte analysiert werden, wer dem Einkaufsgremium des Erwerbenden angehört und an der Entscheidung beteiligt ist. Die Entscheidungsträger sollten dabei direkt adressiert und als Befürworter einer Übernahme akquiriert werden. Bezüglich der Schaffung von Ähnlichkeiten kommt der Personalauswahl demnach eine zentrale Rolle zu (vgl. [13], S. 18 ff.).

Mittelfristig soll das Kundenkontaktpersonal für eine auftragsbezogene Wahrnehmung des Kunden sensibilisiert werden. Das bedeutet, dass das Kundenkontaktpersonal darauf geschult oder danach ausgewählt werden soll, die Erwartungen des Kunden richtig einschätzen zu können. Es soll in der Lage sein, sich in die Wahrnehmung des Nachfragerunternehmens hineinzuversetzen. Das Kundenkontaktpersonal verfügt demnach über die erwünschte soziale Fähigkeit eines kundenspezifischen Einfühlungsvermögens sowie über die fachliche Kompetenz, das Kundenproblem sachlich korrekt auffassen zu können (vgl. [14], S. 285).

VI. Kritische Würdigung

Im Rahmen dieser Studie wurden im Bereich des marktwirtschaftlichen Umfelds vereinfachende Annahmen getroffen. Im Wirkungsfeld der Kommunikation und Information könnte das Start-Up bei der Strategieimplementierung somit auf Schwierigkeiten treffen. Diese Wirkungsfelder stützen sich auf die Erkenntnis, dass im B2B-Sektor Einkaufsentscheidungen in multipersonellen Gremien rationaler getroffen werden, als auf B2C-Märkten. In welchem Maß die Buying Center jedoch tatsächlich rationaler handeln und entscheiden, steht im wissenschaftlichen Diskurs (vgl. hierzu [15], S. 407).

Durch die Verwendung von Annahmen und in der Literatur bislang unzureichende Belege von Ursache-Wirkungs-Beziehungen kann dieser Beitrag lediglich einen Anhaltspunkt für mögliche Strategieimplementierungen und Wirkungsfeldanalysen für Start-Ups darstellen. Diese könnten durch Methoden der quantitativen empirischen Sozialforschung genauer analysiert werden. Grundsätzlich ist des Weiteren festzuhalten, dass aufgrund der Aktualität der Thematik Informationen bezüglich der Situation von Start-Ups in vielen Fällen lediglich online verfügbar sind. Hierbei muss in vielen Fällen die Objektivität der dargebotenen Informationen hinterfragt werden.

VII. Zusammenfassung

A. Handlungsempfehlungen

Wie die B2A-Beziehung die Emanzipationsstrategie des Start-Ups beeinflusst, ist einerseits abhängig davon, um welchen Bedarf an einem Produkt oder einer Dienstleistung es sich ausgehend von der öffentlichen Institution handelt, bzw. welche Anforderungen in Bezug auf die zu erbringende Leistung an das etablierte Unternehmen gestellt werden. Kann das Start-Up genau diese Anforderungen durch spezielles Know-how oder ein spezifisches Leistungsspektrum erfüllen, so kann es diesen Vorteil nutzen, um über eine Kooperation hinaus auf längere Sicht eine Übernahme durch das Partnerunternehmen zu begünstigen.

Andererseits ist dabei der Informationsaustausch innerhalb der B2A-Beziehung wichtig, da das Start-Up detaillierte Informationen über die Leistungsanforderungen sowie generell über die Ausschreibungsvorgaben der öffentlichen Hand vom Geschäftspartner benötigt.

Mit Blick auf die eingangs gestellten Forschungsleitfragen kann aggregiert festgehalten werden, dass ein technologie-orientiertes Start-Up ein mögliches Übernahmeangebot durch ein größeres und etabliertes Unternehmen dann erreichen kann, wenn Maßnahmen des Kompetenzmarketings angewendet werden.

B. Ausblick

In diesem Beitrag wurden unterschiedliche Maßnahmen zur Darstellung und Vermittlung eigener unternehmerischer Fähig-

keiten dargestellt. Hierbei wurden technologie-orientierte Start-Ups betrachtet, die eine Übernahme durch ein etabliertes Unternehmen anstreben, um eine bessere Ausgangsposition im B2A-Segment zu erreichen. Aus Vereinfachungsgründen wurden dabei die unterschiedlichen Branchen der Start-Up-Szene nicht berücksichtigt.

Für die Zukunft stellt sich so die Frage, ob bei einer anzunehmenden hohen Marktdynamik für junge Unternehmen langfristige oder – aufgrund einer besseren Umsetzbarkeit – eher kurzfristige Maßnahmen einen wesentlichen Erfolg versprechen. Diese Frage könnte im Zuge einer statistischen Erhebung analysiert werden.

Auch die Komplexität der angesprochenen Buying Center spielt eine essentielle Rolle. So kann Untersuchungsgegenstand weiterführender Arbeiten schließlich sein, relevante Entscheidungsträger zu identifizieren, und so spezifischere betriebswirtschaftliche Maßnahmen für Start-Ups zu generieren.

LITERATURVERZEICHNIS

[1] Kollmann, T.; Stöckmann, C.: "*Deutscher Start-Up-Monitor 2016*", KPMG Deutschland & Bundesverband Deutsche Startups e.V., Berlin, 2016

[2] Wirtz, B.: "*Gabler Kompakt-Lexikon „eBusiness“*", Verlag Springer, Wiesbaden, 2013

[3] Almeida, P.; Kogut, B.: "The Exploration of Technological Diversity and the Geographic Localization of Innovation", in: "*Small Business Economics*", Vol. 9, S. 21-31, Verlag Springer Science, Heidelberg, 1997

[4] Lee, C.; Lee, K.; Pennings, J. M.: "Internal Capabilities, External Networks and Performance – A Study On Technology-Based Ventures", in: "*Strategic Management Journal*", Vol. 22, S. 615-640, Strategic Management Society, Chicago, 2001

[5] Fueglistaller, U.; Müller, C.; Müller, S.; Volery, T.: "*Entrepreneurship: Modelle – Umsetzung – Perspektiven*", 4. Auflage, Verlag Gabler, Wiesbaden, 2016

[6] Kulnigg, T.; Stalzer, J.: "Vergaberecht – “Startups! Versteckt euch nicht vor öffentlichen Aufträgen”", in: "*der brutkasten – Österreichs Startup- und Innovationsplattform*", Schönherr Legal Services Austria, Wien, 2016

[7] Stalzer, J.: "*Update Vergaberecht – Richtungsweisende Erkenntnisse des EuGH*", Schönherr Publications, Wien, 2016

[8] SoftSelect: "Definition und Erklärung B2A", in: "*Glossar softselect.de – Unser Wissen für Ihren Erfolg*", SoftSelect GmbH, Hamburg, 2017

[9] Scholz, M.: "Business to Administration (B2A) – Electronic Commerce", in: "*competence-site.de*", NetSkill Solutions GmbH, Köln, 2005

[10] Bittner, T.; Drescher, S.; Jost, D.; Müller, U.; Knörle, K.: "*Die öffentliche Hand als Kunde – Aufträge bei Bund, Ländern und Kommunen gewinnen*", 3. Auflage, Industrie- und Handelskammer Auftragsberatungsstelle Baden-Württemberg, Stuttgart, 2013

[11] Rumler, A.: "*Marketing für mittelständische Unternehmen*", Verlag Teia Lehrbuch, Berlin, 2002

[12] Voigt, K.-I.: "*Strategische Unternehmensplanung: Grundlagen – Konzepte – Anwendungen*", Verlag Gabler, Wiesbaden, 1993

[13] Kreutzer, R. T.; Rumler, A.; Wille-Baumkauff, B.: "*B2B-Online-Marketing und Social Media – Ein Praxisleitfaden*", Verlag Springer Gabler, Wiesbaden, 2015

[14] Deckow, F.: "*Vertrauen durch Kompetenzmarketing – Ein ganzheitlicher Ansatz zur Vermarktung von Kontraktgütern*", Deutscher Universitäts-Verlag, Wiesbaden, 2006

[15] Lynch, J.; Chernatory, L.: "The power of emotion – Brand communication in business-to-business markets", in: "*Journal of Brand Management*", Vol. 11, No. 5, S. 403-419, Verlag Springer Science, Heidelberg, 2004

Berlin faces the technology trap 30 yeas after the fall of the wall – the digital transformation influences the follow-up process

Steht Berlin 30 Jahre nach dem Mauerfall vor der „Technologiefalle“? Die Risiken der Digitalisierung für den Nachfolgeprozess!

Dirk Henrich*, Thomas Heupel‡

* FOM Siegen, PhD-Candidate FDIBA Sofia, dirk.henrich@t-online.de

‡ FOM Siegen, Vice-Rector Research, thomas.heupel@fom.de

Abstract **–** On November 9th, 1989 the Berlin Wall fell. Thousands of people flocked to the western part of the city. A year later, the German Democratic Republic (GDR) disappeared from the map. During this phase of entering the free market economy, thousands of small businesses were established in East Germany - particularly in and around Berlin. What is the state of these businesses today? Many of these companies, which were founded roughly 30 years ago, are now facing company succession and are going to undertake this process over the next five years. At the same time, these companies are running the risk of missing out on a major change. They fall into the "technology trap" of company succession. More precisely, they miss out on the shift to „Industry 4.0“ or the change that is associated with "digital transformation". Looking at the process of corporate succession, there are numerous structural breaks. The literature points to the "leadership trap", the "stakeholder trap", the "emotional trap", the "settlement trap" as well as the “modernization- and technology trap". The authors of this paper address this challenge. The aim is to present the various structural breaks of company succession in a theoretical way and to present the "technology trap" as an exposed problem [vulnerability factor] for the East German economy with a focus on Berlin."

Zusammenfassung – Am 9. November 1989 fällt die Berliner Mauer. Tausende Menschen strömen in den Westteil der Stadt. Ein Jahr später verschwindet die DDR von der Landkarte. In der Aufbruchsphase in die freie Marktwirtschaft werden tausende von kleinen Unternehmen im Osten – und vor allem rund um Berlin – gegründet. Und heute? Viele dieser vor rund 30 Jahren gegründeten Unternehmen stehen vor der Unternehmensnachfolge und werden sich diesem Prozess in den nächsten 5 Jahren zuwenden. Einem weiteren wichtigen Wandel verschließen sich diese Unternehmen aber zugleich. Sie tappen in die „Technologiefalle“ der Unternehmensnachfolge . Sie verpassen nämlich den Wandel zur Industrie 4.0 bzw. den Wandel, der mit der „Digitalen Transformation“ einhergeht. Betrachtet man den Prozess der Unternehmensnachfolge, so gibt es zahlreiche Strukturbrüche. Die Literatur kennt hier die „Führungsfalle“, die „Stakeholderfalle“, die „Emotionale Falle“, die „Einigungsfalle“ sowie die Modernisierungs- und Technologiefalle“. An dieser Herausforderung setzten die Autoren des Beitrags an. Ziel ist es, die verschiedenen Strukturbrüche der Unternehmensnachfolge theoretisch darzulegen und die „Technologiefalle“ als exponiertes Problem für die ostdeutsche Wirtschaft – und hier speziell auch Berlins – darzulegen.

I. Ausgangslage der ostdeutschen Wirtschaft

Die deutsche Wirtschaft hat ihre Wettbewerbsfähigkeit seit der Wiedervereinigung Im Jahr 1989 erheblich verbessert: Sie gilt wieder als Vorbild in Europa. Der Aufschwung ist insbesondere ein Verdienst der mittelständischen Unternehmen und ihrer Mitarbeiterinnen und Mitarbeiter [7], [8], [17]. In Ostdeutschland geschah die Transformation der sozialistischen Planwirtschaft hin zu marktwirtschaftlichen Verhältnissen mit einer beeindruckenden Geschwindigkeit und Intensität. Die Bürger haben den Umgestaltungsprozess mit Gründungen und der Bereitschaft zu unternehmerischem Wagnis bewältigt – nicht zuletzt mit der Solidarität der westdeutschen Bundesländer. Bemerkenswert war nicht nur die Gründung von über einer halben Million selbstständiger Existenzen vor allem im Mittelstand, in Handwerksbetrieben und weiteren Dienstleistungsbereichen. Hervorzuheben ist auch die Entwicklung der überwiegend von KMU geprägten ostdeutschen Industrie, die seit 1995 um jährlich rund fünf Prozent gewachsen ist. Diesem Reindustrialisierungsprozess ist es zu verdanken, dass der wirtschaftliche Abstand von Ost- zu Westdeutschland verringert und die Herstellung gleicher Lebensverhältnisse in Ost und West entscheidend vorangebracht wurde. Die ostdeutsche Wirtschaftsleistung hat sich trotz eines erheblichen Bevölkerungsschwundes von 1992 bis heute in etwa verdoppelt [20].

Im Laufe des gewaltigen Transformationsprozesses haben sich jedoch neben Licht- auch Schattenseiten gezeigt: Die Pro-Kopf-Wirtschaftsleistung ist in Ostdeutschland 30 Prozent geringer als im westdeutschen Durchschnitt. Selbst das BIP des wirtschaftsstärksten ostdeutschen Landes Sachsen liegt noch immer 14 Prozent unter dem wirtschaftsschwächsten westdeutschen Bundesland Schleswig-Holstein. Die Arbeitslosigkeit befindet sich zwar auf dem niedrigsten Stand seit der Wiedervereinigung, liegt mit 10,3 Prozent aber weiterhin deutlich über dem westdeutschen Niveau [20].

Ob die Erfolgsstory aber in den nächsten Jahren fortgeschrieben werden kann und ob die aufgezeigten Schwächen gerade in Ostdeutschland an Dominanz verlieren, wird von einem erfolgreichen Nachfolgeprozess abhängen [1], [3], [4]. Viele der in Berlin vor knapp 30 Jahren gegründeten Unternehmen stehen vor der Unternehmensnachfolge. Dies lässt sich alleine schon daraus ableiten, dass viele Menschen zwischen 28 und 35 Jahren sich besonders intensiv für das Thema „Unternehmensgründung" interessieren. Sie haben ausreichend Erfahrung im Erwerbsleben gesammelt und wollen – oftmals frustriert über eine zu geringe Karriereentwicklung – das unternehmerische Geschick selber in die Hand nehmen. Im Alter über 40 Jahren nehmen dann die „Risikobereitschaft" und der „Leistungswille" signifikant ab und viele streben dann eher eine weiterführende Karriere im Anstellungsverhältnis an, da sie oft schon Familie, Haus und gesellschaftliche Verantwortung haben und nurnoch eine definierte Zeit des Tages in die Erwerbsarbeit stecken wollen. Die Gründer von 1989 sind somit heute vielfach zwischen 55 und 65 Jahren alt und suchen nun eine geeignete Nachfolgeregelung [1], [2], [7], [8].

II. Strukturbrüche der Unternehmensnachfolge

Digitalisierung / Internet of things: Da viele der KMU zugleich auch Familien unternehmen sind, ist hier die familieninterne oder auch familienexterne Nachfolge von großer Bedeutung. Gerade aber in Zeiten des demografischen Wandels ist, bei sinkender Geburtenrate, die familieninterne Nachfolge kein Automatismus mehr. Die junge Generation Y begibt sich nur allzu selten in den „Freischwimmprozess" der Unternehmensnachfolge und sieht andere Optionen neben dem „Unternehmertum" als deutlich attraktivere Lebensmodelle an.

Welche Bedeutung daher die „erfolgreiche" Unternehmensnachfolge hat und welche Strukturbrüche durch die Nachfolge innerhalb des Familienunternehmens auftreten können, soll im Rahmen des Beitrags verdeutlicht werden. In Deutschland stehen mehr als 71.000 Unternehmen pro Jahr vor der Aufgabe, einen geeigneten Nachfolger für ihr Unternehmen zu finden. Dabei werden die Probleme, einen passenden Unternehmenseigner für eine Übernahme zu gewinnen, innerhalb der nächsten Jahre aufgrund des demographischen Wandels und dem damit rückläufigen potenziellen Bewerberkreis vermutlich steigen. Der Tatsache, dass für viele der 71.000 Unternehmen wahrscheinlich kein Nachfolger gefunden wird, wird wenig Bedeutung geschenkt. Nicht ohne Grund bedeutet der Generationswechsel insbesondere für kleinere Familienunternehmen eine teilweise größere Gefahr als ein erhöhter Wettbewerbsdruck oder konjunkturelle Schwankungen [17], [18], [21], [8].

III. Digitalisierung als Technologiefalle

Die Digitalisierung ist in mittelständischen Unternehmen in Deutschland noch stark ausbaufähig. Dies belegt die Studie, die das Zentrum für Europäische Wirtschaftsforschung (ZEW) und das infas Institut für angewandte Sozialwissenschaft GmbH im Auftrag der KfW Bankengruppe durchgeführt haben [22], [8].

Nur ein Fünftel der mittelständischen Unternehmen hat digitale Vernetzung von Produkten und Dienstleistungen begonnen und kann damit zu „Vorreitern" gezählt werden.

Auf Basis der Studienergebnisse können die mittelständischen Unternehmen drei Gruppen mit unterschiedlichem Digitalisierungsgrad zugeordnet werden: Knapp ein Fünftel der Unternehmen (19 Prozent) kann zu den „Vorreitern" gezählt werden, etwa die Hälfte der Unternehmen (49 Prozent) befindet sich im „Mittelfeld" und rund ein Drittel (32 Prozent) des deutschen Mittelstands gehört zu den „Nachzüglern" [22]. Unter diesen bestehen Defizite selbst in der grundlegenden digitalen Infrastruktur, wie einer eigenen Webseite oder Enterprise Resource Planning (ERP) Software. Zu dieser Gruppe gehören insbesondere kleine Unternehmen mit weniger als 50 Mitarbeitern. Hierzu zählen also auch die Unternehmen in Berlin – die damit in die Technologie- und Modernisierungsfalle gelangen [22].

IV. Forschungsfragen und Forschungsdesign

Die Autoren haben mit verschiedenen Akteuren in Berlin über das sich abzeichnende Problemfeld gesprochen (Interviews). Im Rahmen des Beitrags sollen Erkenntnisse dieser Analysephase strukturiert dargestellt werden. Es werden unter anderem folgende Fragen behandelt:

- Ist der Wirtschaftsregion Berlin dieses Phänomen bekannt und gibt es konkrete Pläne diese „Technologiefalle" abzumildern. Wie sehen betroffene Unternehmer dieses Thema und spielt es für sie eine Rolle im erfolgreichen Nachfolgeprozess. Ist die Wirtschaftsförderung bzw. die Politik in Berlin mit konkreten Serviceangeboten aktiv unterwegs?
- Können ausdifferenzierte Handlungsempfehlungen für den Nachfolgeprozess abgeleitet werden und wie können diese vor Ort Gehör finden?

Die Vision der digitalen Transformation umfasst verschiedene Komponenten. Zu den technischen Bestandteilen der Industrie 4.0 zählen die Cyber-Physischen-Systeme, wodurch intelligente Objekte, Maschinen und Anlagen entstehen. Wichtig für den reibungslosen und vor allem schnellen Datenaustausch von Produkt zu Maschine und den Maschinen untereinander ist das Internet. Dabei wird zwischen dem Internet der Menschen, dem Internet der Dienste und dem Internet der Dinge unterschieden. Alle drei Bereiche des Internets sind über Cyber-Physische-Systeme miteinander verknüpft und bilden eine große Kommunikationsplattform, über welche Menschen, Maschinen und Produkte fortlaufend miteinander kommunizieren können. Hierbei vernetzt vor allem das Internet der Dinge den Menschen mit Objekten und Systemen - die physische und digitale Welt warden miteinander verknüpft. Neben der reinen technischen Transformation liegen die Chancen des Internets der Dinge in neuen Geschäftsmodellen sowie in der Organisation und Zusammenarbeit der Unternehmen. Die Digitalisierung und intelligente Vernetzung bedeutet die Speicherung der Daten in Echtzeit und jederzeitige Abrufbarkeit über das Internet. Durch die erhöhte Datentransparenz und kontinuierliche Datenverfügbarkeit wird eine neue Möglichkeit der Selbststeuerung geschaffen. Das Internet der Dinge ermöglicht Mitarbeitern in dezentralen Strukturen die für ihren jeweiligen Verantwortungsbereich relevanten Daten unabhängig und selbstständig zu organisieren, auszuwerten und einzusetzen.

Die Fortschritte der digitalen Transformation leisten einen wertvollen Beitrag, die Rahmenbedingungen für die im Behavioral Management Accounting Ansatz und insbesondere im Beyond Budgeting Konzept geforderte Selbststeuerung zu verbessern. Vor diesem Hintergrund stellt sich die zentrale Frage inwiefern der Auftrieb der digitalen Transformation für die zukunftsweisenden Führungskonzepte des Behavioral Management Accounting und Beyond Budgeting durch die nun stärker in das Erwerbsleben rückende Generationen X, Y und Z angenommen und genutzt wird. Wie gestaltet sich die Bereitschaft der Generationen X, Y und Z das Internet der Dinge anzunehmen? Inwiefern werden die Rahmenbedingungen des Internet der Dinge und die Anforderungen der Generationen X, Y und Z an das Arbeitsumfeld zu einem Dissens oder Konsens führen?

LITERATUR:

[1] Albach, H., Freund W., *Generationswechsel und Unternehmenskontinuität – Chancen, Risiken, Maßnahmen*, Bertelsmann Stiftung, 1989.

[2] Baumann, Fitze, Schefer, *Nachfolge im Unternehmen, Eine Herausforderung für Unternehmer, Verwaltungsräte und Familienaktionäre*, Redaktion Suter, USB AG, 2002.

[3] Brun-Hagen, Hennerkes, *Die Familie und ihr Unternehmen*, Campus GmbH, Frankfurt, 2004.

[4] Brüser, J., *Unternehmensnachfolge, Wie sie als Mittelständler den Stab weitergeben*, Cornelsen, Berlin, 2007.

[5] Bullinger H.-J., Kinkel S., "Das Innovationsverhalten der kleinen und mittleren Unternehmen als Basis für künftige inländische Wertschöpfungspotenziale", in: Otmar, F. (Hrsg.), *Mittelstand – Schlüsselakteur im deutschen Innovationssystem, Verl. Wissenschaft & Praxis*, 2009.

[6] Freund, W. *Unternehmensnachfolge in Deutschland – Neubearbeitung der Daten des IfM Bonn*; IfM Bonn (Hrsg.), Jahrbuch zur Mittelstandsforschung 1/2004, Deutscher Universitäts-Verlag, 2004.

[7] Hering, Th. / Olbrich, M., "Unternehmensnachfolge", in: Köhler, R. / Küpper, H.-U./ Pfingsten, A. (Hrsg): *Handwörterbuch der Betriebswirtschaft*, 6. Aufl., Stuttgart, 2007.

[8] Huber, Kirchdörfer und Rotfuß, Schmidt, "Unternehmensnachfolge in Familienunternehmen", in: Schröder, M., Westerheide, P. (Hrsg.), *ZEW Wirtschaftsanalysen, Band 97 der Schriftenreihe des ZEW-Zentrum für europäische Wirtschaftsforschung,* Mannheim, Nomos Verlagsgesellschaft Baden-Baden, 2010.

[9] Huber, H.-G., Sterr-Kölln, H., *Nachfolge in Familienunternehmen, Den Generationswechsel erfolgreich gestalten – Ein Orientierungsbuch für Unternehmerfamilien und ihre Berater*, Schäffer-Poeschel, 2006.

[10] Kempert, W.: *Praxishandbuch für die Nachfolge im Familienunternehmen – Leitfaden für Unternehmer und Nachfolger*, Wiesbaden, 2008.

[11] Letmathe, P., Hill, M., "Strukturbrüche der Unternehmensnachfolge" in: *Zeitschrift für Betriebswirtschaft*, 76 Jg., S.1113-1138, Gabler, 2006.

[12] Martin, A., Martin, T., "Einflussfakoren und Gestaltungsdimensionen bei der Unternehmensnachfolge im Mittelstand", in: Brauweiler, H.-C- (Hrsg.), *Unternehmensführung heute, München,* S. 215-234, 2008.

[13] Menke, M., *Planung der Unternehmensnachfolge*, 2. Aufl., Stuttgart, 1998.

[14] Moog, P., Kay, R., Schlömer-Laufen, N., Schlepphorst, S., *Unternehmensnachfolge in Deutschland – Aktuelle Trends*, IfM Materialien Nr. 216, Bonn, 2012.

[15] Moog, P., Mirabella, D., Schlepphorst, S., "Owner Orientation and Strategies and their Impact on Family Business", in: *International Journal of Entrepreneurship and Innovation Management*, Jan 2011, Vol. 13, Issue 1, pp. 95-112.

[16] Röhl, K.-H., *Die Zukunft der Familienunternehmen in Deutschland, Potenziale und Risiken in der globalen Wirtschaft*, Deutscher Instituts-Verlag GmbH, 2008.

[17] Schlömer, N., Gude, H., Kay, R., "Unternehmensnachfolge im Mittelstand: Auswirkungen der Vor- auf die Nach-Übergabephase", in: KfW, Creditreform, IfM, RWI, ZEW (Hrsg): *Mittelstandsmonitor 2008 – Jährlicher Bericht zu Konjunktur- und Strukturfragen kleiner und mittlerer Unternehmen*, Frankfurt am Mainz, 2008.

[18] Stephan, P., *Nachfolge im mittelständischen Familienunternehmen – Handlungsempfehlungen aus Sicht der Unternehmensführung*, Diss., Wiesbaden, 2002.

[19] Schulte, R., Wille, C., "Unternehmensnachfolgeprozess in Klein- und Mikrounternehmen – Ergebnisse einer empirischen Untersuchung" in: Brost, Faust, Thedeus (Hrsg.) *Unternehmensnachfolge im Mittelstand*, 1. Auflage, Bankakademie-Verlag GmbH Frankfurt, 2005.

[20] Tiefensee, Wolfgang, "*25 Jahre Mauerfall"*, in: *Der Mittelstand*, 1/2014, Unternehmermagazin des BVMW, Herausgeber: BVMW: Bundesverband mittelständische Wirtschaft, Unternehmerverband Deutschlands e. V., Berlin, 2014.

[21] Wottawa, H., "Psychologische Aspekte bei der Entscheidung für eine Unternehmensnachfolge", in: Brost, H., Faust, M. Thedens, C. (Hrsg.), *Unternehmensnachfolge im Mittelstand*, Frankfurt/Main 2005, S. 393-416.

[22] Saam, M.; Viete, S.; Schiel, S. (2016): *Digitalisierung im Mittelstand: Status Quo, aktuelle Entwicklungen und Herausforderungen*, Forschungsprojekt im Auftrag der KfW Bankengruppe, Mannheim, 2016. Provided in Cooperation with: ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

This page is intentionally left blank

Do Digital Transformation and generation change from baby boomer to digital natives enable Beyond Budgeting a new chance?

Erhält das Beyond Budgeting vor dem Hintergrund der Digitalen Transformation sowie des anstehenden Generationenwechsels von Babyboomern zu Digital Natives eine neue Chance?

Sonja Schönauer*, Christian Roth†, Thomas Heupel‡

* FOM Siegen, PhD-Candidate FDIBA Sofia, schoenauer.sonja@t-online.de
† FOM Siegen, PhD-Candidate FDIBA Sofia, roth.christian@gmx.net
‡ FOM Siegen, Vice-Rector Research, thomas.heupel@fom.de

***Abstract* —** The integration of behavioral aspects into management sciences is discussed in the anglo-saxon area more intensively over the last years. Also, the adaption of the new approach 'behavioral management' in Germany hasn't reached a higher level. Nevertheless this new management accounting framework gains a higher reflection of behavioral issues over the last two or three years. Even companies need new management control systems. So, there is still an open door for the new conception of beyond budgeting. Twenty years after Fraser and Hope have developed this concept it becomes more important for the companies. For the successful integration of the new generations this new approach becomes more and more relevant. This contribution deals with the causal relations between the new generations Y and Z, the digital transformation and the chances for a beyond budgeting transformation. A conceptional framework was developed by the research group and is presented at the FDIBA conference. The PhD candidates also provide an empirical pre-study. They have already analyzed the technological affinity of the younger generations and its influence for a digital transformation.

***Zusammenfassung* —** Die Integration der verhaltenswissenschaftlichen Aspekte zur Beantwortung ökonomischer Fragestellungen wird unter dem englischsprachigen Begriff Behavioral Accounting zusammengefasst. Während sich Behavioral Accounting bereits in der angelsächsischen Forschung etablieren konnte, ist eine Adaption in Deutschland noch nicht flächendeckend zu sehen. Dieses neue verhaltenssteuernde Controlling wird aber von Unternehmen sehr ernst genommen und bietet eine offene Eingangstüre für Konzepte wie den Beyond Budgeting-Ansatz. Es ließe sich daher erklären, warum in den letzten Jahren eine stärkere Reflexion des Beyond Budgeting erfolgt und warum mehr Unternehmen diese Konzepte (ggf. auch zur Integration der neuen Generationen Y und Z) in Anwendung bringen. Im Rahmen des vorliegenden Beitrags auf der FDIBA-Tagung in Sofia sollen die kausalen Zusammenhänge zwischen dem Eintritt der neuen Generation Y und Z, dem sich vollziehenden Prozess der Digitalen Transformation und den sich daraus bietenden Chancen für eine Beyond Budgeting-Kultur dargelegt werden. Die hier analysierte Kausalkette bietet den Forschungsrahmen der Arbeitsgruppe. Neben der Darstellung des Wirkungsmodells soll eine erste empirische Analyse zur Technikaffinität der neuen Generationen Y und Z vorgestellt werden.

I. Eine neue Chance für Beyond Budgeting!

Ausgangslage: Die stetig zunehmende Dynamik und Komplexität in Wertschöpfungsprozessen verursacht eine Entwertung der Budgetierung. Die Methode der Budgetierung gilt als antiquiert, kann sie doch ausschließlich unter steuer- und planbaren Rahmenbedingungen gezielt zur Managementunterstützung eingesetzt werden. Je komplexer und dynamischer sich jedoch die Anforderungen des Marktes gestalten desto weiter rücken menschliche Fähigkeiten erneut in den Fokus.

Als Reaktion auf dieses beschriebene Dilemma wurde zu Beginn der 2000er Jahre erstmals der Beyond Budgeting-Ansatz vorgestellt. Dabei folgt Beyond Budgeting nicht der Maßgabe, die Methode der Budgetierung zu verbessern. Vielmehr fordert Beyond Budgeting eine erneute Revolution. Mit kleineren organisatorischen Einheiten, die als Profit- oder Cost-Center eigenverantwortlich handeln, soll das Denken und Handeln in einem sich selbst steuernden Organisationsmodell zusammengeführt werden. Beyond Budgeting verlangt eine Organisation weg von einer hierarchisch aufgestellten Unternehmensführung hin zur radikalen Dezentralisierung [3], [4], [11], [12], [13], [14], [15], [21].

Dieser neue Theorieansatz hat in der Praxis bisher wenig Beachtung erfahren, kann aber vor dem Hintergrund der neuen Generationen Y und Z eine größere Durchsetzungskraft entfalten: Bereits seit der Pionierarbeit von Argyris 1953 wissen Wissenschaftler um die dysfunktionalen Effekte der klassischen Budgetierung, die zu ernsthaften Problemen in den zwischenmenschlichen Beziehungen wie „budgetary slacks", Datenmanipulation oder um knappe Ressourcen konkurrierende Abteilungen führen kann [1], [2]. Obwohl viele Wissenschaftler eine Veränderung der Budgetierung zum Abbau der dysfunktionalen Effekte forderten, indem den Mitarbeiten mehr Mitsprache eingeräumt wird, ist in den vergangenen 40 Jahren wenig passiert. Ein Grund hierfür

könnte in dem größeren Übel der organisatorischen Veränderung im Vergleich zu einem ineffizienten Management-Kontroll-System gesehen werden. Die Aufgabe der Kontrolle zugunsten der Selbststeuerung könnte Manager abschrecken.

Vertrauen wird als ein wichtiger Faktor bei der Implementierung von Beyond Budgeting ausgemacht, da Vertrauen als Basis für eine offene Kommunikation, Transparenz und den Austausch von Wissen gilt. Zudem kann Vertrauen ein positiver Einfluss auf die Leistung von Arbeitnehmern nachgewiesen werden [7], [8], [22].
Wie empirisch nachgewiesen werden konnte, wird das organisatorische Vertrauen dabei maßgeblich durch organisatorische Gerechtigkeit bestimmt. Organisatorische Gerechtigkeit kann als wahrgenommene Fairness bei arbeitsbezogenen Entscheidungen, das heißt bei der Ressourcenverteilung und -zuordnung, verstanden werden. Diese neue Unternehmenskultur wird von den neu in den Arbeitsmarkt eintretenden Generationen Y und Z gefordert. Unternehmen müssen sich im „War for talents“ dieser Herausforderung stellen. Im Kampf um die klügsten Mitarbeiter und der Ausbildung einer starken und attraktiven Arbeitgebermarke werden sich Organisationen wandeln müssen [6].

Daher erlebt das Beyond Budgeting in den letzten fünf Jahren einen enormen Zuspruch. So weist der „Beyond Budgeting Round Table“ mehr als 100 Unternehmen aus, die dieses Konzept aktiv anwenden. Warum erleben wir also gerade – nach der Wirtschaftskrise in den Jahren 2008/2009 – eine fundamentale Zuwendung zu dezentralen Steuerungskonzepten? Könnten nicht insbesondere begünstigende Faktoren wie der Eintritt der Generation Y ins Erwerbsleben sowie die Erkenntnisse der Wirtschaftskrise – Konzernkonglomerate in kleinen Profit- bzw. Cost-Center-Strukturen zu führen – einen Anstoß zu neuem Denken gegeben haben? Und auch die Digitalisierung wird zum “Enabler” des Wandels. Darauf wird nachfolgend eingegangen [21].

II. Die Enabler

A. Digitalisierung / Internet of things

Die Vision der digitalen Transformation umfasst verschiedene Komponenten. Zu den technischen Bestandteilen der Industrie 4.0 zählen die Cyber-Physischen-Systeme, wodurch intelligente Objekte, Maschinen und Anlagen entstehen. Wichtig für den reibungslosen und vor allem schnellen Datenaustausch von Produkt zu Maschine und den Maschinen untereinander ist das Internet. Dabei wird zwischen dem Internet der Menschen, dem Internet der Dienste und dem Internet der Dinge unterschieden. Alle drei Bereiche des Internets sind über Cyber-Physische-Systeme miteinander verknüpft und bilden eine große Kommunikationsplattform, über welche Menschen, Maschinen und Produkte fortlaufend miteinander kommunizieren können. Hierbei vernetzt vor allem das Internet der Dinge den Menschen mit Objekten und Systemen – die physische und digitale Welt werden miteinander verknüpft. Neben der rein technischen Transformation liegen die Chancen des Internets der Dinge in neuen Geschäftsmodellen sowie in der Organisation und Zusammenarbeit der Unternehmen. Die Digitalisierung und intelligente Vernetzung bedeutet die Speicherung der Daten in Echtzeit und jederzeitige Abrufbarkeit über das Internet. Durch die erhöhte Datentransparenz und kontinuierliche Datenverfügbarkeit wird eine neue Möglichkeit der Selbststeuerung geschaffen. Das Internet der Dinge ermöglicht Mitarbeitern in dezentralen Strukturen die für ihren jeweiligen Verantwortungsbereich relevanten Daten unabhängig und selbstständig zu organisieren, auszuwerten und einzusetzen [5], [9], [25].

Die Fortschritte der digitalen Transformation leisten einen wertvollen Beitrag, die Rahmenbedingungen für die im Behavioral Management Accounting-Ansatz und insbesondere im Beyond Budgeting-Konzept geforderte Selbststeuerung zu verbessern.

Vor diesem Hintergrund stellt sich die zentrale Frage, inwiefern der Auftrieb der digitalen Transformation für die zukunftsweisenden Führungskonzepte des Behavioral Management Accounting und Beyond Budgeting durch die nun stärker in das Erwerbsleben rückenden Generationen X, Y und Z angenommen und genutzt wird. Wie gestaltet sich die Bereitschaft der Generationen X, Y und Z das Internet der Dinge anzunehmen? Inwiefern werden die Rahmenbedingungen des Internets der Dinge und die Anforderungen der Generationen X, Y und Z an das Arbeitsumfeld zu einem Dissens oder Konsens führen [16], [17], [18], [19]?

B. Generationenmanagement und Implikationen für die Generationen X, Y und Z

In den Medien und wissenschaftlichen Publikationen wird seit geraumer Zeit von der Generation Y gesprochen, die aktuell auf den Arbeitsmarkt eintritt oder teils bereits eingetreten ist. Oft wird betont, dass sich diese in vielen Aspekten von den Vorgängergenerationen unterscheidet. Infolge des demografischen Wandels treten mehr Arbeitnehmer aus dem Berufsleben in den Ruhestand als neue Arbeitnehmer auf den Arbeitsmarkt eintreten, wodurch die Generation Y extrem an Bedeutung für die Unternehmen gewinnt.

Dieser Generation werden aber andere Attribute und Eigenschaften zugesprochen als wir dies für die schon bekannte Generation X und die ganz neue, nächste Generation Z anführen können. Ein Arbeitgeber sollte jedoch attraktiv für alle sein und sich im Umgang mit den verschiedenen „Anspruchsgruppen“ gut aufstellen [10], [16], [20], [24].

Viele Branchen haben Nachwuchssorgen und junge, gut ausgebildete Arbeitskräfte sind gesucht. Nach dem Prinzip von Angebot und Nachfrage sind Arbeitgeber zunehmend in der Pflicht, für diese jungen Arbeitskräfte attraktiv zu sein, allerdings ohne dass dies auf Kosten der älteren Mitarbeiter geschieht. Um hier erfolgreich zu sein, ist es wichtig zu wissen, welche Generationen auf dem Arbeitsmarkt präsent sind, welche Ansprüche und Bedürfnisse sie haben, aber auch, wie man ihre Stärken am besten nutzen kann.

Generation Z: Diese Generation ist auch in der Literatur noch die Neueste. In Publikationen, die vor dem Jahr 2014 erschienen sind, kommt die sogenannte Generation Z so gut wie nicht vor. Wenn überhaupt tritt sie als nicht explizit benannter, jüngster Teil der Generation Y in Erscheinung. Über das früheste Geburtsjahr herrscht Uneinigkeit. Es wird zwischen 1991 [16] und 2000 angesetzt, meistens jedoch mit 1995 [23], [24] angegeben. Ein Enddatum gibt es noch nicht. Die prägenden Jahre beginnen um 2007 und dauern noch an. Diese Generation ist schwierig abzugrenzen und zu definieren, da der nötige Abstand fehlt. Wie befinden uns noch mitten in den prägenden Jahren, sodass es zudem schwer abzusehen ist, welche Ereignisse wichtig sind und wie sich diese letztlich auswirken werden [19]. Allerdings ist festzuhalten, dass sie, stärker noch als die Millennials, in einem Zeitalter ständiger Krisen aufwachsen – die Gesellschaft, in der sie leben, wird als kompliziert, undurchsichtig und zunehmend ungerecht wahrgenommen; den politischen und gesellschaftlichen Geschehnissen stehen sie größtenteils machtlos gegenüber. Oft wird ihnen daher Politikverdrossenheit attestiert, also ein Desinteresse an jeglicher politischer Teilhabe [24]. Es ist nicht

ganz einfach, eine einheitliche Charakterisierung zu bestimmen. Allen Publikationen gemein ist aber die Benennung der Generation Z als „Digital Natives“, also als Menschen, die mit digitalen Technologien aufgewachsen sind und diese quasi intuitiv anwenden. Diese werden als selbstverständlich vorhanden und Mittel zum Zweck angesehen, was unter anderem zur Folge hat, dass sich eher nicht aktiv mit technischen Neuerungen auseinandergesetzt wird. Dies ist ein wesentlicher Unterschied zur Generation Y, deren Mitglieder digitale Technologien noch eher aktiv anwenden und mitgestalten wollen [19], [24].

III. Das kausale Wirkungsmodell – der Forschungsrahmen

Ziel des Führungsansatzes von Beyond Budgeting ist es, Unternehmen flexibler agieren und erfolgreicher werden zu lassen. Darunter fällt auch, leistungsfähige und talentierte Mitarbeiter durch den Anreiz auf mehr Entscheidungsfreiheit und höhere Flexibilität zu gewinnen. Die Mitarbeiter, die den Beyond Budgeting-Ansatz zukünftig umsetzen und leben müssen, beginnen mit der Generation Y. Daraus abgeleitet kann sich ein erfolgreiches Beyond Budgeting ausschließlich im Einklang der Anforderungen der Beyond Budgeting-Methode und den Persönlichkeitseigenschaften der Generation Y einstellen. Beyond Budgeting kann insbesondere für wissensbasierte Organisationen – die Generation Y – enorme Vorteile bieten.

Der Generation Y ist offen gegenüber modernen Medien und Technologien. Ihr wird eine hohe Informationalisierung zugeschrieben, das heißt durch eine angeeignete Informationsbeschaffungs- und -verarbeitungskompetenz sind (fast) alle Informationen (fast) immer und (fast) überall verfügbar. Dieser Medien- und IT-Zugang der jüngeren Generation Y wird vor allem vor dem Hintergrund von Industrie 4.0 bzw. digitaler Transformation besonders wichtig.

Insgesamt ergeben sich drei Kausalzusammenhänge:

- Die Digital Natives machen durch ihre Technikaffinität eine dezentrale Steuerung möglich (A)
- Zugleich fordern Sie auch dezentralere Strukturen an Ihrem Arbeitsplatz und wollen dort eigenverantwortlich arbeiten (B)
- In Summe erfahren verhaltenssteuernde Management-Konzepte einen großen Zuspruch – insbesondere in jungen, dynamischen Unternehmen. (C)

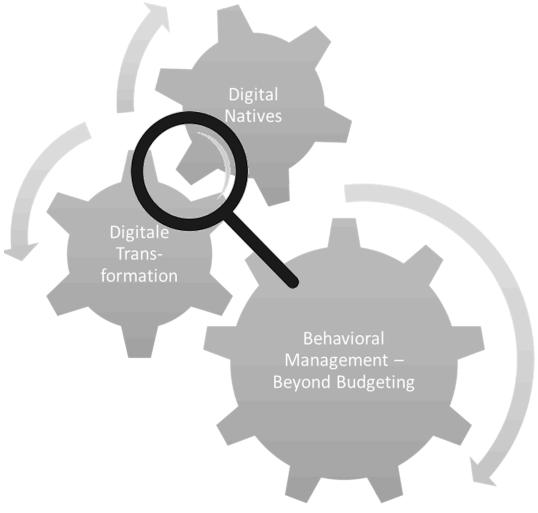


Abb. 1: Forschungsrahmen – Kausalmodell

Die vorstehende Abbildung (Abb. 1) veranschaulicht diesen Zusammenhang. Im Rahmen der FDIBA Tagung soll explizit der Zusammenhang von Digital Natives und digitaler Transformation als ein „Enabler“ des Veränderungsprozesses beleuchtet werden. Hier sollen erste empirische Ergebnisse einer Vorstudie vorgestellt werden.

Forschungsfrage:

Theoretischer Diskurs: Leistet die Hinwendung auf ein Behavioral Management eine Forcierung des Beyond Budgeting Ansatzes?

Begriffe: Behavioral Management Accounting, Digitale Transformation, Generation X und Z, Controlling, Organisationsveränderung, Selbststeuerung, Profit-Center

References

[1] Argyris, C., Miller, F. B., *The Impact of Budgets on People*, New York, 1952.

[2] Argyris, C., “Human Problems with Budgets”, in *Harvard Business Review*, Vol. 31 No. 1, 1953, S. 97-110.

[3] Bisbe, J., Otley, D., “The effects of the interactive use of management control systems on product innovation”, in *Accounting, Organizations and Society*, Vol. 29 No. 8, 2004, S. 709-737.

[4] Bogsnes, B., *Implementing Beyond Budgeting: Unlocking the Performance Potential*, John Wiley & Sons, Hoboken, NJ, 2009.

[5] Bundesministerium für Bildung und Forschung Referat IT-Systeme, *Zukunftsbild "Industrie 4.0"*, Bundesministerium für Bildung und Forschung Referat IT-Systeme, Bonn, 2013.

[6] Chin, J. L., Desormeaux, L., Sawyer, K., Making way for Paradigms of diversity leadership, in *Consulting Psychology Journal: Practice and Research*, No. 68, 2016, S. 49-71.

[7] Colquitt, J. A., Scott, B. A., LePine, J. A., “Trust, trustworthiness, and trust propensity: A meta analytic test of their unique relationships with risk taking and job performance”, in *Journal of Applied Psychology*, Vol. 92 No. 4, 2007, S. 909-927.

[8] Costanza, D. P., Badger, J. M., Fraser, R. L., Severt, J. B., Gade, P. A., Generational Differences in Work-Related Attitudes. A Meta-analysis, in *Journal of Business & Psychology,* No. 27, 2012, S. 375-394.

[9] Drath, R., Matthias, B., YuMI – Robotics for Industry 4.0, in *Albus Schäffer, A. O et. al. ., Internet of Things - Executive Handbook 2016 - Inspiration - Ideas – Insights, WYZE Projects GmbH*, London, 2016, S. 26-29.

[10] Ewinger, D., Ternès, A., Koerbel, J., Towers, I., *Arbeitswelt im Zeitalter der Individualisierung – Trends: Multigrafie und Multi-Option in der Generation Y*, Wiesbaden, 2016.

[11] Gillenkirch, R. M., Arnold, M. C., State of the Art des Behavioral Accounting, in *wirtschaftswissenschaftliches Studium,* 2008, S. 128-134.

[12] Hope, J., Bunce, P., Röösli, F., *The Leaders's Dilemma: How to build an empowered and adaptive oragnization without losing control,* Jossey-Bass, Chichester, UK, 2011.

[13] Hope, J., Fraser, R., “Beyond budgeting.”, in *Management Accounting: Magazine for Chartered Management Accountants*, Vol. 75 No. 11, 1997, S. 20.

[14] Hope, J., Fraser, R., Beyond Budgeting: How Managers Can Break Free from the Annual Performance Trap, in *Harvard Business School Press*, Boston Massachusetts, 2003a.

[15] Hope, J., Fraser, R., “Who Needs Budgets?”, in *Harvard Business Review*, Vol. 81 No. 2, 2003b, S. 108-115.

[16] Kapil, Y., Roy, A., “A Critical Evaluation of Generation Z at Workplaces”, in *International Journal of Social Relevance & Concern*, Vol. 2, No. 1, 2014, S. 10-14.

[17] Kupperschmidt, B. R., Multigeneration employees: Strategies for effective management, in *The Health Care Manager*, No. 19, 2000, S. 65-76.

[18] Lyons, S. T., Schweitzer, L., Ng, E. S.W., How have careers changed? An investigation of changing career patterns across four generations, in *Journal of Managerial Psychology*, No. 30, 2015, S. 8-21.

[19] Mangelsdorf, M., *Von Babyboomer bis Generation Z. Der richtige Umgang mit unterschiedlichen Generationen im Unternehmen*, Offenbach, 2015.

[20] Petroulas, E., Brown, D., Sundin, H., Generational Characteristics and their Impact on Preference for Management Control Systems, in *Australian Accounting Review,* No. 20, S. 221-240.

[21] Pfläging, N., *Beyond Budgeting, Better Budgeting*, 2. Aufl., Hamburg, 2011.

[22] Rousseau, D. M., Sitkin, S. B., Burt, R. S., Camerer, C., "NOT SO DIFFERENT AFTER ALL: A CROSS-DISCIPLINE VIEW OF TRUST.", in *Academy of Management Review*, Vol. 23 No. 3, 1998, S. 393-404.

[23] Ruthus, J., *Arbeitgeberattraktivität aus Sicht der Generation Y. Handlungsempfehlungen für das Human Resources Management.* Wiesbaden, 2014.

[24] Scholz, C., *Generation Z. Wie sie tickt, was sie verändert und warum sie uns alle ansteckt*, Weinheim, 2014.

[25] Weber, J., Linder, S., "Neugestaltung der Budgetierung mit Better und Beyond Budgeting?: Eine Bewertung der Konzepte", in *Advanced Controlling*, Vol. 68, 1st ed., WileyVCH, Weinheim, 2008.

X-Matrix Combining with Lean/Six Sigma for Company Great Results

Die Kombination zwischen X-Matrix und Lean/Six Sigma für ein hohes Unternehmensergebnis

Nikolay Kambosev

TU Sofia, FDIBA
Sofia, Bulgaria, kambosev@gmail.com

Abstract — The big competition between companies leads to usage of various methods for assuring qualitative products in a short run. Therefore, some of the large companies on the market decide to use combination between well-known process improvement methods and create their own on the basis of existing ones to achieve great company results.

Zusammenfassung — Die große Konkurrenz zwischen den Unternehmen führt zu der Verwendung der verschiedenen Methoden für die kurzfristige Qualitätssicherung der Produkte. Deswegen verwenden die großen Unternehmen auf dem Markt eine Kombination zwischen bekannten Methoden der Prozessverbesserung und Methoden für ein hohes Unternehmensergebnis.

I. Introduction

The purpose of this article to show X-Matrix interaction with the other process improvement methods: Lean and Six Sigma. It also aims to describe how the combination of these three process improvement methods usage lead to significant improvement of the business environment and management control. When focused on the highest value projects and supported by the right performance improvement infrastructure, definitely the company achieves great results in: inventory working capital optimization, procurement optimization, focusing in strategic marketing for rapid products development based on the voice of the customers, improvement of company's position in high growth regions. All this progress is achieved by well-developed breakthrough initiatives based on Hoshin Kanri X-Matrix supported by Six Sigma, an effectiveness based process that brings the company process under statistical control, and Lean, an efficiency based process that improves dramatically the process speed or reduced the invested capital.

II. Hoshin Kanri X-Matrix

The likelihood of success in strategic planning can be increased significantly when the priorities are agreed executive level.

Hoshin Kanri Planning process takes into account execution of every step of the way, providing the company with a systematic method of strategic planning and managing progress toward achieving the company's strategic goals. Hoshin Kanri Planning process is used in combination with business development and innovation tools. This process aligns the company strategic objectives with the specific resources and action plans needed to make them happen. Through a back-and-forth refinement system known as "catchball," the entire organization becomes involved in delivering a combination of breakthrough performance and daily management [2].

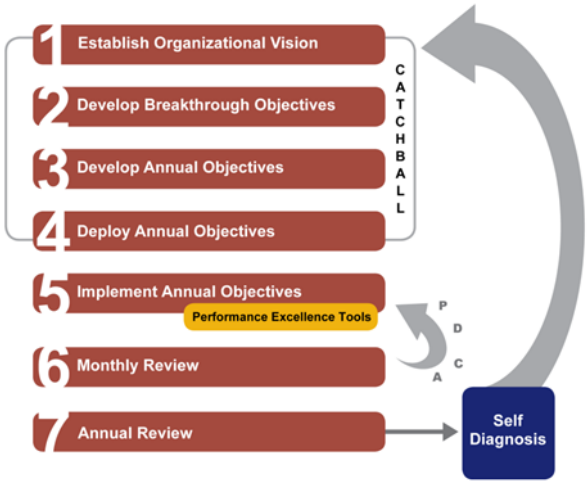


Fig. 1. Hoshin Kanri Planning 7 Steps [2]

Hoshin Kanri Planning is a flexible method that focused on the company's needs to make it successful. It helps the business to be connected in a right track that generates great results.

The X-Matrix is used by the companies to develop top-level improvement priorities and then to apply metrics to them. Then, it is created business-specific second- and third-level targets to improve that tie directly to the top-level priorities. Each level goes down into more and more detail to the end-product manufacturing and sale. This alignment keeps people focused and integrates different departments, making sure everything across the entire company aligns back up to the strategic objectives. The strategy centered thinking drives innovation to create breakthrough opportunities and creates balance in short term and long term objectives. All functions and stakeholders operate as a unified organization to drive accountability. The capabilities required to achieve breakthrough initiatives are defined and resourced. A strategy deployment process is in place from the top level of the business to the point of execution.

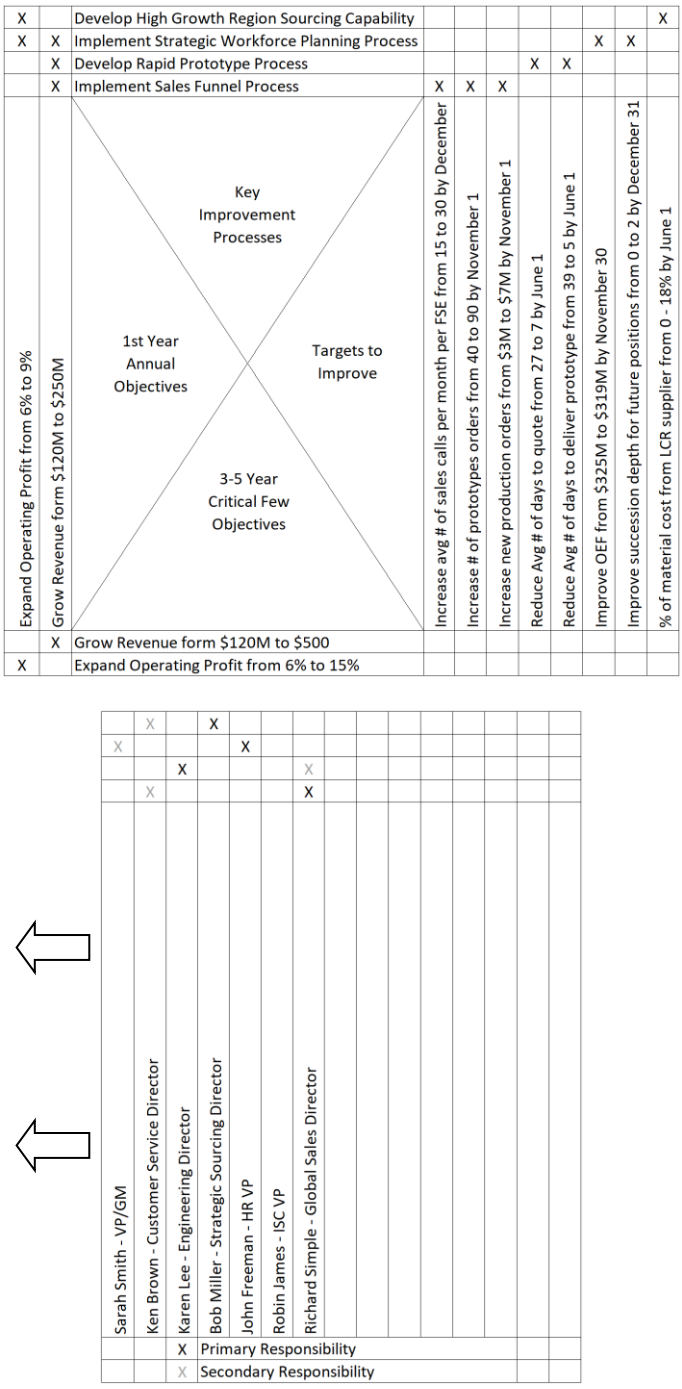


Fig. 2. Hoshin Kanri X-Matrix

The Hoshin Planning Matrix, or X matrix, captures the objectives and cascading priorities. Other process improvement methods Lean/Six Sigma maximizes shareholder value by achieving the fastest rate of improvement in customer satisfaction, cost, quality, process speed, and invested capital.

III. Combining Six Sigma Quality with Lean Speed

Six Sigma, including the Design for Six Sigma is an effectiveness focused method based on the voice of the customers, decision making process based on data and facts; variation reduction to near perfection levels and provides analytical and statistical rigor.

Lean, on the other hand, is an efficiency focused method that provides guiding principles in the company's business based on operating system. It eliminates all the waste. It create process flow and demand pull. It optimizes the resources throughout the organization.

Combining Six Sigma Quality with Lean Speed is extremely powerful in improving the quality and speed of all types of transactional processes, including sales and marketing, quotations/pricing/order processing, product development, financial, administrative, and human resources. These methods are already being used and widely endorsed by companies such as Caterpillar, GE, Honeywell, International Truck, ITT Industries, NCR, Northrop Grumman, Lockheed Martin, Rockwell, Raytheon, and many others [1].

TABLE I. Lean/Six Sigma Method (own figure)

Strategic Targets	**Methods for Achieving the Strategic Targets**			
Growth				Lean Six Sigma
Innovation			Design for Six Sigma	
Quality		Six Sigma		
Speed	Lean Management			
Cost				

IV. Honeywell Example of Implementation of X-Matrix and Lean/Six Sigma Methods

Honeywell present themselves on their webpage as "A Fortune 100 company with approximately $40B in sales in 2015, Honeywell creates solutions that improve quality of life for people around the globe — generating clean, healthy energy, and using it more efficiently; increasing our safety and security; enabling people to connect, communicate, and collaborate; and equipping their customers to be even more productive. Honeywell Great Positions in Good Industries have been a huge driver of its portfolio development and organic growth across industries, including homes and building, aviation, defense and space, oil and gas, industrial, chemicals and vehicles" [3].

Honeywell implements among others Six Sigma, Lean and X-Matrix in the form of Honeywell Operating System (HOS) Gold.

Fig. 3. HOS Gold [4]

"HOS Gold focuses on growing sales and becoming more productive, integrating all of Honeywell major internal process initiatives into a total business operating system. It is an end-to-end business management process focused on customers and markets, strategy development and execution, robust management, standardized work and cross-functional engagement" [5].

"Honeywell Operating System (HOS) drives sustainable safety, quality, delivery, cost, and inventory improvements that ultimately will give us a 20-year competitive advantage" [5].

"Functional Transformation (FT) is doing more with less, standardizing the way Honeywell works, reducing costs and improving service quality across Honeywell administrative and back-office functions" [5].

"Sales, Inventory & Operations Planning (SIOP) enables integration and collaboration among Honeywell Integrated Supply Chain, Sourcing and Engineering teams, allowing Honeywell to respond quickly to changes in demand and supply, and driving accountability to reach one set of numbers" [5].

"Velocity Product Development (VPD) brings together all of the functions necessary to successfully launch new technologies - R&D, Manufacturing, Marketing and Sales - to ensure Honeywell delivers the right products at the right price, and faster than Honeywell competitors" [5].

"Honeywell User Experience (HUE) is a systematic process focused on developing solutions for Honeywell users, maintainers, installers, consumers and employees that are easier to use, more intuitive, more efficient and more productive" [5].

"Capability Maturity Model Integration (CMMI) Level 5 certification confirms that a robust and reliable software development process has been achieved" [5].

"Cycle Time initiatives, improved through the application of HOS, are helping to reduce inventories and improve customer delivery" [5].

"Six Sigma/Lean provides the best process analysis tools available, enabling millions of dollars in productivity savings for Honeywell each year" [5].

X-Matrix captures the breakthrough objectives and cascading priorities within Honeywell.

Implementation of combined methods X-Matrix, Lean and Six Sigma leads to great improvement within company management and control, rapid product development based on the voice of the customers, growth, available operating capital, income and cash for a company success.

References

[1] Michael L George, *Lean Six Sigma: Combining Six Sigma Quality with Lean Speed*, The McGraw-Hill Companies USA, *2002*

[2] William Waldo, "Seven Steps Hoshin Planning", 2017, Online. Available at https://www.leanmethods.com/resources/articles/seven-steps-hoshin-planning

[3] Honeywell International, "Welcome to Honeywell", 2017, Online. Available at https://www.honeywell.com/who-we-are/overview

[4] ---- "HOS Gold Building a Best Practice Enterprise", 2017, Online, Available at https://www.honeywell.com/who-we-are/our-initiatives/hos-gold

[5] ---- "Honeywell Fact Sheet 2015". Online. Available at: http://phx.corporate-ir.net/External.File?item=UGFyZW50SUQ9MjczNDk5fENoaWxkSUQ9LTF8VHlwZT0z&t=1

This page is intentionally left blank

Challenges for New Business Models Based on Big-Data Analysis in the Machining Industry

Herausforderungen für neue Geschäftsmodelle auf Basis Big-Data-Analysen in der zerspanenden Fertigung

Dr.-Ing. Thomas Mücke

TDM Systems GmbH, Derendinger Str. 53, D-72072 Tübingen

***Abstract* —** The retrieval of information from the production area leads to a constant optimization of product and manufacturing costs by means of a methodical preparation of real-time data. This reduces production and change costs, coupled with improved workflows in planning and production. The connectivity of ERP, PLM and MES processes creates new possibilities to optimize the technical and business areas through data analytics. Digitization leads through connections via IoT platforms (Internet of Things) to new business models. It shows how the infrastructure has to change in order to achieve new business models through the integration of suppliers and customers in the process chain.

***Zusammenfassung* —** Die Informationsrückgewinnung (retrieval of information) aus der Produktion führt durch eine methodische Aufbereitung von Echtzeitdaten zu stetigen Produkt- und Herstellkostenoptimierungen in der Planung. Dadurch werden Produktions- und Änderungskosten reduziert, verbunden mit verbesserten Workflows im Planungs- und Produktionsbereich. Durch die Vernetzung von ERP-, PLM- und MES geprägten Prozesse entstehen neue Möglichkeiten, welche die technischen und betriebswirtschaftlichen Bereiche ganzheitlich mittels Data-Analytics optimieren. Die Digitalisierung führt durch die Vernetzung über IoT-Plattformen (Internet of Things) zu neuen Geschäftsmodellen. Es wird aufgezeigt, wie sich die Infrastruktur ändern muss um neue Geschäftsmodelle, durch Einbindung der Lieferanten und Kunden, in der Prozesskette zu erzielen.

I. Einführung

Durch Echtzeitdaten sind in der gesamten Wertschöpfungskette neue Potenziale möglich. Dies führt in der spanenden und spanlosen Bearbeitung zu notwendigen standardisierten Maschinenintegrationen auf Basis der VDI-Richtlinie 5600, bzw. dem RAMI 4.0-Adapter mit detaillierten Maschineninformationen zu Prozess- und Bauteildaten, um über die horizontale Vernetzung die Kunden und Lieferanten in den Gesamtprozesskette einzubinden und resultierend neue Effekte zu erzielen. Hierzu müssen die Daten aus der Maschine strukturiert und verdichtet werden, damit die Echtzeitdaten korrekt in Verbindung mit den Informationen aus verschiedenen Systemen in der Gesamtprozesskette optimiert werden können.

Dadurch ist es möglich, die unterschiedlichen Bereiche zu verbessern und mittels der Echtzeitdatenverarbeitung Kunden- und Lieferantenprozesse in die Gesamtprozesskette einzubinden. Basierend auf dieser Informationsverarbeitung sind neue Geschäftsmodelle mit Servicekonzepten realisierbar, welche zu erheblichen Einsparungen führen. Somit können auf Basis der Reduzierung von Produkt- und Herstellkosten individualisierte Produkte mit Losgröße 1 zu Kosten von Großserien gefertigt werden. Desweiteren ist es möglich, dass sich produzierende Unternehmen auf Ihr Kernbusiness konzentrieren, da betriebstechnisch notwendige Serviceleistungen verstärkt durch Lieferanten in neuen Businessmodellen unterstützt werden können, welche zu einer weiteren Kostenreduzierung im Gesamtumfeld führen.

II. Herausforderung für neue Effekte

Um neue Effekte zu erzielen ist die Echtzeitdatenverarbeitung einer der wesentlichen Schlüssel. In der Zerspanung müssen die Informationen in detaillierter Form aus der Maschine ausgelesen werden und die Daten aus den Planungsbereichen assoziierend zusammengeführt werden. Somit ist es zwingend notwendig, die Daten nicht nur zu sammeln, sondern auch assoziativ mit anderen Daten zu verschmelzen, um u.a. die Qualität stetig zu verbessern.

III. Herausforderung für neue Geschäftsmodelle

Zur Einführung neuer Geschäftsmodelle sind die aufgeführten Effekte mittels Big Data-Analytics eine Voraussetzung [1]. Um Kunden und Lieferanten ins gesamte Netzwerk einzubinden ist eine peer-to-peer-Vernetzung nicht mehr möglich. Somit müssen die verschiedenen Bereiche und Systeme in eine Plattform-Integration wie in Abb. 1 dargestellt, überführt werden. Hierzu ist die vertikale und horizontale Integration zur Einbindung der Lieferanten in die Wertschöpfungskette notwendig. Diese Kommunikationstechnik bildet den Endkunden die Basis für eine hohe Flexibilität in der IT-Landschaft, da somit Systeme mit der offenen Architekturanbindung über APIs relative einfach und kostengünstig geändert werden können, da die Schnittstellenabhängigkeiten wesentlich reduziert werden.

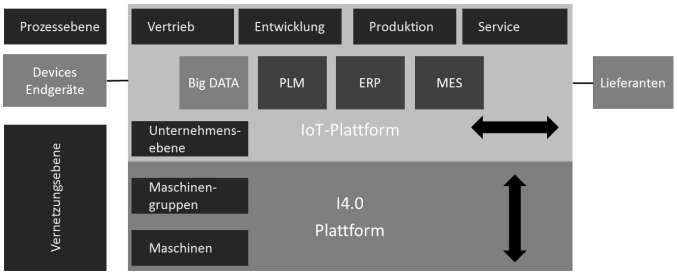


Abb. 1. Vertikale und horizontale IoT-Plattform nach Mücke

IV. Realisierung zur Ermittlung neuer Effekte

Mittels Echtzeitdaten aus der Maschine können neue Effekte erzielt werden. Dies wird am Beispiel von Technologiedaten erläutert.

Der grundsätzliche heutige Ablauf im technischen Bereich des CAD/CAM-Prozesses mit den aktuellen Schwachstellen ist in Abb. 2 dargestellt [2].

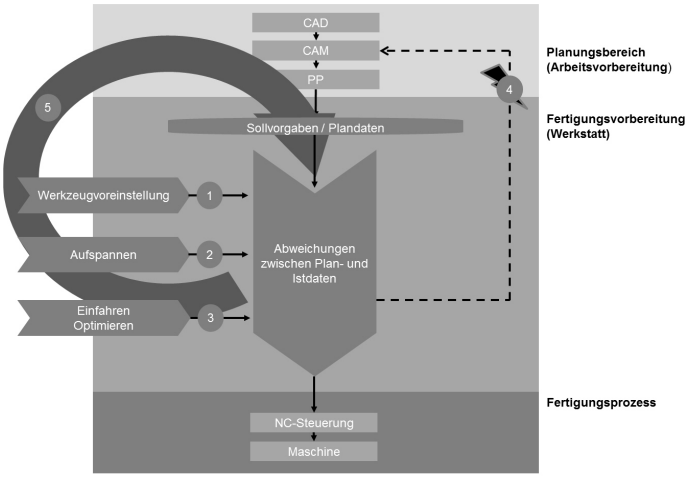


Abb. 2. Heutiges iteratives Vorgehen innerhalb der CAD-CAM-NC-Prozesskette in Anlehnung an Vitr

Legende:
1=Abgleich Werkzeug-Nennmaße mit gemessenen Einstelldaten
2=Abgleich Soll-Lage mit realer Lage entsprechend den Rohteilabmaßen
3=Kollisionsprüfung, Test von geänderten Anfahrbedingungen oder Schnittwerteanpassungen
4=Programmänderungen im G-Code können nur zurückgeführt werden, wenn die Quelldaten aktuell zum NC-Programm gehalten werden
5=manuelle Änderungsmitteilung an die Planung

Das NC-Programm überträgt die Informationen zur Bearbeitung i.d.R. nach DIN 66025 in die Steuerung der Werkzeugmaschine. Das NC-Programm hat somit alle Informationen aus technologischer Sicht; u.a. welche Werkzeuge bei welcher Bearbeitung verwendet werden. Diese Informationen sind den NC-Sätzen zugeordnet. Aufgrund dieser Informationen ist auch die technologische Assoziativität zum CAM-Modell gegeben. Die Steuerung kann ebenso aufgrund vom NC-Satz die zugehörigen Istdaten in Echtzeit aus der Maschine auslesen. Durch eine Big-Data-Analyse kann über Regressionsanalysen die Standzeitkonstante ermittelt werden, welche nicht vom Bediener und dessen Änderungsmöglichkeiten an der Maschine beeinflussbar sind.

Somit können kosten- oder zeitoptimierte Technologiedaten ermittelt werden um die Gesamtkosten der Bearbeitung in Abhängigkeit des Werkzeuges, Werkstoffs, der Spindelleistung der Maschine, etc. zu beeinflussen. Dies ist die Basis um bei kleinen Losgrößen dem NC-Programmierer optimierte Technologiedaten vorzuschlagen. Die Verknüpfung der Systeme mit deren Optimierung ist übergeordnet in Abb. 3 dargestellt.

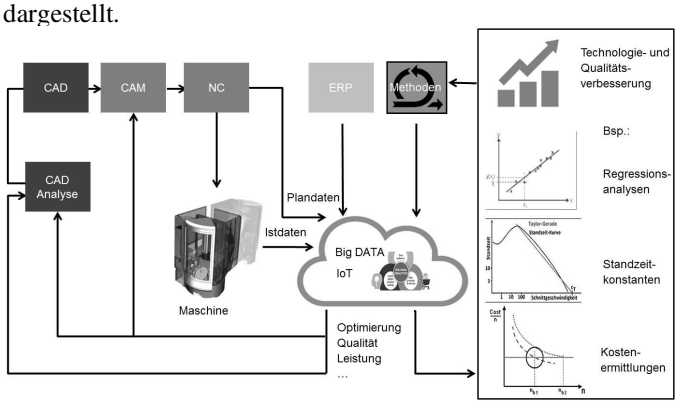


Abb. 3. Übersicht eines Optimierungsprozesses mittels IoT

Basierend auf den Echtzeitdaten und der Historien sind nicht nur Anomalien zu erkennen, sondern auch die Möglichkeit die geeigneten Werkzeuge, bzw. deren Fertigungsverfahren für die Bearbeitung zu ermitteln [3].

Nachdem die Entwicklung und Konstruktion 70-80% der Produktkosten definiert, werden die optimierten Daten aus der Produktion im kompletten Planungsbereich verwendet.

Durch einen integrierten Ansatz kann durch Plug-Ins im CAD-System direkt eine Kostenkontrolle stattfinden, wie in Abb. 4 der Kalkulationsgrundlage zu entnehmen ist. Ebenso wird aufgrund einer Bauteilanalyse und dementsprechendem Regelwerk geprüft, ob das Bauteil herstellbar ist oder welche Technologien notwendig sind. Somit wird nicht nur eine Vorkalkulation auf Knopfdruck Realität und die Produktionsabläufe wesentlich verbessert, sondern auch zugleich der Arbeitsplan für das ERP-System generiert, unter Reduzierung von Änderungskosten durch wesentlich verbesserte Abläufe in der Produktion [4].

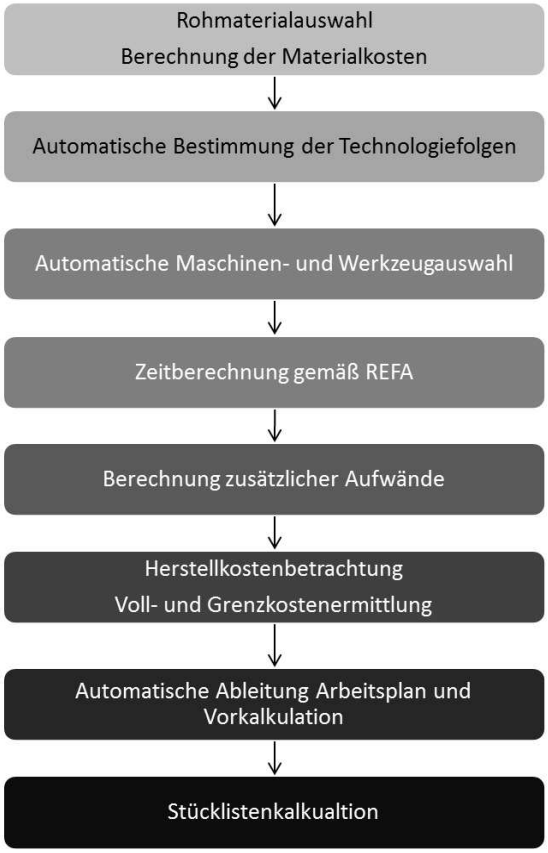


Abb. 4. Abfolge der Berechnungsschritte und deren Ergebnisse in Anlehnung an Weisskopf

Die Echtzeitdaten und deren Auswertung wird zukünftig sicherlich die Möglichkeit bieten, direkt in der Konstruktion die Fertigungstechnologie zu beeinflussen. Es ist von elementarer Bedeutung, dass aufgrund der optimierten Fertigungsinformationen in der Konstruktion definiert werden kann, ob das Bauteil z.B. additiv oder konventionell zerspanend hergestellt wird. Dies muss aufgrund von unterschiedlichen Bauteildefinitionen in der Konstruktion erfolgen, so dass der Anwender direkt im CAD erkennt, welches Fertigungsverfahren aufgrund der Kostensituation gewählt werden sollte.

Durch die Vernetzung der verschiedenen Systeme können die Prozessabläufe im CAD/CAM-Umfeld optimiert werden. Mittels den NC-Sätzen sind durch Bearbeitungsfeatures technologische Assoziativitäten möglich, wodurch heutige manuelle Korrektur von NC-Programmen mittels eines Texteditors aufgrund Änderungen durch den Werker entfallen können. Somit können Änderungen von technologischen Informationen über Analysen direkt in den Planungsbereich ohne manuelle Eingriffe zurückgeführt werden, wie in Abb. 5 zu sehen [5].

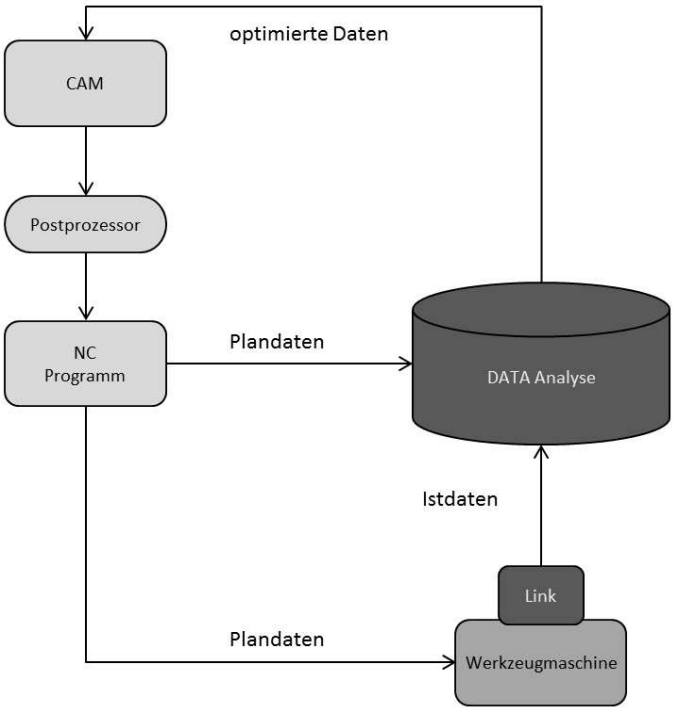


Abb. 5. Rücktransfer von Technologiedaten in das CAM-Projekt nach Mücke

V. Ergebnisse der Effekte

Im Rahmen der Dissertation zum Thema "Informationssystematik zur Optimierung von Konstruktions- und NC-Prozessen" wurden die Möglichkeiten der Echtzeitverarbeitung validiert. Hierbei wurde anhand von Fräs- und Bohrbearbeitungen bei dem Material Aluminium und allg. Baustahl mittels Testreihen eine Validierung mit den Echtzeitdaten durchgeführt. Dadurch konnte nachgewiesen werden, dass durch die Ermittlung der Standzeitkonstanten auf Basis von Echtzeitdaten aus der Maschine verbesserte Technologiewerte, beim gleichzeitig besseren Einsatz von Ressourcen möglich sind. Diese Verbesserung bildet aus Sicht von technologischen Informationen die Basis um ganzheitlich in der Wertschöpfungskette die Prozesse zu optimieren und neue Geschäftsmodelle daraus abzuleiten.

VI. Möglichkeiten für neue Geschäftsmodelle

Aufgrund Echtzeitanalysen im ERP (Feinplanung), der Vernetzung von PLM-geprägten Abläufen und dem Shopfloorbereich durch MES sind neue Servicekonzepte realisierbar. Aufgrund der Echtzeitdatenverarbeitung und der notwendigen Plattformstrategie, wie in Kapitel III dargestellt, sind neue Businessmodelle möglich.

Ein Beispiel eines solchen Businessmodelles im Bereich der Fertigung sind Werkzeuge im Zerspanungsumfeld. Heute kann man bereits davon ausgehen, dass zukünftig direkt aus der CAM-Programmierung die Werkzeuge aufgrund der zu bearbeitenden Geometrie beim Werkzeughersteller über eine Plattform angefragt werden kann, um eine kostenoptimierte Bearbeitung erzielen zu können. Dieser kann, wie in Abb.6 dargestellt, das Komplett-Werkzeug als Datensatz für die Programmierung dem Kunden anbieten und aufgrund der Auftragssicht das Werkzeug, bzw. die notwendigen Komplettwerkzeuge für das Bauteil mit den vermessenen Istwerten JiT oder JiS direkt der Bearbeitung auf Mietbasis zur Verfügung zu stellen.

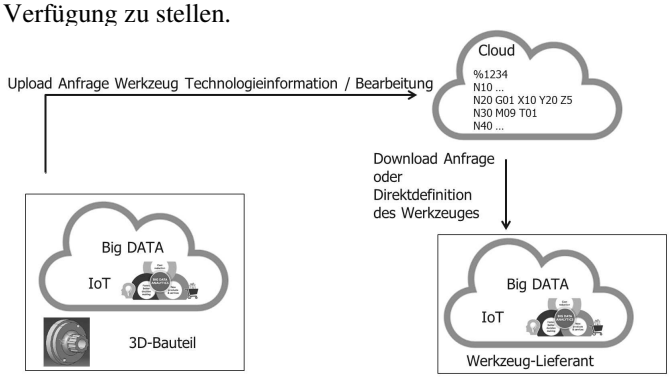


Abb. 6. Übergeordnete Darstellung zur Einbindung von Lieferanten mit neuem Businessmodell

Neue Servicekonzepte reduzieren nicht nur Produkt- und Herstellkosten, sondern verringern u.a. auch die

- Kapitalbindung bei Werkzeugen, da ein Werkzeuglager nicht mehr benötigt wird
- Lohnkosten im Bereich der Werkzeuglogistik, da diese nicht mehr notwendig sind
- Kosten im Werkzeugmanagement, da vorbereitende Funktionen in den Planungsstellen überflüssig werden
- Dienstleistung zur Beschaffung, Instandsetzung, Vermessung und Vorbereitung von Werkzeugen im Shopfloor, da diese nicht mehr notwendig sind.

Neue Businessmodelle bilden somit die Basis zur stetigen Kostenoptimierung auf Basis von neuen Effekten mittels Data-Analytics.

VII. Voraussetzung für neue Geschäftsmodelle

Damit neue Geschäftsmodelle durch eine verstärkte Lieferanten- und Kundenintegration möglich sind, werden IoT-Plattformen benötigt. Somit können z.B. Echtzeitdaten von Lieferanten in die Plattform überführt werden, welche zur Nachverfolgbarkeit von erheblicher Bedeutung sind. Hierzu müssen Daten aus der Maschine über die vertikale I4.0-Integration an die horizontale IoT-Plattform zur Erhöhung der Wertschöpfung über die gesamte Prozesskette angebunden sein, wie in Abb. 1 dargestellt.

VIII. Schlussfolgerung

Die beschriebenen Effekte durch die Vernetzung und Auswertung verschiedenen Datenquellen wird ein Bestandteil von Industrie 4.0 sein, welche in den Schlussfolgerungen

dargestellt werden. Daten müssen in den produktiven Systemen im Zugriff sein. Dadurch wird auch deutlich, wie in Abb. 7 zu sehen, dass Systeme zur reinen planerischen Verwaltung von Daten voraussichtlich an Bedeutung abnehmen werden [6]. IoT-Plattformen müssen dezentrale Systeme, soweit notwendig, aus technischer-, qualitäts- und betriebswirtschaftlicher Sichtweise verknüpfen.

Durch eine zentrale IoT-Lösung werden desweiteren die Anzahl der Schnittstellen trotz dezentraler Systeme reduziert [7].

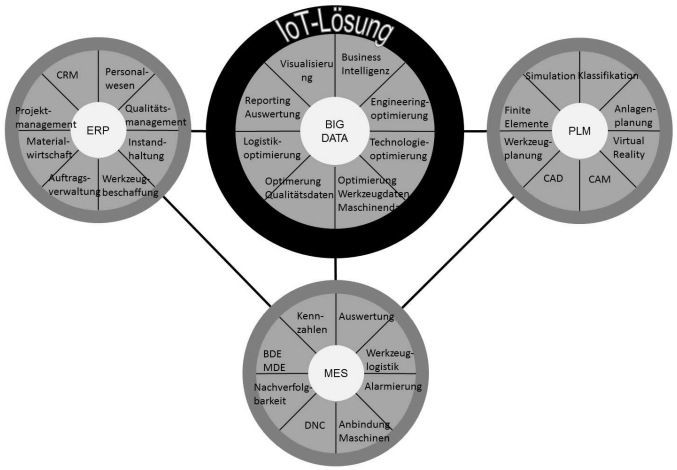


Abb. 7. Hauptsysteme / Funktionen in den Prozessen

Die Einzelprozesse, welche größtenteils in den Bereichen dargestellt sind, werden durch Datenanalysen dementsprechend verbessert, so dass eine bessere Datenqualität erzielt wird. Hierbei ist es von eminenter Bedeutung Daten aus unterschiedlichen Prozessen und Systemen ganzheitlich zu betrachten und zu analysieren, da ansonst keine stetige Verbesserung im Gesamtprozess zur Reduzierung der Produkt- und Herstellkosten zu erzielen ist.

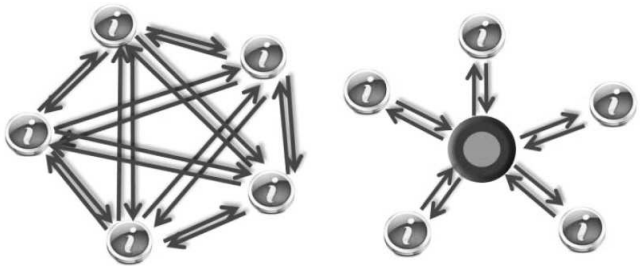


Abb. 8. Reduzierung der Schnittstellen in Anlehnung an Hofmann

Auf Grund aktueller Entwicklungen können bereits u.a. qualitätsrelevante Echtzeit- und OEE-Auswertungen mit technologischen Auswertungen zu weiteren Optimierungen verknüpft werden.

Hierbei ist es von erheblicher Bedeutung, Schnittstellen und Komplexitäten zu reduzieren. Somit ist es sinnvoll, dass z.B. der Maschinenkonnektor von MES auf Basis RAMI 4.0, bzw. VDI 5600 realisiert ist, damit nicht mehrere Systeme auf die Steuerung, bzw. auf die Maschine zugreifen.

Durch die „Middleware“ auf Basis der IoT-Plattform können identische Endgeräten (Devices) verwendet werden, welche auf unterschiedliche Datenbanken zugreifen, ohne dass der End-User dies bemerkt.

Desweiteren ist durch die Reduzierung der Schnittstellenkomplexität eine einfachere Austauschbarkeit der Sub-Systeme möglich, welche zur Flexibilität im IT-Umfeld beiträgt.

Die Vernetzung ist der Innovationstreiber in der gesamten Weltwirtschaft, um Produkte weiterhin kostengünstiger herzustellen. Wenn es uns gelingt, die Digitalisierung als Basisinnovation in die Ökonomie von morgen zu etablieren ist sicherlich der Wohlstand für morgen weiterhin gesichert.

LITERATURVERZEICHNIS

[1] Mücke, T.,*Informationssystematik zur Optimierung von Konstruktions- und NC-Prozessen*, BoD, Mai 2017

[2] Vitr, M.,*CAM-NC-Kopplung für einen durchgängigen, bidirektionalen Informationsfluß zwischen Planung und Fertigung*, Juni 2012

[3] Mücke, T.,“Wenn die virtuelle und die reale Welt verschmelzen“, *IT&Production*, TeDo Verlag, April 2017

[4] Weisskopf, J.,“Automatische Vorkalkulation und Arbeitsplan-generierung für eigenkonstruierte Bauteile“, 2015

[5] Mücke, T.,“integrated digital engineering through real-time data and big-data analytics”, *Vol 9, No 16 (2016): International Collection of scientific proceedings «European Cooperation»,* CLM Consulting PL, September 2016

[6] Mücke, T.,“Mit Daten Schätze heben“, *dispo*, Weka-Verlag, Oktober 2016

[7] Hofmann, J.“Der Entwicklungspfad zur digitalen und resilienten Fertigung im Sinne von INDUSTRIE 4.0“, *VDE-Kongress 2016 – Internet der Dinge,* VDE-Verlag, November 2016

Concept and Development of a VR-Application

Konzept und Entwicklung einer VR-Anwendung

Miglena Dontschewa*, Denise Stamatov‡, Marin B. Marinov†

* University of Applied Sciences Vorarlberg, Department for Computer Sciences and Research Center for User-Centered-Technologies, Hochschulstraße 1, AT-6850 Dornbirn, Austria, don@fhv.at

‡ University of Applied Sciences Vorarlberg, Department for Computer Sciences and Research Center for User-Centered-Technologies, Hochschulstraße 1, AT-6850 Dornbirn, Austria, denise.daniele@web.de

† Technical University of Sofia, Faculty of Electronic Engineering and Technologies, Department of Electronics, 8, Kliment Ohridski Blvd., BG-1756 Sofia, Bulgaria, mbm@tu-sofia.bg

***Abstract* —** The discussed work concept and develops a VR application under circumstances of a study led by the "Research Center User-Centered Technologies" at Vorarlberg University of Applied Sciences with aiming at inspecting rooms while making use of the HTC-VIVE in 360° mode. Therefore, a solution has to be worked out as well as an alternative will be analyzed. The goal is to present and evaluate techniques to capture real rooms. The captured data then shall be examined with given 360° observation tools by wearing a set of VR goggles.

***Zusammenfassung* —** Die vorgestellte Arbeit konzipiert und entwickelt im Rahmen einer Studie im „Forschungszentrum Nutzerzentrierte Technologien" an der Fachhochschule Vorarlberg eine VR-Anwendung, um einen Raum mit der HTC-VIVE in 360°-Modus inspizieren zu können. Es wird die Realisierung einer kompletten Lösung und eine Alternative berichtet. Die Arbeit beschäftigt sich mit Technologien zur Aufnahme von realen Räumen. Diese Aufnahmen sollen in 360°-Betrachtertools verwendet werden mit dem Ziel sie als VR-Sets mit VR-Brillen begutachten zu können.

I. Einführung

Anliegen dieser Arbeit ist gegenwärtige Entwicklungen im Bereich der „Virtuellen Realität" gerecht zu werden. Diese Arbeit verlangt eine interdisziplinäre Herangehensweise, um technische Systeme zu erörtern, die zur Erfüllung der Aufgabenstellung beitragen könnten. Die Herausforderung besteht zum einen darin, dass man die Schnittstellen herausfinden soll und entsprechende Software-Pakete zu finden und/oder notwendige Software-Lösungen zu entwickeln, und zum anderen die Hardware- und Softwarekomponenten dementsprechend zu verbinden, um die Anforderungen der Aufgabenstellung zu erfüllen. Einige dieser Komponenten sind neue Entwicklungen sind und somit durchaus fehleranfällig, instabil und nicht direkt miteinander kombinierbar sind.

Im Rahmen eines geförderten Projekts im Forschungszentrum „Nutzer zentrierte Technologien" (UCT) an der Fachhochschule Vorarlberg wurde eine Studie durchgeführt, um die Frage zu klären, wie es möglich ist, potentielle Nutzer eines bestimmten Präsentationsgerätes (in diesem Fall Cynap von WolfVision) auf diesem möglichst effektiv hinzuweisen. Die grundlegende Frage hiermit ist: Wie macht man es möglich einen beliebigen Raum dem Menschen in einer virtuellen Realität 360° zugänglich zu machen? Zu diesem Zweck wurden VR-Szenarien konzipiert und entwickelt, entsprechende VR-Geräte und Tools untersucht und Usertests durchgeführt. In diesem Schriftstück wird über die Arbeit berichtet, die eben oben genannte Frage bearbeitet. Es wird auf Related Works nicht eingegangen, denn im Rahmen der Recherche keine gefunden wurden und von partiellen Lösungen abgesehen wird. Die Arbeit ist komplex und behandelt neuartige Kombinationen.

II. Methodik

Um die Aufgabenstellung zu erfüllen wurde eingehend eine Analyse der möglichen Lösungen durchgeführt und aus verschiedenen Gründen wurde eine VR-Anwendung favorisiert. In Abb. 1 sieht man die bearbeiteten Komponenten, Schnittstellen, Prozesse, Varianten und Workflow.

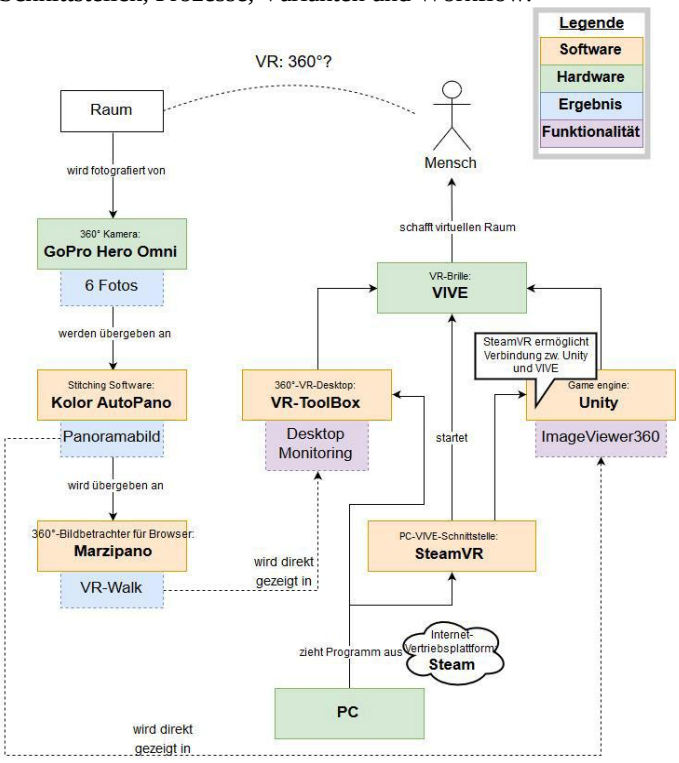


Abb. 1. Einzelne Komponente des Systems und dessen Verbindungen und Workflow

Die Grafik ist so gestaltet worden, dass je nachdem welche VR-Brille benutzt werden soll, sie so angepasst werden kann, dass sie den Ablauf auch für andere VR-Brillen darstellen kann. Dafür müssen die einzelnen Hardware-Pakete, Kamera, VR-Brille und Software-Pakete so verändert werden, dass sie zur bestimmten VR-Brille passen. Um VR-Raum zu generieren sind Panorama - Bilder des darzustellenden Raum notwendig, die Basis der VR-Walk sind. Es sind zwei Möglichkeiten (Virtueller Desktop und Imageviewer360°) in der Grafik zu sehen, die einen bestimmten Raum im virtuellen Raum zu betrachten und mit realen Bewegungen mit dem Raum zu „kommunizieren“ möglich machen.

III. Versuchsaufbau

A. Hardware

HTC-VIVE

Die HTC-VIVE ist ein HeadMountedDisplay mit einem Sichtfeld von 110° bestehend aus zwei OLED mit je einer Auflösung von 1080x1200 Pixel und eine Aktualisierungsrate von 90Hz. Die HTC-VIVE kann sitzend, stehend und raumfüllend bis zu ca. 25m² (ca. 5m x 5m) benutzt werden [1, 2, 3].

Abb. 2. Bestandteile der HTC-VIVE

Die einzelnen Bestandteile der HTC-VIVE sind in Abb. 2 dargestellt, die aus drei Komponenten besteht: einem VR-Headset bzw. HeadMountedDisplay, zwei Controllern und zwei sogenannten Basisstationen. Aktuell sind einige VR-Brillen auf dem Markt, wie zum Beispiel Meta2000, OculusRift, Hololens usw. Die ausgewählte VR-Brille entspricht den Zielen der vorgestellten Studie relativ gut.

GoProHero Omni

Die „GoPro Hero Omni“ ist eine Kamera, die die Möglichkeit gibt Panoramafotos zu erzeugen. Der Omni-Würfel, wie in Abb. 3 zu sehen, besteht aus einem leichten und mehrteiligen Aluminiumgestell, welches ein synchronisiertes Array aus sechs „GoPro Hero4“ Kameras schützt.

Abb. 3. GoProHero Omni

Diese Synchronisation erleichtert die Nachbearbeitung und das Zusammenfügen der einzelnen Bilder, auch Stitching genannt. Durch die GoPro Smart Remote-Fernbedienung, die dem Fotografen ermöglicht aus der Ferne Fotos zu erzeugen ohne sich im Bild zu verstecken, kann die Haupt- bzw. Masterkamera bis zu einer Entfernung von 180m bedient werden, die dann gleichzeitig den Auslöser für alle Kameras auslöst, da im eingebauten Schaltkreis alle sechs GroPro-Kameras automatisch synchronisiert sind.[4] Es können auch andere Kameras verwendet werden, die Panoramafotos erstellt werden. Alternativen dazu sind unter [5] aufgelistet.

B. Software

VR-Tool-System:Unity, ImageViewer360

Unter Unity ist eine Entwicklungsumgebung des Unternehmens Unity Technologies zu verstehen und hat als Zielplattformen PCs, Spielkonsolen, mobile Geräte und Webbrowser. Mithilfe dieser GameEngine ist es möglich interaktive 3D-Grafik-Anwendungen und Spiele auf unterschiedlichen Betriebssystem, also Windows, Linux und macOS, zu entwickeln.

Der ImageViewer360, der in dieser Arbeit verwendet wird, ist ein Unity-Projekt, welches ermöglicht Panoramabilder in 360° sehen zu können. Angelegt an diesem Blog [6] wird später dieses Unity-Projekt dafür verwendet, durch die HTC-VIVE ein Panoramabild im 360°-Modus zu sehen. Dieser ImageViewer360 funktioniert indem er im Inneren einer Sphäre ein Panoramabild hat und im Zentrum dieser Sphäre dann eine Kamera sitzt, die, wenn die VIVE durch Unity gestartet wird, das VIVE-Headset entspricht.

Zuerst wird eine Sphäre in die Szene gesetzt. Danach muss ein einfaches Gameobject, welches als Kamera-Container dienen wird, im Zentrum der Sphäre positioniert werden. Danach muss das Bild, wie bereits gesagt im Sphäreninneren gelegt werden. Um das zu machen, muss ein sogenannter Shader *Pano360Shader.shader* in der eigenen Programmiersprache von Unity geschrieben werden.[7,8, 9] Dieser Shader hat hier zum Zweck, dass das Panoramabild, welches sich in den Assets vom Projekt befinden sollte, auf die Innenseite der Sphäre gesetzt wird und es nicht zu sehr gestreckt wirkt. In Abb. 4 ist der Codeausschnitt zeigt, der das Panoramabild für die Innenseite der Sphäre spiegelt.

```
36        IN.uv_MainTex.x = 1 - IN.uv_MainTex.x;
```

Abb. 4. Screenshot Codeausschnitt Shader: Panoramabild auf Innenseite der Sphäre spiegelt

Wenn das Projekt dann über die VIVE läuft und das Headset bewegt wird, muss die Kamera diesen Bewegungen folgen.

Dafür wird dem Kameracontainer ein C#-Script *CameraInputRotator.cs* gegeben, welches die Bewegungen in X- und Y-Richtung der Maus der Kamera im Sphäreninneren weitergibt und somit das Panoramabild angeschaut werden kann.

```
13        float horizontal = Input.GetAxis("Mouse X");
14        float vertical= Input.GetAxis("Mouse Y");
```

Abb. 5. Screenshot Codeausschnitt: Maus-Input speichern

Der Codeausschnitt in Abb. 5 zeigt im eben genannten Script, wie der Input der Mauskoordinaten in einer Variablen gespeichert wird.

Um dann die Kamera VR und somit HTC-VIVE tauglich zu machen, muss für die eigentliche Kamera ein weiteres C#-Script *AugmentedRealityCamera.cs* erstellt werden, welches die Information des Gyroskops bzw. Bewegungssensors der VR-Brille so bearbeitet, dass die Kamera in der Sphäre sich bewegt und somit der VR-Brillenbenutzer das Panoramabild 360° im Raum betrachten kann. In Abbildung 9 ist ein Codeausschnitt zu sehen, wie der Input des Gyroskops erhalten und weiterverarbeitet wird.

```
50        transform.localRotation = Input.gyro.attitude * initialRot;
```

Abb. 6. Screenshot Codeausschnitt: Gyroscope-daten weiterverarbeiten

Zum Schluss muss der Sphäre das Material gegeben werden, welches das Panoramabild ist, indem das Bild aus den Assets, in Abb. 7 zu sehen, direkt auf die Sphäre in der Object-Hierarchy im Unity mit der Maus darauf gezogen wird. Da es drei unterschiedliche Sets gibt, werden drei Sphären erstellt, wie Abb. 8 zu sehen, die jeweils die drei unterschiedlichen Panoramabilder als Material haben.

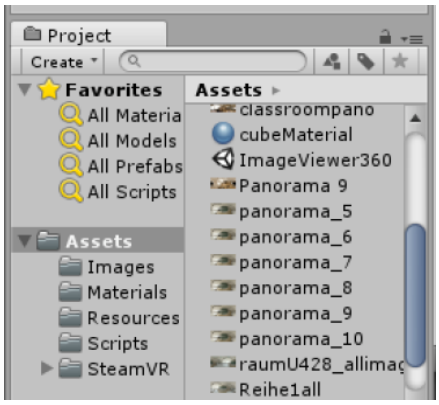


Abb. 7. Screenshot Project-Assets: Panoramabilder

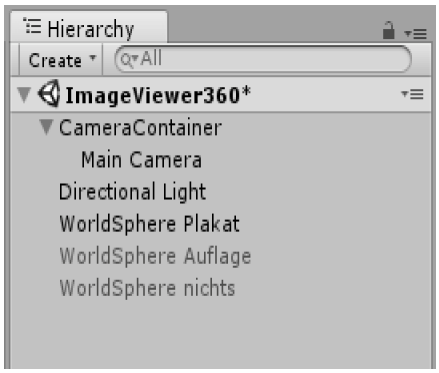


Abb. 8. Screenshot Hierarchie: Unity-Projekt

Steam, SteamVR, VRToolBox

Das Unternehmen Valve entwickelte und betreibt die Internet-Vertriebsplattform Steam, die im Laufe der Jahre für die Betriebssysteme Windows, MacOSX, Android, iOS und Linux erhältlich wurde. Steam vertreibt unter anderem Computerspiele, Software, Filme und Serien. Um die HTC-VIVE überhaupt verwenden zu können, muss momentan ein Account bei Steam angelegt werden, auch wenn die HTC-VIVE die OpenVR API verwendet. Eine OpenSource Implementierung könnte eine Anmeldung umgehen und muss erst geschrieben werden [10].

Wie auch Steam ist auch SteamVR eine Valve-Entwicklung, die eine VirtualReality-Plattform-Erweiterung von Steam ist, d.h. diese Software kann über Steam heruntergeladen werden und ermöglicht eine 360° - VR-Erfahrung.[11] Sie unterstützt nicht nur die HTC-VIVE, sondern auch die Oculus Rift.

SteamVR ermöglicht dem Entwickler die Spielentwicklungsumgebung Unity mit der HTC-VIVE zu verbinden, indem ein Plugin SteamVR aus dem Unity-Assetstore, hier [12], heruntergeladen werden muss. Mithilfe dieses Plugins kann dann zum Beispiel in einem C#-Script die Information erhalten werden, welche dann an einem bestimmten Bestandteil der Unity-Szene gegeben wird, ob das Trackpad eines VIVE-Controllers gedrückt wird. Dieser Input kann dann so viele Würfel im VR-Raum erscheinen lassen, wie das Trackpad gedrückt wurde. Dieses Plugin erweitert somit die Funktionalität von Unity, da der Entwickler einen Zugriff auf die Sensoren, Knöpfen, etc. der HTC-VIVE erhält.

Über Steam im Browser über den Link [13] oder über die Desktopvariante kann das Programm SteamVRToolBox heruntergeladen werden. Dieses Programm schafft für den Benutzer der HTC-VIVE einen virtuellen Desktop. Mit den Controllern kann der Benutzer im VR-Desktop interagieren, das heißt er kann mit den Controllern den Desktop personalisieren und seinen Wünschen nach einrichten.

Bildverarbeitung: Kolor Autopano, Marzipano

Generell können auch andere Stitching-Software verwendet werden, die Alternativen zu den hier vorgestellten Programmen können unter [14] entnommen werden.

Kolor Autopano

Programm Kolor Autopano, das über [15] heruntergeladen werden kann, ist eine Stitching-Software für Fotos und Videos. Stitching bzw. Zusammenfügen bedeutet, wenn aus mehreren einzelnen Fotos eine ganzes Großes erstellt wird.[16] Sie wird zum Erstellen von Panoramafotos (Abb. 9) oder Gesamtansichten von Objekten verwendet. Die Fotos, die zum Beispiel nach dem Stitching ein Panoramafoto darstellen sollen, können direkt per drag&drop ins Programm gezogen werden. Danach wird überprüft, ob die Fotos ein Panorama erzeugen können, indem auf dem Play-Knopf geklickt wird. Das mögliche Resultat wird dann rechts von den einzelnen Fotos dargestellt. Durch Klicken auf das Rädchen öffnet sich ein Fenster, was dazu dient, Einstellungen vorzunehmen und Informationen zu geben, wie Dateiname und -format, Speicherort etc., damit das Rendern gestartet werden kann [17].

Abb. 9. Mögliches Panoramabild

Marzipano

Marzipano ist ein 360°-Bildbetrachter, der im Browser unter[18], verwendet werden kann. Indem das Tool [19] auf dieser Seite verwendet wird, können VR-Walks (Abb.10) aus Panoramafotos erstellt werden. Hierbei wird das Panoramabild per drag&drop ins Browserfenster gezogen und nach einer Lade- und Aufbereitungsphase, können Info- und Link-Hotspots, Startblickpunkt, View-Kontroll-Knöpfe, Fullscreen-Knopf und anderes eingestellt werden.

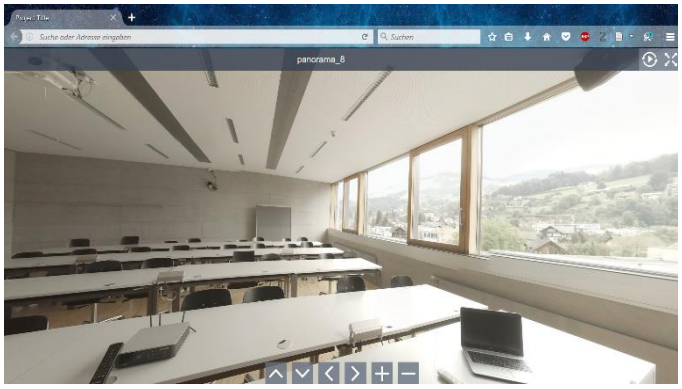


Abb. 10. Screenshot: VR-Walk

Danach kann dieses 360°-Panoramabild als Paket, unter anderem mit einer .html-Datei, exportiert und gespeichert werden.

IV. Experimentelle Ergebnisse

Die Handhabung der 360°-Kamera birgt eine besondere Spezifik, die verschiedene Einstellungen (Abb. 11) und Tests für die Positionierung notwendig macht. Die Position der Kamera im Prozess der Erstellung der Bilder bestimmt die nachfolgende Perspektive.

Abb. 11. Kamerapositionierung

V. Zusammenfassung

Die VR-Anwendung in der jetzigen Form ist funktionsfähig. Eine Erhöhung der Bildqualität in der VR-Brille, sowie eine performantere Hardwareunterstützung, ist wünschenswert. Es sollten alternative Lösungen bearbeitet werden, um die Wartbarkeit zu gewährleisten, denn viele Komponenten der derzeitigen Lösung sind von Drittanbietern und zum Teil als Blackbox zu betrachten. Weitere Anwendungsgebiete mit der aufgestellten Lösung sollten recherchiert werden und entsprechend untersucht werden. Die Unity-Lösung kann und soll weiterentwickelt werden, um die Qualität zu erhöhen und weitere Funktionalitäten einzubauen. Eine mögliche Funktionalität könnte die Begehbarkeit des Raumes sein, um die gezeigten Raumbestandteile von der Nähe betrachten zu können. Auch Interaktivität könnte ausgebaut werden. Die Einsatzfähigkeit der VRToolBox-Lösung soll in ähnlichen Test untersucht werden. Controllerfreie Entwicklung in diesem Zusammenhang sollen stattfinden, zum Beispiel mit LeapMotion oder Kinect.

Literaturverzeichnis

[1] Patrick Schulze, “HTC Vive: So kannst Du in die virtuelle Realität eintauchen“, Online, 05.02.2016. Available at https://goo.gl/Gvwkr7

[2] Darius Schuiszill, “HTC Vive im Test: Abtauchen und Berge erklimmen“, Online, 14.12.2016. Available at https://goo.gl/EcDdg7

[3] HTC, “HTC Vive”, Online, 21.10.2017. Available at https://goo.gl/W9VMgs

[4] “Omni”, Online, 31.10.2017. Available at https://goo.gl/LxSmuo

[5] ”20 best alternatives to GoPro Omni as of 2017“, Online, 31.10.2017. Available at https://goo.gl/AR8RTF

[6] Verónica Valls, “How to make a 360°image viewer with Unity3D”, Online, 14.03.2016. Available at https://goo.gl/rqcfMz

[7] Unity Technologies, “ShaderLab Syntax”, Online, 06.10.2017. Available at https://goo.gl/WNPFrC

[8] “Getting Started with Shaders”, Online, 04.11.2013. Available at https://goo.gl/sPTgBM

[9] “How to write unlit surface shader?”, Online, 23.06.2012. Available at https://goo.gl/9h3VzV

[10] “Vive Port: HTC bringt eigene App… Steam-Account Pflicht?”, Online, 05.08,2016. Available at https://goo.gl/62EWro

[11] ”SteamVR”, Online, 17.06.2016. Available at https://goo.gl/fz3TtA

[12] “SteamVR Plugin”, Online, 29.06.2017. Available at https://goo.gl/9kFBTk

[13] Steam, “VRToolbox: 360 Desktop”, Online, 26.12.2016. Available at https://goo.gl/epsgUH

[14] Abhijith N Arjunan, “8 Best Photo Stitching Software”, Online, 24.03.2017. Available at https://goo.gl/tEft7y

[15] Kolor, “Autopano”, Online, 31.10.2017. Available at https://goo.gl/M36GaF

[16] “Stitching”, Online, 26.07.2017. Available at https://goo.gl/XXooAZ

[17] Kolor, “Autopano Documentation”, Online, 31.10.2017. Available at https://goo.gl/n7D3tS

[18] Marzipano, “Marzipano”, Online, 31.10.2017. Available at http://www.marzipano.net/

[19] Marzipano “Marzipano Tool”, Online, 31.10.2017. Available at http://www.marzipano.net/tool/index.html

[20] D. Stamatov, Concept and Development of a VR-Application, Dornbirn, Austria, 2017.

Non-traditional procedure for the organization of information control systems for intelligent production machines

Nicht-traditionelles Verfahren zur Organisierung von Informationssteuerungssysteme für Intelligente Produktionsmaschinen

Sasho Guergov*, Aleksandar Tsenov†

Technical University of Sofia, Bulgaria,
*Faculty of Mechanical Engineering, email: sguergov@tu-sofia.bg,
†FDIBA, email: aleksandar.tsenov@fdiba.tu-sofia.bg

***Zusammenfassung* —** In den letzten Jahren der industrielle Fortschritt ist mit der, auf der Messe in Hannover im 2013 erwähnten Plattform „Industrie 4.0“. Das Konzept ist basierend auf führenden Lösungen von SIEMENS für digitale Produktentwicklung und Produkterstellung. Diese Lösungen sollen die Bedingungen sowie zur Automatisierung des gesamten Lebenszyklus eines Produkts, als auch zur Integrierung von Cyber-Physical Systemen, die eine virtuelle Darstellung der reellen Welt abbilden um die physikalische Produktionsprozesse zu verwalten, um zu kommunizieren und um selbst verteilte Entscheidungen zu treffen.

***Abstract* —** In recent years the industrial progress is with the platform "Industrie 4.0" mentioned at the trade fair in Hannover in 2013. The concept is based on leading solutions from SIEMENS for digital product development and production. These solutions are designed to address the conditions, as well as the automation of the entire life cycle of a product, as well as the integration of cyber-physical systems, which represent a virtual representation of the real world in order to manage physical production processes and to make distributed decisions itself.

I. Einführung

Es existieren mehrere charakteristische Besonderheiten der gegenwertigen Entwicklung der Wirtschaft. Einerseits das kurze Lebenszyklus der Produkte und die rasche technologische Entwicklung ändern das Verständnis, dass in einigen Sektoren die komplexe Verarbeitung sich nicht in einem stabilen Funktionszustand befindet, wenn diese geändert werden soll. Andererseits besteht das Trend in den großen Unternehmen, die Flexibilität als organisationale Lösung zu betrachten um der neuen Marktanforderungen schnell und effektiv entgegenkommen zu können. Eine schnelle Reaktion auf diese Anforderungen lohnt sich nur dann, wenn einen geprüften Produktionsprozess kurzfristig angewendet wird.

Dies erfordert die Entwicklung eines Konzepts, bei dem die Produktionssysteme mehr Intelligenz und Flexibilität umfassen, sowie bei der Erfüllung den technologischen Prozessen, als auch bei dem selbstständigen Entschei-dungstreffen während einer kontinuierlichen Änderung des Produktionsprogramms [1], [2].

Bei diesen Voraussetzungen kann man die folgenden Charakteristiken der Produktionsmaschinen festlegen [3],[4]:

1. Modulare Struktur
2. Skalierbarkeit
3. Leichte und genaue Modulintegration
4. Optionen zur funktionale Transformation
5. Flexibilität und Anpassungsfähigkeit
6. Automatische on-line Diagnostik und Korrektion von Systemparameter

In diesem Kontext kann man die Produktion als Rechnerintegrierten Prozess, bei dem die Intelligenz alle Hardware- und Softwarekomponente umfasst.

Die Änderungen in der Systemkonfiguration führen zur Änderungen in den Parametern der Maschinen, und davon folgt, dass die Steuerung in der Lage sein muss sich selbst zu konfigurieren um sich an die neuen Bedingungen anzupassen.

II. Realisierung von Intelligenten Produktionsmaschinen

Um die Realisierung eines intelligenten, technologischen Prozesses der mechanischen Bearbeitung bei einer vollen Automatisierung zu gewährleisten, soll das Produktionssystem wie folgt enthalten:

1. Identifikation des zu bearbeitenden Objekt;
2. Zuordnung des Objekts zu einer Detailklasse;
3. Auswahl eines typischen Technologieprozesses;
4. Datenwahl und Anhäufung neuer Informationen;
5. Auswahl einer Route für die Reihenfolge technologischen Operationen;
6. Wahl zusätzlicher technologischen Ausrüstung;
7. Wahl eines technologischen Operationsprozesses;
8. Wahl eines Optimierungskriteriums zur Umstrukturierung/Rekonfigurierung der Maschine;
9. Wahl des Strukturmodels der Maschine;
10. Verwaltung des vorläufigen Strukturstandes der Maschine;
11. Wahl eines Steuermodels zur Umstrukturierung des Systems;
12. Steuerung der Maschine.

Abbildung 1 stellt eine Beispielsstruktur des Informationssteuersystems zur Verwaltung und Koordination aller Funktionen einer Intelligenten Produktionsmaschine (IPM) vor.

1. Datenbank zur formalen Beschreibung und Identifikation der Details (Di). Jedes Detail ist durch eine Menge

geometrischer Parameter dargestellt ({GP}), die durch Untermengen von Basisprimitiven ({BP}) und zusätzlichen Primitiven ({DP}) wie folgt formalisiert werden kann: Di ϵ {BP} ∪ {DP};

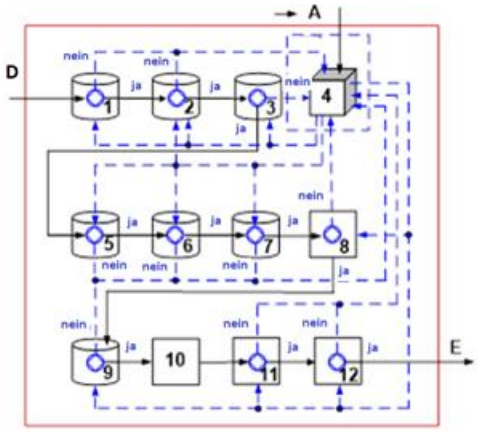


Abb. 1. Struktur des Informationssteuerungssystems

2. Datenbank für Detailklassen: KL_{ij}. Jedem parametrisierten Objekt von KL_{ij} wird einen entsprechenden Code zugeordnet A_{KLij};

3. Datenbank typischer technologischen Prozessen (TP). Jeder TP wird als tj = ({Ei}j, {Ci}j), wo {Ei}j die Menge der Elementen TPtj (Menge de Operationen, Übergängen und Zügen) und {Ci}j – Menge der Verbindungen zwischen den Elementen von TPtj sind. Die Suchprozedur $Px = P_2RP_1$, wo R eine Reichenprozedur ist; P1 – Suchprozedur für die Produktmenge {D}, die identische TP haben; P_2 – Suche in der Menge $\{T_x\}$ von TP nach der Abbildung von P2, das heißt {D} $\Rightarrow \{T_x\}$.

4. Informationssuchsystem zur Datenwahl und Anhäufung neuer Informationen.

5. Datenbank, die Routen für die Reihenfolge technologischen Operationen auf dem Detail enthält. Jeder $t_j = (\{E_i\}_j, \{C_i\}_j$ entspricht eine Route die Reihenfolge technologischen Operationen vom Typ M_n: $\{M_{ij}^{kl}\}$ mit zugeordneter Adresse $A_{KL_{ij}\,KL_{ij}}^{M_n}$ in Datenbank 5.

6. Datenbank mit Informationen für die technologische Ausrüstung. Die Menge {TE} von parametrisch dargestellten Untermengen von Werkzeugen: Gama-Futter ({T}), Greifer ({GR}), Satelliten ({S}) und andere, dass heißt ({TE}U{T},{GR}), {S},…). Jedem Objekt von der Menge {Di} entspricht Menge vom Typ{Ti,GRi,Si}, das heißt

$$D_i \overset{TE}{\rightarrow} \{TTi, GRi, SI\} \tag{1}$$

7. Datenbank mit Informationen über technologischen Operationsprozesse und Steuerprogramme. Bei dieser Prozedur soll den passenden Operationsprozess gewählt werden

$$OT_i \in \{NJ,RJ, YJ\}j_j \subset \{Q\},$$

wo {NJ} ist die Menge der Schneidwerkzeugen; {RJ} ist die Menge der Schneidregimen; {YJ} ist die Menge der Schneidbedingungen;{Q} ist eine indexierte Menge, die den Operationsprozess für den jeweiligen Detail entspricht.

8. Modul zur Wahl des Strukturmodels der Maschine. Die Strukturvariante (SV_i) wird durch die Menge der Module {Mod} und Subsystemen {SS}, die eine intelligente Produktionsmaschine zur Erfüllung konkretes technologischen Prozesses bei minimalen technologischen Wertes (TS_{min}) aufbauen., das heißt

$$SV_i \in \{Mod\} \cap \{SS\} \rightarrow TS_{min}. \tag{2}$$

9. Datenbank, die verschiedene strukturelle Konfigurationen/Varianten der IPM - { SVi}.

10. Abspeicherung vorläufiger Maschinenstruktur. Diese Funktion ist im Falle Erfüllung einer Priorität-Funktion. Zu diesem Zweck ist die Option vorgesehen, den Maschinen- und den Programmstand abzuspeichern, so dass bei einer zukünftigen Umstrukturierung zu dem abgespeicherten Stand, das gestoppte Programm bis zum Ende zu erfüllen. Bei dem Übergang von einem vorigen Zustand zu einem gegenwärtigen, es gilt:

$$P_k^j : m_k(Q) \rightarrow m_j(Q) \tag{3}$$

wo Q ein Operator ist, der den ausgeführten Teil des Produktionsprogramms in % abspeichert.

11. Modul für Wahl/Synthese des Steuermodels der Maschine. Das Model wird wie folgt dargestellt:

$$F_{UM}\{M, P_k^j, y(t), w(p), h(p)\} \tag{4}$$

wo M die Module repräsentiert, die in der jeweiligen konstruktive Variante beteiligt sind, P_k^j beschreibt den vorläufigen Zustand des Systems, *y(t)* ist die Steuerfunktion, die die innere und die äußere Störungen in Echtzeit zählt, *w(p)* ist die Übertragungsfunktion der elektrischen Antriebe der Module, *h(p)* die jeweilige Übertragungsfunktion der Änderung der Eingangsparametern der endlichen Ausführungseinheiten.

12. CNC/PC Steuersystem der IPM.

Die Eingänge A und E beziehen sich auf die Möglichkeit interaktiver Betreiberinterventionen und Informationen über die technischen und wirtschaftlichen Indikatoren der Produktionsaufgabe. Ausgangsanschluss E ist ein Mehrkanal-Eingang für das Maschinen-System. In Abbildung 1 werden gestrichelte Linien die Links in der Verwendung von "vergangener Erfahrung" gegeben und mit unterbrochener "Akkumulation neuer Informationen".

III. Nicht-traditionelles Verfahren zur Organisierung von Informationssteuerungssysteme für Intelligente Produktionsmaschinen

Die wichtigsten Herausforderungen an einer intelligenten Maschine können wie folgt festgelegt werden:

- Zuverlässige Produktion mit wenigem Ausschuss;
- Geringere Energiekosten;
- Vereinfachte Einrichtung und Wartung;
- Höhere Lebensdauer der Maschinen.

Diese Herausforderungen benötigen die Anwendung passenden intelligente Lern- und Funktionsmethoden in den Maschinen, so dass diese selbst den gesamten Produktionsprozess (oder einen großen Teil davon) ohne Anwesenheit und/oder Teilnahme an den Prozess eines Menschen notwendig wäre. Eine intelligente Maschine soll z.B.:

- Abweichungen in den Produktionsprozess und in den Werkstoffen erkennen und korrigieren können;
- Vernetzt mit anderen Maschinen arbeiten können;

- Den Energie- und den Werkzeugverbrauch durch Aufzeichnung und Auswertung von Daten optimieren können und andere.

Alle, bis hierher erwähnte Herausforderungen und Eigenschaften der intelligenten Maschinen, führen ganz einfach zur Entscheidung: die gegenwärtigen Informationstechnologien sollen breit in den Produktionsprozessen eingeführt werden.

Auf der Abbildung 1, sind sowie eine große Menge von Datenbanken, die Daten zur Vorbereitung und Ausführung des technologischen Prozesses enthalten, als auch einige Module vorgesehen, die die wichtigen Aufgaben zur Entscheidungstreffen im Laufe des Produktionsprozesses dienen. Diese Entscheidungen können nur aufgrund komplexer Bewertung großer Menge von Parametern getroffen werden.

In einer vorigen Arbeit wird vorgeschlagen die Entscheidungen durch im Voraus erzeugten Routen für die technologischen Operationsprozesse zu nehmen [5]. Das erfolgt durch Anwendung von heuristischen Algorithmen (Genetic Algorithms), wobei von der immensen Menge möglicher Entscheidungen, die optimalen zur Produktion bestimmten Details selektiert werden und in einer Reinfolge von Operationen umgewandelt werden.

Gegenwärtig reicht dieser Vorgang zu erreichen hochintelligenten Prozessen nicht.

In dieser Arbeit ein integriertes Steuerungs- und Wartungsprozess zur Gewährleistung optimalen Produktion, aufgrund neusten Errungenschaften der M2M (Machine to Machine) Kommunikation und des IoT (Internet of Things) Managements ist vorgeschlagen.

Die, im vorigen Teil beschriebene Module 4, 8, 10, 11 und 12 werden, können als vernetzte Elemente betrachtet werden, wobei diese Vernetzung den Austausch wichtigen für die Funktionen der Module Informationen und Parameterwerten ermöglicht. Dafür braucht man eine Ausrüstung der Module mit passende Netzwerkschnittstellen ausgerüstet werden. In den IP basierten Netzwerke bedeutet das Netzwerkkarten, di sowie mit Kabelverbindungen als auch Wireless zueinander angekoppelt werden können, Die Controller der Netzwerkkarten sollen nicht sehr komplex sein, der Datenaustausch erfolge auf der Schicht 2 der OSI – Referenzmodels und deswegen benötigt man keine Adressierungsprotokolle der höherwertigen Schichten nicht. In den Kontrollern soll eine Funktionalität integriert werden die unten beschrieben ist.

Das Konzept ist basierend auf neue Managementarchitekturen für Internet of Things. Im standardisiertem Management in Internet verwendet man die Manager Agent Architektur, implementiert in dem Management Protokoll SNMP. Es handelt sich um einen intelligenten Manager, der vollständigen Kenntnisse über eine Netzwerkarchitektur hat und je nachdem Entscheidungen bezüglich Konfiguration und Verwaltung der Komponente der Architektur nehmen kann. Der Manager selbst kann aber keine Handlungen selbst ausführen, sondern nur entsprechende Befehle an einer Menge untergeordneten Agenten vergeben. Die Agenten sind die Tatsächlichen Ausführungselemente in der Managementarchitektur und üben die Auswirkung auf die zu menagierenden Komponenten (NE – Network Elements) aus. Die Agenten selbst haben keine „Intelligenz“ selbst zu entscheiden, was zu machen ist (Abb. 2).

Im IoT Managementkonzept sind sogenannte eingebettete Agenten (AA - Auxilary Agents) vorgesehen, die die tatsächliche Funktion Agenten der Agenten unterstützen. (Abb 3).

In den eingebetteten Agenten sind die Modelle der Funktion der NE implementiert. Das bedeutet, die Strukturmodelle, die Funktionsweise und die zu steuernden Parameter sind den eingebetteten Agenten vollständig bekannt und sie können dann selbst die richtige Entscheidung für den Prozessverlauf nehmen und ausführen.

Dies erfolgt unter der Beobachtung seitens Managers, der schon nicht mit der Entscheidungsnahme belastet ist, sondern nur den gesamten Verlauf des Produktionsprozesses verfolgt. Dies passiert mittels von den Agenten generiert und zu ihm gesendeten Mitteilungen.

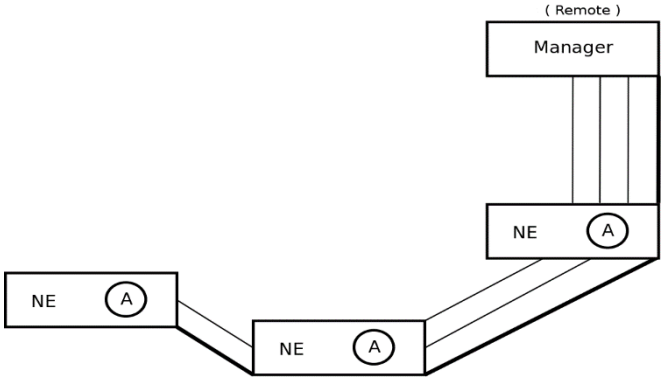


Abb. 2. Standardisierte Struktur des Managementsystems

Da es sich hier um eine Erweiterung einer bereits bestehenden Lösung handelt, ist im Voraus bekannt, welche Funktionalitäten der Manager den Agenten ausgesetzt ist (und umgekehrt), um eine Reihe von Aufgaben zu decken. Daher können diese Funktionalitäten in eine Schnittstelle (Quasi-Manager) abstrahiert und durch den eingebetteten Agenten belichtet werden, so dass die Befehlskette von (NE-A) -M auf (NE-A) -AA-M erweitert werden kann.

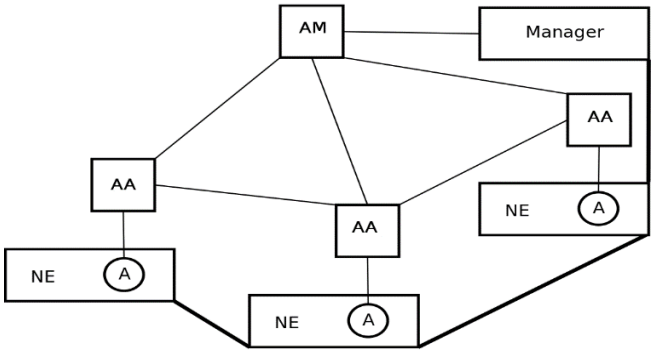


Abb. 3. Vorgeschlagene Struktur des Managementsystems

Ein Hilfsmanager (AM – Auxilary Manager) kann verwendet werden, um komplexere Richtlinien zu berücksichtigen und AAs zu koordinieren.

Aus der Perspektive eines Agenten werden Nachrichten an einen AA gesendet und empfangen, der sich in der gleichen Weise wie der Manager verhält. Aus diesem Grunde sind keine Änderungen in der inneren Logik der Agenten notwendig. Dasselbe gilt für Fernmanager.

Dies beseitigt jedoch keineswegs die Koexistenz zwischen einer direkten A-M und einer A-AA-M-Kette. Abb. 4 zeigt eine weitere Möglichkeit zur Modifikation der Kette, wobei AAs als Alternativen zu Agenten verwendet werden können.

In diesem Fall ist eine Teilmenge der Funktionalitäten des Agenten in das AA eingebettet, so dass es unabhängig überwachen und mit den verwalteten Ressourcen interagieren kann.

Dies stellt das Modell aus einer Sicht dar, die auf dem Managementaspekt basiert. Allerdings sollten auch Bedenken in einer Weise verteilt werden, die für IoT natürlich ist, um einige wichtige Vorteile zu nutzen. Agenten können die Sensor- und

Aktor-Rollen verkörpern, da sie bereits äquivalente Funktionalitäten im traditionellen Manager-Agent-Paradigma enthalten. AAs wird die wichtigste Asset-Device-to-Device-("agent-to-agent"), Konnektivität bereitstellen und eine Out-of-Band-Kommunikation zwischen lokalen Elementen ermöglichen

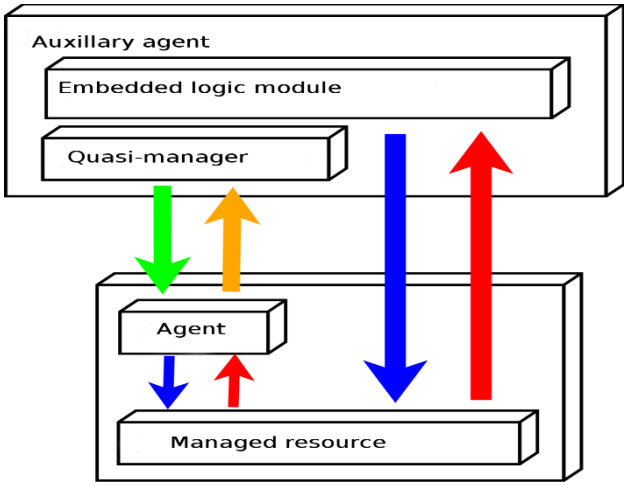


Abb. 4. Vorgeschlagene Struktur des Informationsaustauschsystems in den Modulen 4, 8, 10, 11 und 12.

Die zu menagierenden Ressourcen sind die Struktur der Module, die die vorgeschlagene Struktur unter denen Informationen austauschen lässt, die Funktionalität verwaltet und die Parameterwerte konfiguriert oder ändert.

IV. Schlussfolgerungen und zukünftige Arbeiten

Die gegenwärtige Entwicklung der Informationstechnologien macht es möglich das Informationsmanagement in der Industrie ein rasches Eindringen in Gebiete, die auf erster Sicht nicht mit der Informationsverwaltung zu tun haben. Die Produktionsprozesse werden immer mehr komplexer und es werden passende Methoden, Verfahren und Werkzeuge der Herausforderungen entgegenzukommen.

Diese Arbeit bietet ein integriertes Verfahren zur gemeinsamen Mitwirkung unterschiedlichen intelligente Techniken zu Erhöhung der Effizienz der Produktionsmaschinen. Der Vorschlag umfasst Erstellung neuer Maschinenstruktur, die selbst intelligente Funktionskomponente beinhaltet, Anwendung von Heuristiken um optimale Parameter zu bestimmen und am Ende – Ansatz einer Managementarchitektur, die gleichzeitig den Informationsaustausch und das Management der intelligenten Maschine ausführt.

Als weitere Aufgaben, die zu lösen geblieben sind, kann man die Erstellung geeigneten Datenstrukturen, Datenbehandlungsverfahren und Optimierung des Informationsaustauschprozesses in Richtung Echtzeitfunktion und durch maschinelles Lernen Erzeugung neuer Maschinenstrukturen, bezeichnen

Referenzen

[1] Mehrabi, M., A. G. Ulsoy, Y. Koren, “Reconfigurable Manufacturing Systems and Their Enabling Technologies”, *International Journal of Manufacturing Technologies and Management*, Vol. 1, № 1, 2000, pp. 113-130

[2] Koren, Y., Heisel, U., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G., and Van Brussel, H., “Reconfigurable Manufacturing Systems”, *Annals of the CIRP*, 1999, Vol. 48/2, pp. 527–540

[3] Guergov, S., “Integriertes verteiltes selbstorganisiendes System für mechanische Bearbeitung,” *IVth International Congress “Mechanical Engineering Technologies’04”*, September 23-25.2004, Varna, Proceedings, Volume 5, pp. 149-154

[4] Guergov, S., “Multioperations technological System with complex functions, reconfigurable structure, high integrity and self-organizing elements,” *Proceedings V International Congress“ Computer Science for Design and Technology”*, 5-7 Oct., 2005, Stankin, Moscow, Russia, pp. 31-36

[5] Guergov S., “Information System for Data Selection and New Information Acquisition for Reconfigurable Multifunctional Machine Tools,” *International Journal Of Intelligent Technology*, Vol. 2, No 1, 2007, ISSN 1305-6417

On the issue of technology process for intelligent machine tools

Auf die Frage der Automatisierung der technologischen Vorbereitung für intelligente Werkzeugmaschinen

Sasho Guergov*, Aleksandar Tsenov†

Technical University of Sofia, Bulgaria,
*Faculty of Mechanical Engineering, email: sguergov@tu-sofia.bg,
†FDIBA, email: aleksandar.tsenov@fdiba.tu-sofia.bg

Abstract — In recent years the industrial progress is with the platform "Industrie 4.0" mentioned at the trade fair in Hannover in 2013. The concept is based on leading solutions from SIEMENS for digital product development and production. These solutions are designed to address the conditions, as well as the automation of the entire life cycle of a product, as well as the integration of cyber-physical systems, which represent a virtual representation of the real world in order to manage physical production processes and to make distributed decisions itself.

Zusammenfassung — In den letzten Jahren der industrielle Fortschritt ist mit der, auf der Messe in Hannover im 2013 erwähnten Plattform „Industrie 4.0“. Das Konzept ist basierend auf führenden Lösungen von SIEMENS für digitale Produktentwicklung und Produkterstellung. Diese Lösungen sollen die Bedingungen sowie zur Automatisierung des gesamten Lebenszyklus eines Produkts, als auch zur Integrierung von Cyber-Physical Systemen, die eine virtuelle Darstellung der reellen Welt abbilden um die physikalische Produktionsprozesse zu verwalten, um zu kommunizieren und um selbst verteilte Entscheidungen zu treffen.

I. Introduction

The enhancement of modern machine-building production is closely related to the development of technology as a result of the development of information technology, technics and the ever-increasing requirements of the market.

The design of the technological processes is one of the main activities of the technological preparation of the production, which at the present stage of development is characterized by a number of peculiarities resulting from the changes in the mechanical engineering production. Hence, there are a number of requirements for manufacturers to increase the level of automation of all technological processes in the field of small and serial production, and to shorten the timeframe for technological preparation as well as the whole production cycle.

One of the most important procedures in order to fully automate the technological preparation for the realization of a manufacturing process for machining is:

1) Formalization and identification of the production site;
2) Addressing and assigning the object to a relevant class of details;
3) Choice of type / group technological process from database;
4) Selection of sequence of technological operations / route technology from the database.

The purpose of this development is to demonstrate the feasibility of implementing the above procedures by using popular software such as Microsoft Office Access 2003-SP2.

II. Formulation and Implementation

The identification of the object to be processed is carried out by its parametric description, addressing and assignment to the relevant type of detail in the database for which the relevant type technology (DB_3) has been developed. The description of the object can be performed based on typical primitives of typical surfaces (cone, cylinder, parallelepiped, sphere, etc.) by parameterization of its main characteristics to a given coordinate system so that it is uniquely identified in the database.

If the object of the production task is designated by X, it can be represented as a union of the classes of the rotational symmetric parts KL_1, the planar surfaces PS, the additional primitives DP1 and the plurality of the prismatic-shell elements KL_2, i.:

$$X \in KL1 \cup KL2. \quad (1)$$

The statistical distribution of the details in machine building is relatively constant over time, with the relative share of rotational symmetrical details and prism-body details being greatest (about 90%). Therefore, more attention will be paid below them.

The set of rotational symmetrical details (KL_1) can generally include subsets of primitive parts with different surfaces: cylindrical (C), cone (Co), spherical (S) or planar (P) surfaces. The subdivision of these subsets is necessitated by the fact that the respective surfaces are analyzed analytically with different parameters and this facilitates their formalization.

In that regard, it follows that

$$KL_1 = \{C, Co, S, P\} * DP_1. \quad (2)$$

The set of additional primitives can be represented by the expression:

$$DP1 = Ph \cup K \cup KZP \cup ShK \cup ShL \cup RRO \cup ZRO \cup O \cup R, \quad (3)$$

where: Pf is the set of chamfers; K - the set of channels; KZP - the set of zigzag ring grooves; ShK - the set of key channels; ShL - the set of slots; RRO - the set of radially spaced holes; ZRO - a set of centrally located openings; O - the set of holes displaced by the symmetry axis; R - the set of threads.

Each code description of the object of type (2) corresponds to the code KD_i^3 and address $A_i^{KD^3}$ in DB_1.

The plane / detail (PS) parameterization is considered as a composite figure {i, j}, calculated as the sum of the parameters of the individual elements i and j in the right-aligned two-coordinate system (corresponding to the orientation and layout of the work piece).

Each code description of the object corresponds to the code KD_i^2 and address $A_i^{KD^2}$ in BD_2.

In Figure 1 is given the organization of DB_1 for parameterization and description of the object processing unit of type KL_1. The structure of DB_1 consists of three parts: database name, catalog, and memory. The name of the database contains data about the physical address of the catalog, memory address, and free space and memory address. In this way, it is possible to analyze the correctness of storing information and possibly overloading the memory.

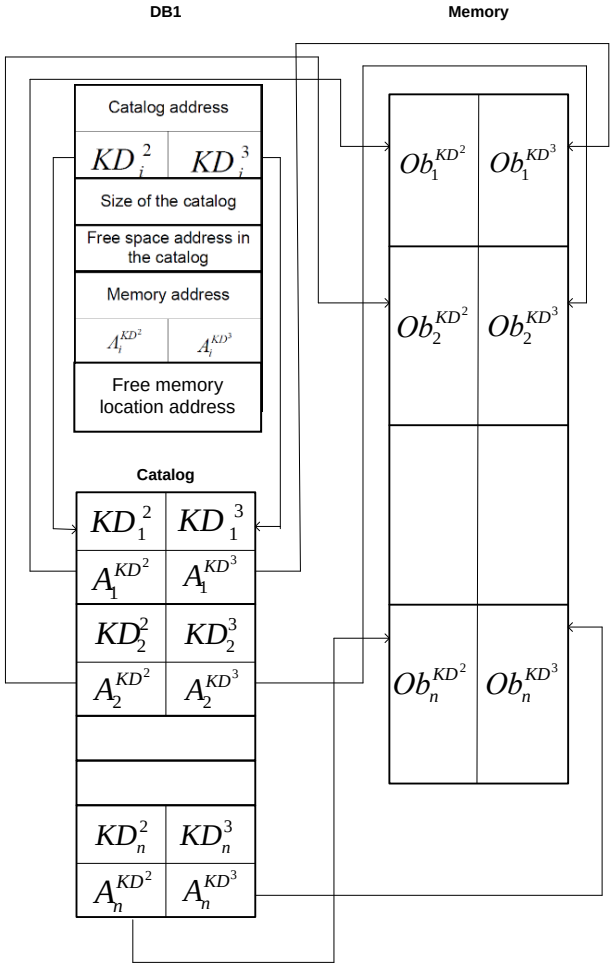


Fig. 1. Organization of DB_1 for parameterization and description of object processing unit of type KL_1

The DB_1 catalog is intended to determine the match between the block name symbol and its physical address. It has a sequential file structure, each record containing a symbolic name of the parametric object model, the physical address in the memory, and the length of the block in which the object model is recorded.

For complex objects of the prismatic-corporeal details, the formalisations are made by primitives (Pr); surfaces (Su); ribs (Ri); resulting from the intersection of two surfaces; vertices (V) - points obtained from the intersection of three surfaces and a boundary (D) corresponding to the line of the line graph defining the order of the peaks. In this way, the object model consists of two parts: B_i, Xi, Yi, Zi coordinates and a topology of their union set by cycles or boundary outlines in the order of their crawling with a type of cycle (surface or hole entry).

The mathematical model M of a complex prismatic-corpus piece (PR) composed of n primitive can be represented by the dependence

$$M = \bigcup_{i=1}^{n} PR_i \tag{4}$$

Obtaining the coordinates of the points of attachment of the individual primitives needs to be transformed into the coordinate system of the final object (FO), i.

$$Xcv = M_{transf} M_{rot} B_{\kappa.vertices} \tag{5}$$

where X_{cv} is the vector of the coordinates of the vertices of the Pr in the coordinate system of the FO; M_{transf} - a matrix of 4 x 4 for transforming *Pr*; M_{rot} - a matrix of 4 x 4 of the rotation of the coordinate system of the Pr in that of the FO; $B_{\kappa.vertices}$ - a vector of the coordinates of the vertices of the PR in its own coordinate system.

Thus, the position of Pr_i is determined by six numbers: the transform vector $\vec{T}_i = (\Delta x, \Delta y, \Delta z)$ and the angles of rotation $(\alpha_i, \beta_i, \gamma_i)$ of the coordinate system relative to that of the FO. Hence, in the node B of the G column of the FO, the origin of the coordinate systems of the primitives is located, and the frames of the graph Pg depict the type of the relations between the primitives.

The algorithm for parametric description of complex KL_2 objects is given in Figure 2.

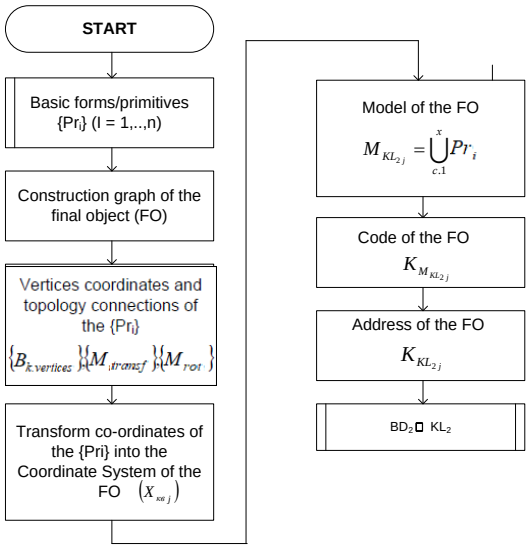


Fig. 2. Block diagram of the algorithm for parametric description of complex objects of type KL_2

In organizational terms, the structure of DB_2 is analogous to that of Figure 1.

III. PRACTICAL IMPLEMENTATION

In this part the Assignment a site to relevant details and class and the data organization for technological processes is described. Figure 3 shows the general program structure and the input screen for describing the parameters.

On the right side of the screen, there is an example of the developed in the Microsoft Office Access 2003-SP2 environment.

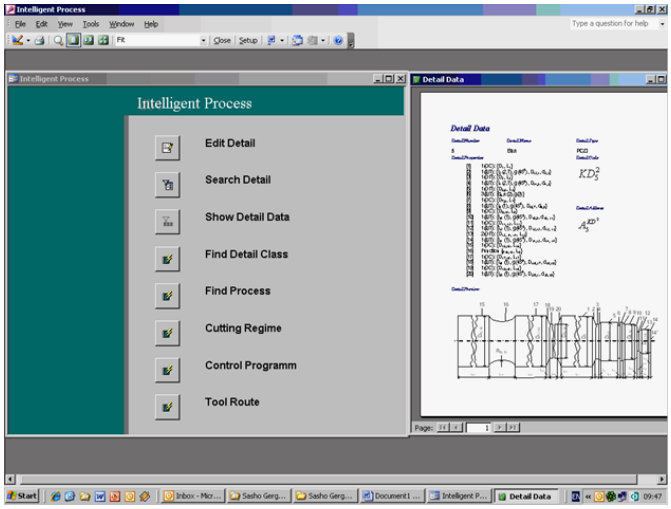


Fig. 3. Input data to describe the detail

On each parameterized object KL_{1i} or KL_{2i}, as seen above, an appropriate code $A_{KL_{1i}}$ or $A_{KL_{2i}}$ is assigned. With the help of the detail code and the corresponding search procedure, there is a BD_2 object that contains two parts - respectively for objects of class rotational parts (KL_{1i}) or prismatic parts (KL_{2i}).

The object code A_{KL} is described by elements of algebraic parameterization as a system

$$A_{KL\,ij} = \{E_{KL_{ij}}\}, \tag{6}$$

where $\{E_{KL_{ij}}\}$ is the set of objects elements..

If there is an analogue detail in BD2, dependence (6) accepts the species

$$A_{KL_{ij}} = (e_1, e_2, \ldots, e_n) \tag{7}$$

where e_i are the elements of the j object of the i-th kind of details.

In the case of incomplete matching of the elements e_i with those of the analogue part, the new detail is obtained by adding or subtracting corresponding elements from the closest analogue object (detail-type representative).

The search procedure $P_{KL}$is the type

$$P_{KL} = R_{KL} P_{KL,i}, \tag{8}$$

where R_{KL} is a ranking procedure and $P_{KL,i}$ is the procedure for searching an object with identical address.

The ranking procedure helps to reduce the analyzed set of analog objects. Ranking is done in accordance with the rules.

$$(N_n, B_n) = \{A_{KL}\} \tag{9}$$

where N_n are the number of elements (parameters) in the algebraic description of the object and B_n is the type of elements in the algebraic description of the object.

The input data for describing the part parameters apart from serving for its identification and based on them, it receives the corresponding code and address in the BD_1 (Fig.4).

Through the code and address thus obtained and the corresponding procedure an analogue object of the relevant class KL_{1i} (rotational parts) or KL_{2i} (prismatic parts) is searched in DB_2 (Fig.5).

In the event of finding such an analogue, the type of technological process is detected by the BD_3 workpiece address (Figure 6).

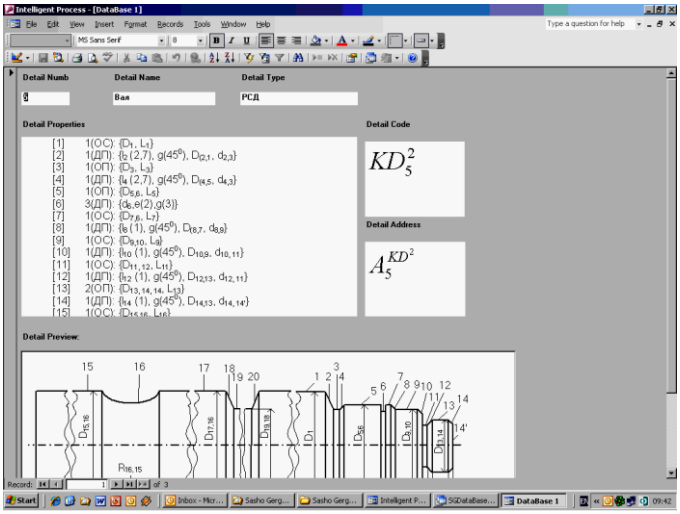


Fig. 4. Parameters, code and batch address of BD_1

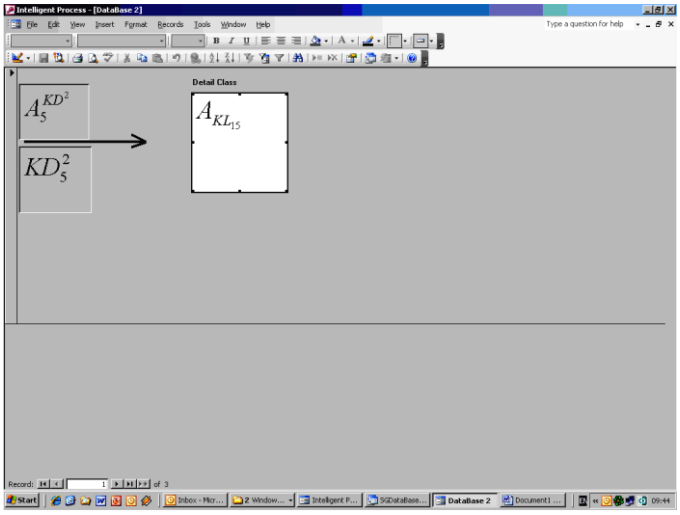


Fig. 5. Assigning the workpiece to a relevant class (BD_2)

Next, it is necessary to determine the sequence of the route of the technological operations (BD_5), but because for the given example they are only turning in Fig. 7, only the cycle of movement of the tool in the workpiece is shown.

In the case of an undetected detail-analog in DB_2, a control procedure is performed by the M-CREATE operator and by appropriate procedure a new object is created, which is recorded in the BD_2. The procedure is then repeated for the other DBs (BD3 and BD_4).

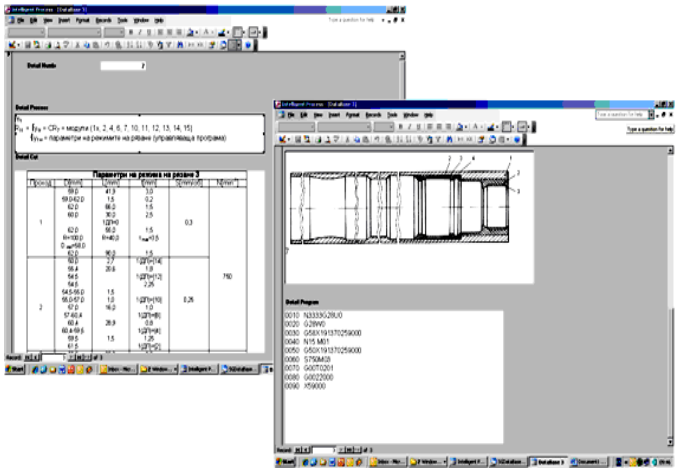


Fig. 6. Choosing a Type / Group Technology Process

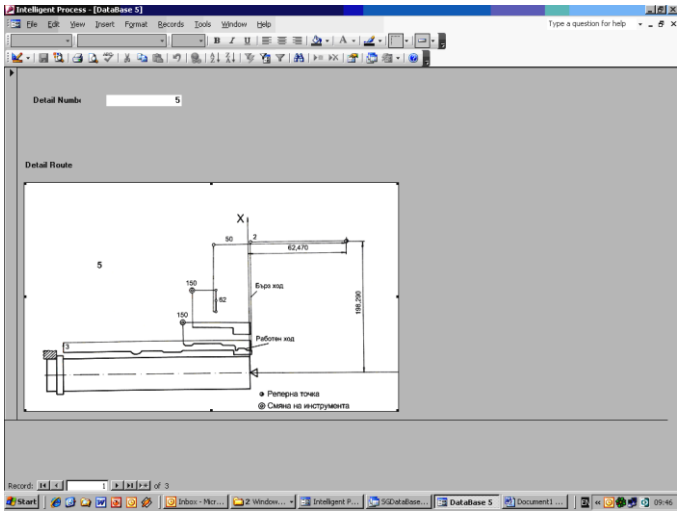


Fig. 7. Cycle of tool movement (BD_5)

IV. Conclusions

The current development of information technology makes it possible for information management in the industry to penetrate rapidly into areas that are not related to information management at first sight. Production processes are becoming more and more complex, and appropriate methods, procedures and tools will meet the challenges.

This work provides an integrated process for jointly contributing different intelligent techniques to increase the efficiency of the production machines. The proposal includes the creation of new machine structure, which includes intelligent function components, application of heuristics to determine optimal parameters, and finally, the approach of a management architecture that simultaneously performs the information exchange and management of the intelligent machine.

Other tasks that can be solved are the creation of appropriate data structures, data handling procedures and the optimization of the information exchange process in the direction of real-time function and by machine learning the creation of new machine structures

References

[1] Mehrabi, M., A. G. Ulsoy, Y. Koren. “Reconfigurable Manufacturing Systems and Their Enabling Technologies”, *International Journal of Manufacturing Technologies and Management*, Vol. 1, № 1, 2000, pp. 113-130

[2] Koren, Y., Heisel, U., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G., and Van Brussel, H. “Reconfigurable Manufacturing Systems”, *Annals of the CIRP*, 1999, Vol. 48/2, pp. 527–540

[3] Guergov, S. “Integriertes verteiltes selbstorganisiendes System für mechanische Bearbeitung,” *IVth International Congress “Mechanical Engineering Technologies’04”*, September 23-25.2004, Varna, Proceedings, Volume 5, pp. 149-154

[4] Guergov, S. “Multioperations technological System with complex functions, reconfigurable structure, high integrity and self-organizing elements”, *Proceedings V International Congress“ Computer Science for Design and Technology”*, 5-7 Oct., 2005, Stankin, Moscow, Russia, p. 31-36

[5] Guergov S., “Information System for Data Selection and New Information Acquisition for Reconfigurable Multifunctional Machine Tools,” *International Journal Of Intelligent Technology*, Volume 2, Number 1, 2007, ISSN 1305-6417

Estimation of Magnetic Field Maps With Mobile Platforms

Erstellung von Magnetfeldkarten mit mobilen Plattformen

Stefan Hensel*, Marin B. Marinov‡

* University of Applied Sciences Offenburg, Department for Electrical Engineering,
Badstraße 24, D-77652 Offenburg, Germany, stefan.hensel@hs-offenburg.de
‡ Technical University of Sofia, Faculty of Electronic Engineering and Technologies, Department of Electronics,
8, Kliment Ohridski Blvd., BG-1756 Sofia, Bulgaria, mbm@tu-sofia.bg

***Abstract* —** The aim of this project is to study the extent to which local magnetic fields can be used for the exact localization in indoor environments, for example in railway stations, airports or shopping centers. For this purpose, we investigate whether and how low-cost magnetic field sensors and mobile robot platforms can be used to create maps that improve the accuracy and robustness of later navigation with smartphones or other devices.

***Zusammenfassung* —** Für die genaue Positionsbestimmung in Innenräumen, beispielsweise in Bahnhöfen oder Einkaufszentren, soll in dem beschriebenen Projekt untersucht werden, inwiefern lokale Magnetfelder genutzt werden können, um Genauigkeit und Robustheit zu erhöhen. Hierzu wird untersucht, ob und wie kostengünstige Magnetfeldsensoren und mobile Roboterplattformen genutzt werden können, um Karten zu erstellen, die eine spätere Navigation, beispielsweise mit Smartphones oder mit anderen mobilen Geräten.

I. Einführung

Während die Positionsbestimmung unter freiem Himmel seit längerer Zeit auf etablierte Weise durch eine Kombination von Satellitennavigation und Inertialsensorik erfolgt, ist die präzise Ortsbestimmung innerhalb von Gebäuden weiterhin der Gegenstand aktiver Forschung. Im Labor für Mess- und Sensortechnik erfolgt die Untersuchung von kostengünstigen Sensoren zur Verringerung oder Vermeidung der unweigerlich auftretenden Drift der inertialen Navigation.

Das natürlich auftretende Erdmagnetfeld wird in Gebäuden durch die verwendeten Baumaterialien und Einrichtungsgegenständen teils massiv verzerrt. Diese Störungen der magnetischen Flussdichte $\vec{B}$ sind lokal stark ausgeprägt und weitestgehend stationär. Sie können daher als Merkmal für die Positionsbestimmung verwendet werden [1]. Vorab kalibrierte Magnetfeldsensoren messen diese Merkmale und verwenden sie als Orientierungshilfe in der Navigation [2].

Die von den Sensoren gewonnene Information wird dem Nutzer sinnvoller Weise in Form einer Karte zur Verfügung gestellt. Das beschriebene Projekt untersucht, wie diese Karten mit Hilfe mobiler Roboter erstellt werden können. Besonderes Augenmerk wird auf das Zusammenführen mehrerer individuell erstellter Karten gelegt, so dass entweder ein Schwarm von Robotern eingesetzt werden kann oder ein einzelner Roboter die Karten in mehreren Durchläufen erstellen kann.

II. Methodik

Für die gegebene Problemstellung wurde die Kartenerstellung in zwei parallele Stränge aufgeteilt. Die Positionsbestimmung des Roboters erfolgt durch die Fusion von Inertialsensoren, d.h. Beschleunigungs- und Drehratensensoren, mit Odometriedaten und den Entfernungsdaten aus einer Tiefenkamera. Der eingesetzte Algorithmus ist ein Partikelfilter, der eine simultane Lokalisierung und Kartierung (Simultaneous Localization and Mapping - SLAM) ermöglicht und als Ausgangsdaten die Position sowie eine zweidimensionale Rasterkarte der Umgebung liefert. Aufbauend auf dieser Position wird mit Messdaten eines handelsüblichen Magnetfeldsensors die Magnetfeldkarte berechnet.

Die Erstellung der Karte muss die Unsicherheit der Position aufgrund von Umgebungseinflüssen und verrauschten Sensoren ebenso berücksichtigen wie die unsicherheitsbehafteten und störanfälligen Messungen des Magnetfeldsensors. Da die Messungen des Sensors punktuell auf der gefahrenen Trajektorie des Roboters erfolgen, muss eine Interpolation und Extrapolation zwischen den Punkten und über diese hinaus durchgeführt werden. Die Aufgabe der Kartierung mit geringer Datenmenge ist in Abb. 1 exemplarisch für den vorliegenden zweidimensionalen Fall dargestellt.

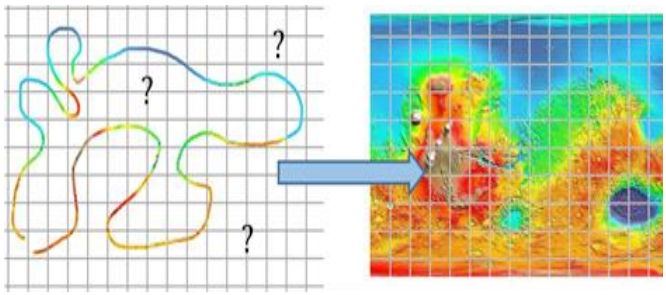


Abb. 1. Abb. 1. Prädiktion der magnetischen Flussdichte $\vec{B}$ in unbekannten Bereichen auf Basis verteilter Messungen

Für die Vorhersage dieser Werte und die Verarbeitung der zugehörigen Unsicherheit in Form einer Kovarianz wurde in der Arbeit der Einsatz von sog. Gaußprozessen untersucht.

A. *Gaußprozess-Regression und Interpolation*

Gaußprozesse [3] sind eine Verallgemeinerung der multivariaten Gaußverteilung und können genutzt werden, um eine stetige und glatte Funktion aus Messdaten zu schätzen. Ein Spezialfall der Gaußprozess-Regression (GPR), der auch in der vorliegenden Arbeit Verwendung findet, ist in der Geostatistik

unter dem Namen Kriging bekannt. Daneben können Gaußprozesse auf eine Vielzahl Probleme des maschinellen Lernens angewandt werden und finden Einsatz in der Glättung mit Splines, Klassifikationsproblemen oder der Datenvorhersage. Ein veranschaulichendes Beispiel für die Regression auf Basis weniger Messungen ist in Abb. 2 für den Fall von einer, zwei und drei Messungen dargestellt.

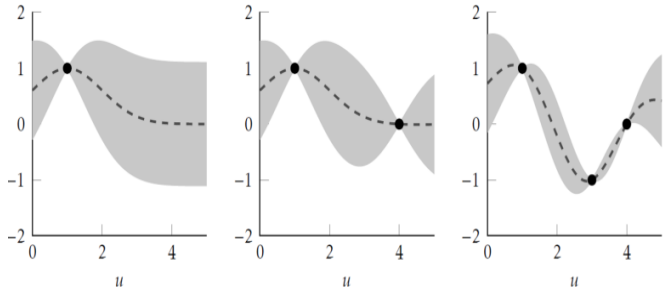


Abb. 2. Prinzip der Gaußprozess Regression.

Messpunkte sind schwarz dargestellt, der grau mattierte Bereich gibt das sog. 95% Konfidenzintervall (die zweifache Standardunsicherheit der geschätzten Funktion) auf Basis der gegebenen Messwerte wieder.

In der hier verwendeten GPR werden die Modellfunktionen $f(\boldsymbol{x})$ als Realisierungen eines stochastischen Prozesses gesehen, dessen Kovarianzfunktion $\kappa(\boldsymbol{x}, \boldsymbol{x}')$ a priori vorgegeben wird. Die Bestimmung der Funktion erfolgt auf Basis von n gegebenen und unsicheren Messungen $y_1, y_2, \dots, y_n$, die an den Stellen $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ beobachtet werden. Man erhält dann das Modell

$$f(\boldsymbol{x}) \sim GP(0, \kappa(\boldsymbol{x}, \boldsymbol{x}')), \qquad y_i = f(x_i) + \varepsilon_i, \tag{1}$$

bei dem angenommen wird, dass der Fehler ε einem weißen Rauschen entspricht. Damit sind das Mess- und a priori Modell normalverteilt und das geschätzte a posteriori Modell wird wiederum normalverteilt sein.

B. *Wahl der Kovarianzfunktion*

Die Wahl einer spezifischen Kovarianzfunktion drückt das Domänenwissen über den unterliegenden Prozess aus, im vorliegenden Fall über das Verhalten des Magnetfeldes in geschlossenen Räumen. Aufgabe der Inferenz ist es, aus den Messdaten die Parameter dieser Kovarianzfunktion zu schätzen. Die Parameter werden hierbei als sogenannte Hyperparameter bezeichnet, der Typ der Kovarianzfunktion wird in Form von Modellwissen a priori festgelegt. Eine der meistgewählten Kovarianzfunktionen ist eine stationäre Exponentialverteilung (auch radiale Basisfunktion RBF) der Form

$$\kappa_{RBF}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_{RBF}^2 \cdot \exp\left(-\frac{\|\boldsymbol{x}, -\boldsymbol{x}'\|^2}{2l_{RBF}^2}\right) \tag{2}$$

mit den Hyperparametern σ_{RBF}^2 und l_{RBF}, die die Skale und Granularität (Bandbreite) der Funktion ausdrücken.

Ein Beispiel für den Einfluss des Hyperparameters l_{RBF} auf die geschätzte Funktion mit Konfidenzintervall ist in Abb. 3 gezeigt.

Die Schätzung der Hyperparameter auf Datenbasis erfolgt im optimalen Sinne durch Bestimmung der Maximum a posteriori Wahrscheinlichkeit.

Die Verwendung von Gaußprozessen ermöglicht nicht nur die Einbindung aller in der Kartierung auftretenden Unsicherheiten, sondern erlaubt die anschließende Fusionierung der erstellten Karten auf Basis der geschätzten Genauigkeit. Dies ist möglich, indem die Karte jeweils mit allen Messungen neu berechnet wird oder indem mehrere interpolierte Karten durch eine gewichtetes arithmetisches Mittel auf Basis der Kovarianz zusammengeführt werden, die zu beliebigen Zeiten und mit beliebigen Sensoren aufgezeichnet wurden.

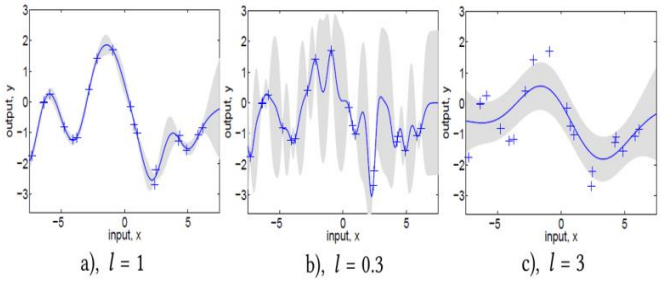


Abb. 3. Einfluss des Hyperparameters l auf die Schätzung der Funktion [3].

III. Versuchsaufbau

A. *Hardware*

Turtelbot

Für die Erprobung wurde ein mobiles Robotersystem vom Typ Turtlebot verwendet. Als Middleware kam das Robot Operating System (ROS) [4] zum Einsatz, das eine einfache Treiberanbindung erlaubt. Das ROS ist eine open source Softwareumgebung, welche Entwicklern verschiedene Bibliotheken und Werkzeuge zur Verfügung stellt, um die Entwicklung von Anwendungen in der Robotik zu erleichtern. So werden zahlreiche Hardwareelemente von Herstellern vertrieben, wofür es Gerätetreiber unter ROS gibt.

Ein Turtlebot ist eine kleine Roboterplattform im Niedrigpreissegment. Der Roboter wird, wegen seine Flexibilität und seinem einfachen Zugang, überwiegend für Realisierbarkeitsstudien und in der Forschung oft eingesetzt. Er besteht aus einer flachen mobilen Trägerplattform, dem Kobuki von Yujin Robot, einer Kinect von Microsoft sowie einem Aufbau um Platz für Laptops, weitere Sensoren oder sonstige Aufbauten zu bieten.

Inertiale Messeinheit – Razer IMU 9 Degree of Fredom

Die Lokalisierung des Systems wurde mit dem *gmapping* SLAM-Algorithmus [5] realisiert. Für die Aufzeichnung des Magnetfeldes wird ein Dreiachs-MEMS-Magnetfeldsensor genutzt, wie er beispielsweise in Smartphones zu finden ist. Hier wurde ein HMC5883L drei-Achsen Magnetfeldsensor verwendet. Zusätzlich ist eine vollständige inertiale Messeinheit (IMU), bestehend aus einem Dreiachsen-Drehratensensor und Dreiachsen-Beschleunigungssensor, in dem Chip integriert. Die IMU erlaubt über die Berechnung der Lage die Projektion der dreidimensionalen Magnetfeldmessung in ein globales Koordinatensystem und damit die Bestimmung der Flussdichtekomponenten $\vec{B} = (B_X, B_Y, B_Z)^T$.

Ein weiterer Vorteil der IMU ist, dass, je nach Anwendung, ein Teil der Daten- und Signalverarbeitung bereits von der IMU auf einen ATmega328 Mikrocontroller übernommen werden kann. Die Auflösung des Magnetfeldsensors beträgt 0,5 μT bei einem Messbereich von $\pm 800\ \mu T$ und eine Abweichung von 0,2 μT.

Abb. 4 zeigt den verwendeten Roboter sowie den Versuchsaufbau mit Rechner und Magnetfeld-IMU-Sensor.

Abb. 4. Mobile Roboterplattform (links) und Versuchsaufbau mit Magnetfeldsensor-IMU-Kombination (rechts).

B. *Software*

Für die Schätzung der Parameter des Gaußprozesses werden komplexe Optimierungsverfahren benötigt, die in der Regel die Komplexität $O(n^3)$ aufweisen. Da für die Interpolation des Magnetfeldes hunderte bis tausende Messwerte vorliegen, wurde daher auf approximative Verfahren zurückgegriffen. Als a priori Annahme der Messfunktion wurde ein Gauß'sches Rauschen angenommen, die Kovarianz wurde a priori zu einer Matérn-Funktion festgesetzt, die eine generalisierte Version der RBF darstellt und einen zusätzlichen Hyperparameter zur Festlegung der Glattheit enthält [6]. Die Berechnung erfolgte offline im Batch-Verfahren, d.h. für den gesamten Datensatz nach der Aufzeichnung der Messwerte. Eingangsdaten sind die Magnetometermessungen $\vec{B}$ zusammen mit der Position, Orientierung und der Kovarianzmatrix, die durch den SLAM-Algorithmus bestimmt wurden.

Datenvorverarbeitung

Die magnetische Flussdichte wurde in x ,y und z gemessen $\vec{\boldsymbol{B}} = (B_X, B_Y, B_Z)^T$. Um das Signal vorerst von Messrauschen und einzelnen Ausreißern zu bereinigen, wurde es mit einem Medianfilter 5. Ordnung geglättet. Da sich das Magnetfeld jedoch durch die begrenzte Geschwindigkeit des Turtlebots nicht besonders schnell ändern kann, wurden anschließend die Messungen mit einem Glättungsfilter geglättet. Die Größe des Kernels variiert hierzu zwischen den einzelnen Messungen und muss gelegentlich angepasst werden. Für einen Großteil der Messungen im Labor wurde ein Kernel der Ordnung $n \approx 100$ gewählt, da hier ein sehr unregelmäßiges Magnetfeld vorliegt. In Bereichen, in denen größtenteils ein gleichmäßiges Magnetfeld vorhanden ist und es nur einzelne Regionen gibt in denen das Magnetfeld stark schwankt, muss der Filterkernel verkleinert werden um Messungen nicht vollständig zu entfernen. Im Foyer der Hochschule Offenburg wurde hierfür ein Kernel der Größe $n \approx 20$ gewählt.

Da die IMU und die Odometrie des Turtelbots mit unterschiedlichen Abtastfrequenzen abgerufen werden, mussten diese, obwohl sie bereits durch ROS einen gemeinsamen Zeitstempel haben, einander noch zugewiesen werden [7].

Kartierung

Um nun aus der gefahrenen Bahn eine flächige Magnetfeldkarte zu erstellen, existieren einige Interpolationsverfahren. Matlab selbst bietet mit der Funktion *griddata* die Möglichkeit, Funktionen $f(x, y)$ über ein zuvor definiertes Gitter zu interpolieren. Die hier verwendete Methode basiert auf Delaunay-Triangulation. Hierbei werden Dreiecke zwischen Punktepaare im Raum gebildet. Diese Dreiecke sind so gewählt, dass lange, dünne Dreiecke vermieden werden. Da die Dreiecke das gesamte Gitter abdecken und sich nicht überlappen, kann jedem Punkt auf dem Gitter nun ein Wert zugewiesen werden [8].

Kartenfusionierung

Um nun aus mehreren Messreihen eine Karte zu erzeugen, ist eine Fusion der Messwerte nötig. Diese können entweder durch das mehrfache Abfahren einer Umgebung mit einem einzelnen Roboter oder durch paralleles Abfahren einer Umgebung mit mehreren Robotern erzeugt werden. Für eine Interpolation mit Gaußprozessen gestaltet sich diese Fusionierung einfach. Es können dem Gaußprozess alle Messwerte mit einer zugeordneten Unsicherheit als Trainingsdaten zugefügt werden. Die daraus resultierende Interpolation ist bereits gewichtet und liegt einer Verteilungsfunktion zugrunde. Es macht hierbei keinen Unterschied, ob dieselbe Stelle mehrmals bei einer Messung durchfahren wird, oder ob mehrere Messungen fusioniert werden. Beides hat zur Folge, dass die entstehende Karte Messungenauigkeiten oder Messfehler weniger stark gewichtet.

Bei der gewöhnlichen Interpolation würde ein solches Vorgehen allerdings nicht zum Richtigen Ergebnis führen, da jeder Wert mit der gleichen Gewichtung in das Ergebnis eingeht.

Um nun mehrere Karten gemäß ihrer Unsicherheit zu fusionieren, wird jedem Punkt $\tilde{P}$ neben einem Wert z_{mag} der die magnetische Feldstärke an der Position x und y ausdrückt, ein Wert σ_{max} zugewiesen:

$$\tilde{P}_{map} = \tilde{P}_{map}(x, y, z_{mag}, \sigma). \tag{3}$$

Sind Messungen vorhanden, die die gesamte Fläche im Raum dicht abdecken, lässt sich die Kartenfusionierung mittels eines gewichteten Mittelwertes durchführen. Die Vorteile dieser Methode gegenüber den Gaußprozessen liegen hierbei in der geringeren Rechenlast. Ein grundlegendes Problem stellt allerdings die Gewinnung der Messdaten in ausreichender Dichte dar, so dass auf eine intelligente Inter- und Extrapolation der Messwerte nicht verzichtet werden kann [9].

IV. Experimentelle Ergebnisse

In einer ersten Untersuchung wurde die Eignung der Gaußprozesse für die Interpolation der Messdaten betrachtet. Zum Vergleich wurde eine klassische zweidimensionale lineare Interpolation herangezogen. Obgleich letztere nicht die Möglichkeit der Kovarianzangabe und -schätzung bietet, ist die Berechnungszeit um mehrere Größenordnungen geringer.

Allerdings wird schnell deutlich, welch qualitativer Vorteil durch die Interpolation und Prädiktion mit Gaußprozessen erreicht werden kann. Abb. 5 zeigt die lineare Interpolation auf der linken Seite auf Basis der befahrenen Trajektorie, die in schwarz dargestellt ist. Das Resultat des GP für den gleichen Datensatz ist auf der rechten Seite dargestellt.

Für die Arbeit wurden Magnetfeldkarten auf Basis mehrerer Messfahrten in verschiedenen Bereichen der Hochschule aufgezeichnet. Abb. 6 zeigt exemplarisch die qualitativen Ergebnisse für einen kleinen Bereich des Labors für Mess- und Sensortechnik an der Hochschule Offenburg.

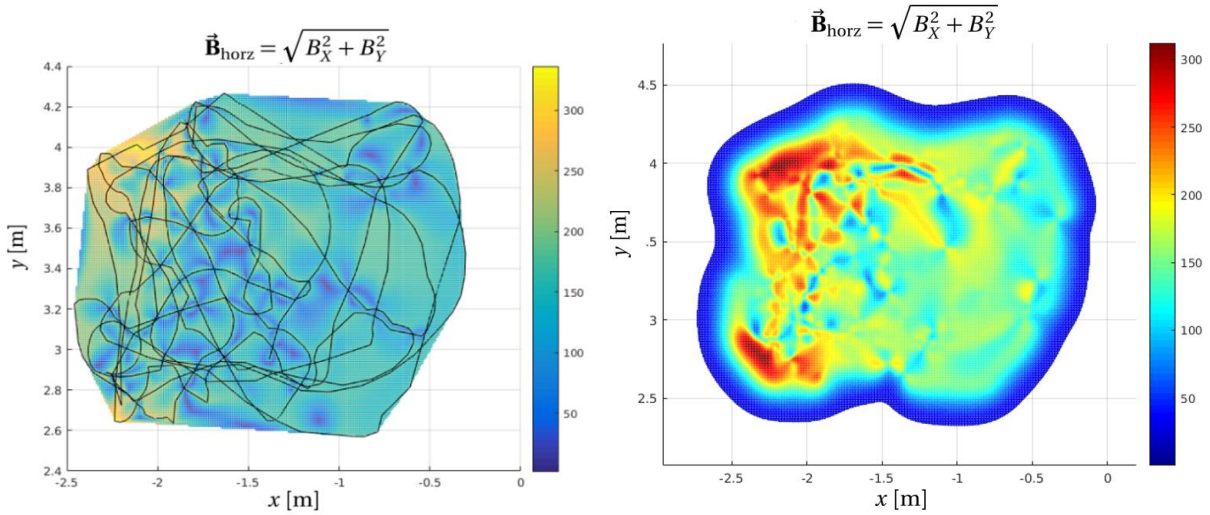


Abb. 5. Erstellte Karte der horizontalen Flussdichte.

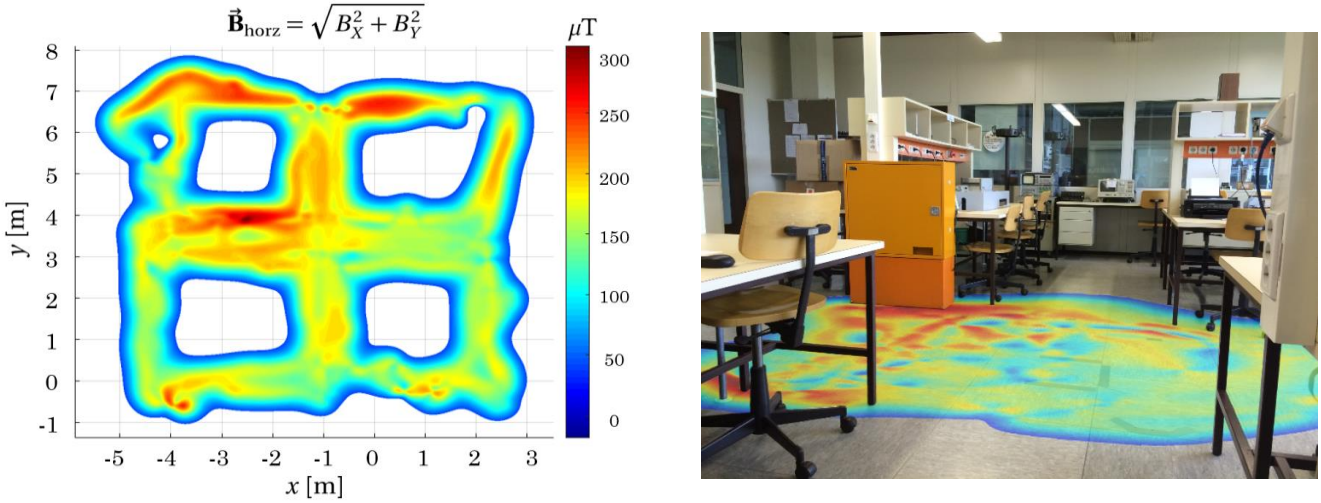


Abb. 6. Exemplarische Karte der horizontalen Flussdichte des Labors für Mess- und Sensortechnik.

V. Zusammenfassung

Die Untersuchung zeigt, dass die Gaußprozess-Regression grundlegend geeignet ist, um aus einzelnen Messungen eine vollständige und nutzbare Karte kontinuierlicher Merkmale zu erzeugen. Die erstellten Karten weisen einen guten Detaillierungsgrad auf und zeigen beispielsweise Schaltschränke oder Heizkörper. Die Vorteile gegenüber weniger aufwendigen Verfahren liegen in der einfachen Einbindung neuer Messungen, der Verfügbarkeit eines Qualitätsmaßes in Form einer Kovarianzmatrix und darauf basierend der natürlichen Integration in stochastische Lokalisierungsverfahren. Die Modellierung erlaubt zudem das Einbringen von Randbedingungen und Modellwissen. Dies alles geht allerdings zu Lasten einer schnellen Berechnung.

Die Karten wurden mit einer Auflösung von 5cm erstellt, das Potenzial für eine Positioniergenauigkeit liegt im Rahmen einiger Dezimeter bis Meter, was für die kontinuierliche Verbesserung und Driftkorrektur inertialer Sensoren bereits ausreicht.

Literaturverzeichnis

[1] B. Li, T. Gallagher, A. Dempster and C. Rizos, "How feasible is the use of magnetic field alone for indoor positioning?," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, Australia, 2012.

[2] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. Julian und M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and nonlegged locomotion“," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Tokyo, Japan, 2013.

[3] C. E. Rasmussen and C. K. I. Williams., Gaussian processes for machine learning, Cambridge, MA, USA: MIT press, 2006.

[4] M. Quigley, B. Gerkey and W. Smart, Programming Robots with ROS, Sebastobol, CA, USA, : O'Reilly Media, 2015.

[5] G. Grisetti, C. Stachniss and Wolfram Burgard, "Improved Techniques for Grid Mapping with Rao- Blackwellized Particle Filters," *IEEE Transactions on Robotics*, vol. 23, pp. 34-46, 2007.

[6] B. Minasny and B. McBradley, "Spatial prediction of soil properties using EBLUP with the Matérn covariance function," *Geoderma*, vol. 140, pp. 323-456, 2007.

[7] Z. Kokolanski, "Hardware techniques for improving the calibration performance of direct resistive sensor-to-microcontroller interface," *Metrology and Measurement Systems* , vol. 20, no. 1, p. 529–542, 2013.

[8] I. Rachev, "MATLAB Toolbox for Quality Assessment Processing in Optical Systems,," *Journal of the TU Sofia, branch Plovdiv "Fundamental Sciences and Applications"*, vol. 14, pp. 545-550, 2009.

[9] T. Tashev, V. Monov, R. Tasheva,, "High Performance Computations for Study the Stability of a Numerical Procedure for Crossbar Switch Node," in *Lecture Notes in Computer Science, International Conference on Numerical Analysis and Its Applications*, Springer, Cham, Vol. 10187, pp. 665-673, 2017.

Effect of Nanoadditives on the Characteristics of Bainitic Transformation in Spheroidal Graphite Cast Irons

Die Wirkung der Nanoadditiven auf die Eigenschaften der bainitischen Umwandlung in kugelgraphitischen Gußeisen

Julieta Kaleicheva[*], Valentin Mishev[*], Zdravka Karaguiozova[†]

[*]Faculty of Industrial Technology, Technical University of Sofia,
Sofia, Bulgaria, e-mail: jkaleich@tu-sofia.bg, v_mishev@tu-sofia.bg
[†]Space Researche and Technology Institute, Bulgarian Academy of Sciences
Sofia, Bulgaria, e-mail: karazuzi@yahoo.com

***Abstract* —** In the present study austempered ductile cast irons (ADI) with upper bainitic structure are investigated. Nanosized particles (50nm) of titanium carbite+titanium nitride (TiCN+TiN), titanium nitride TiN and cubic boron nitride cBN are added to the casting volume. The samples microstructure is studied by means of optical metallography, scanning electron microscopy and X-Ray analysis. The influence of the nanosized additives on the kinetic of the bainitic transformation and on the morphology of the bainitic structure is investigated. The abrasive wear testing, hardness measurements and impact strength are carried out. It is established that the presence of nanoadditives in the bainitic irons leads to the changes in their microstructure which increases their mechanical characteristics and abrasive wear resistance. The studied nanocomposite materials expand the potential for new ADI applications in the industry.

***Zusammenfassung* —**In dem vorliegenden Artikel werden isotherm gehärtete kugelgraphitische Gusseisen (ADI) mit oberbainitischer Struktur untersucht, die Naniaditiven (50nm) – Titancarbonitrid + Titannitrid (TiCN+TiN), Titannitrid TiN und kubisches Bornitrid cBN enthalten. Die Mikrostruktur der Proben wird durch optische Metallographie, Rasterelektronenmikroskopie und Röntgenstrukturanalyse erforscht. Der Einfluß der Nanoadditiven auf die Kinetik der bainitischen Umwandlung und auf die Morphologie der bainitischen Struktur der Gußeisen ist studiert. Es wird eine Untersuchung des abrasiven Verschleißes, der Härteprüfung und der Schlagzähigkeit durchgeführt. Es ist festgestellt, dass die Änderungen der Mikrostruktur der bainitischen Gußeisen unter der Wirkung der Nanoadditiven eine Erhöhung der mechanischen Kennzahlen und der abrasiven Verschleißfestigkeit der ADI verursachen. Die erforschten Nanokomposite erweitern die Möglichkeiten für neue industrielle Anwendungen der ADI.

I. Introduction

Bainitic irons are spheroidal graphite ductile cast irons with a bainitic structure of the metal matrix, obtained after austempering in the bainitic area. Austempering of the iron-carbon alloys in the bainitic area is a heat treatment with a wide practical application in the processing of the structural steels and ductile cast iron [1,2]. Austempered spheroidal graphite ductile irons with a bainitic structure combine high strength with increased toughness and plasticity. Alloyed spheroidal graphite irons have a high wear resistance due to a graphite phase and a bainitic metal matrix. Depending on the austempering mode, a lower bainitic structure (220÷280^{o}C) or upper bainitic structure (350÷450^{o}C) of cast irons are obtained.

Austempered spheroidal graphite irons (ADI) find application in the production of responsible details in the transport equipment and other industries.

The possibility for a wider application of these materials in the industry includes investigation of ADI with additions of a small quantity of nanosized particles during the casting process, which could improve their structure and properties.

Nanosized particles added to the casting of cast iron in small quantities can alter the graphite morphology from laminated to vermicular [3], also increase the amount of graphite and reduce the size of the graphite sphere [4], alter the structure of the metal matrix of cast irons and austempered ductile irons (ADI), which increases the mechanical and tribological properties of the cast irons [3-7].

The aim of this study is to investigate the peculiarities of the structure formation in the bainitic area, mechanical and tribological properties of spheroidal graphite cast irons containing additives of nanosized particles - titanium carbonitride + titanium nitride *(TiCN + TiN)*, titanium nitride *TiN* and cubic boron nitride *cBN*.

II. Materials and investigation methods

The composition of the spheroidal graphite cast iron samples is:Fe-3,55C-2,67Si-0,31Mn-0,009S-0,027P-0,040Cu-0,025Cr-0,08Ni-0,06Mg wt.%. The samples without and with nanoadditives of (*TiCN+TiN), TiN* and *cBN* are tested (Table 1). The *TiCN, TiN* and *cBN* nanosized particles are coated by electroless nickel coating EFTTOM-NICKEL prior to the edition to the melt [8]. The nickel coating improves the particles wetting into the melt and their uniformity distribution into the casting volume.

The spheroidal graphite iron samples are undergone to austempering, including heating at 900°C for an hour, after that isothermal retention at 380°C, 0,5 h, 1 h, 2 h, 4 h, 6 h. The austempered ductile iron samples' microstructure is observed by means of an optical metallographic microscope GX41 OLIMPUS. The samples surface is treated with 2 vol.% solution of HNO3 in ethanol. The microstructure of the cast samples before austempering is tested by scanning electron microscopy (SEM). The scanning microscope EVO MA10 „Carl Zeiss" is used. The hardness testing is performed by Vickers method. The austempered ductile iron samples are tested by X-Ray diffraction analysis the retained austenite quantity in the structure to be defined. X-ray powder diffraction patterns for phase identification are recorded in the angle interval 22÷104° (2θ), on a Philips PW 1050 diffractometer, equipped with Cu Kα tube and scintillation detector. Data for cell refinements and quantitative analysis are collected in θ-2θ, step-scan mode in the angle interval from 22 to 104° (2θ), at steps of 0.03° (2θ) and counting time of 3 s/step. Quantitative analysis is carried out by BRASS - Bremen Rietveld Analysis and Structure Suite [9].

For the samples with an upper bainitic structure, obtained for 2 hours of austempering at 380 °C the impact toughness and abrasive wear tests are performed.

The impact strength test is carried out with Charpy hammer. The experimental study of the wear is carried out using method and device for accelerated testing in kinematic scheme "thumb-disc" under friction over a fixed abrasive [6]. The impregnated material Corundum 220 is used for the austempered ductile iron samples. The test data are: nominal contact pressure, $Pa=0,4.10^6$ (Pa); average sliding speed, V=24.5(cm/s); nominal contact surface, Aa=50.24 (mm^2); density, $\rho=7.80.10^3$ (kg/m^3) [7].

III. Experimental results and discussion

The examined cast iron structure consists of ferrite, pearlite and graphite after casting. The nanosized additives change the quantity and size of the graphite phase, as the pearlite and ferrite quantity in the cast iron structure. Nanosized additives in the spheroidal graphite cast iron don't change the graphite shape. They decrease the average diameter of the graphite sphere D_{mid} from 11, 00 to 10, 34 μm. The increase of the quantity of the graphite phase with 35÷94% and the change of the pearlite and ferrite quantity in the cast iron structure are observed in the presence of the nanosized particles [4].

SEM analysis of the fracture of the impact destructed cast iron sample with nanoparticles additives (Fig.1), show the nanoparticles presence in the graphite. These results and that achieved from the quantitative metallographic analysis [4] prove the modified influence of the nanoparticles on the graphite phase in the cast iron samples.

The spheroidal graphite cast irons with and without nanosized additives are subjected to austempering in order to obtain a bainitic structure of the metal base. The austempering regimens include austenitization at 900°C, 1 h and subsequent isothermal retention at 380°C, 0.5 h, 1 h, 2 h, 4 h and 6 h. As a result of this heat treatment, the cast iron is given an upper bainitic structure (Fig. 2).

The binite is an oriented structure consisting of α-phase needles (bainitic ferrite), carbides and unconverted austenite. The α-phase is formed by a martensitic mechanism from the austenitic regions of low carbon content [1, 2]. During cooling from the temperature of the isotherm to room temperature, part of the unconverted austenite undergoes martensitic transformation, and another part remains in the structure as retained austenite A.

The bainitic transformation of the austenite begins with the formation of separate α-phase needles (bainitic ferrite) and develops with the formation of new oriented needles placed close to each other and forming a package of alternating α-phase plates and unconverted carbon enriched austenite $A_{(c)}$ [1,2].

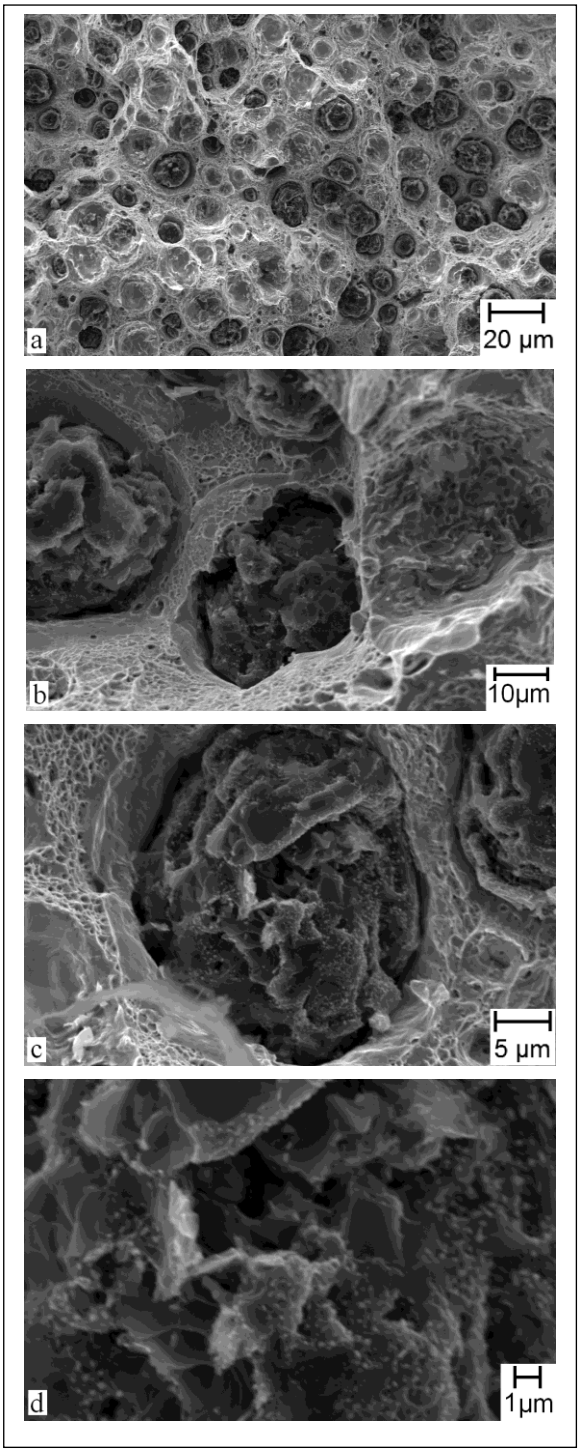


Fig. 1. SEM analysis of the fracture of the impact destructed sample from ductile cast iron with nanoadditives at different magnification (a,b,c,d).

By an optical metallographic analysis, this package appears as a separate needle in the lower bainitic structure. In the upper bainitic structure α-phase and $A_{(c)}$ in the package are perceived as separate phases. The carbide phase is formed as a result of self-release of the α-phase or directly from $A_{(c)}$. Silicon in ductile cast iron (2÷3%) hinders the process of a carbide formation. Isothermal retention up to 2÷4 hours produces a

formation of a bainitic ferrite structure and carbon-enriched unconverted austenite $A_{(c)}$. This structure is characterized by high mechanical properties. At austempering modes over 4÷6hours, it is possible to observe carbides separation directly from $A_{(c)}$ or decomposition of the carbon-enriched austenite $A_{(c)}$ to the ferrite-carbide mixture (α+carbide), which reduces the mechanical properties of the cast iron and it is not actually carried out in practice [1,2].

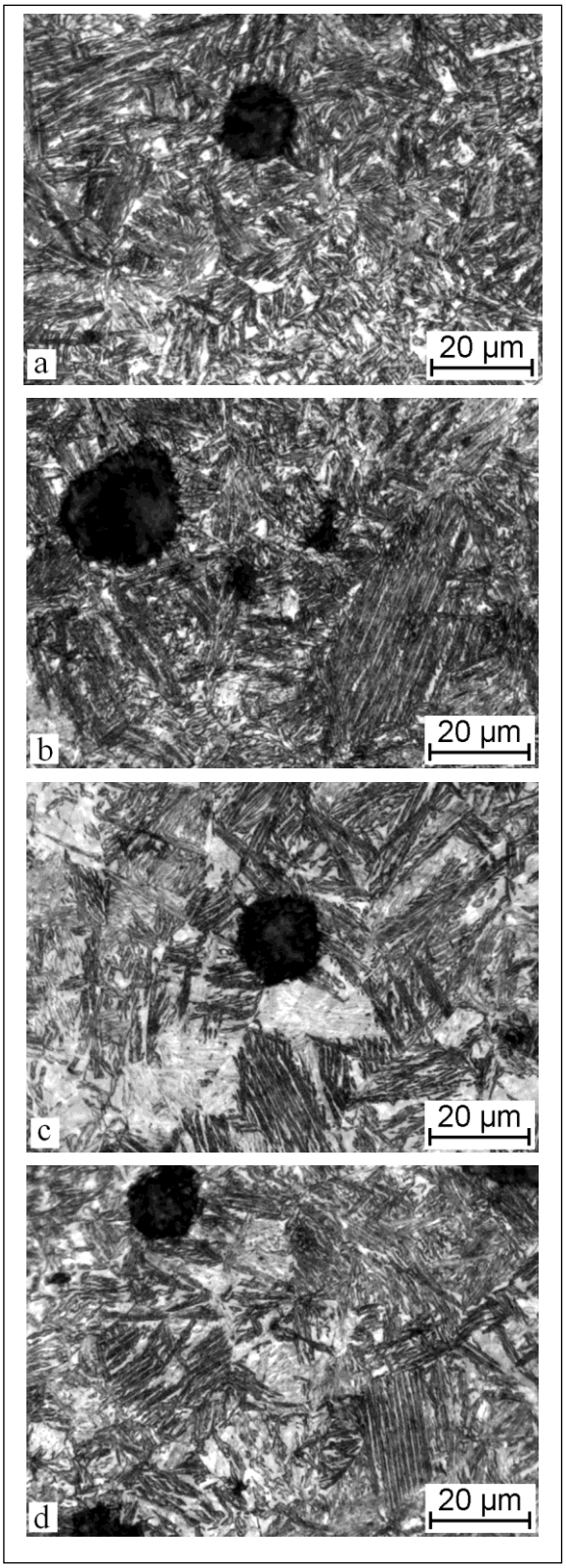


Fig. 2. Upper bainitic structure in ADI samples without nanoadditives (a) and with nanoadditives of (*TiCN+TiN)* (b), *TiN* (c) and *cBN* (b) after isothermal retention for 30 min at 380 ° C.

Nanosized particles accelerate the transformation of austenite to bainite. The amount of the retained austenite in the samples without nanoparticles additives for 2 hours isothermal retention at 380°C is 40.4%. In the samples with nanoaditives, the amount of the retained austenite decreases to 30.2÷27.1% for 2 hours of isothermal retention at 380°C.

With the development of bainitic transformation from 2 to 6 hours the amount of retained austenite in all analyzed samples decreases, which can be explained or with the separation of carbide phase directly from the carbon enriched austenite $A_{(c)}$ or with its decomposition to the ferrite-carbide mixture (α+carbide) (Table 1, Fig. 3). The hardness of the tested ADI is in the range 294÷322 HV10 (Table 1).

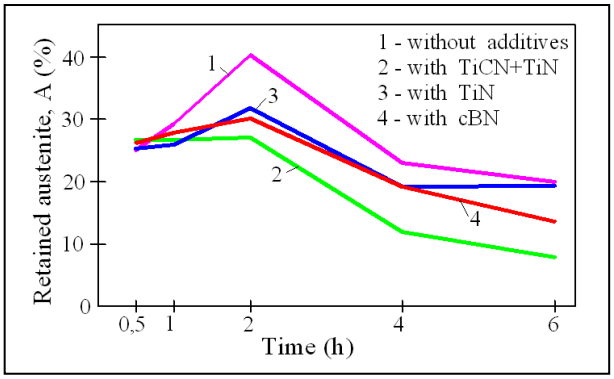


Fig. 3. Dependance of the retained austenite quantity on the time of austempering τ at 380ºC (upper bainitic structure) for spheroidal graphite ductile irons without and with nanoadditives of (*TiCN+TiN*), *TiN* и *cBN*

After final austempering, a complex structure is formed in the tested samples consisting of bainitic ferrite (α-phase), carbides, retained austenite and martensite formed by the unconverted carbon enriched austenite $A_{(c)}$ during cooling from 380°C to an ambient temperature. The properties of the austempered ductile cast irons depend on the type of the phases and the quantitative ratio between them.

The samples after austemepring for 2 hours at 380° C are put to test for impact toughness and abrasive wear (Table 2). The increase of the impact toughness and wear resistance of ADI samples with nanoadditives in comparison to these one without nanoadditives is established (Fig. 4 and 5). The quantity of the retained austenite in the irons samples before and after tribological testing is defined (Table 2).

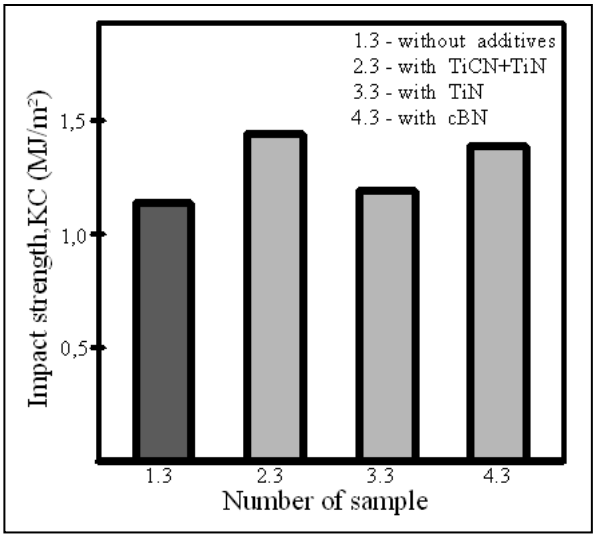


Fig. 4. Impact strength KC of austempered ductile iron samples without nanoadditives (1.3) and with nanoadditives of (*TiCN+TiN)* (2.3), *TiN* (3.3) and *cBN* (4.3) after isothermal retention for 2 h at 380 °C.

TABLE I. RETAINED AUSTENITE QUANTITY *A* AND HARDNESS *HV10* OF AUSTEMPERED DUCTILE IRONS ADI.

№ of sample	Nanosized additives	Time of austempering									
		0,5 h		*1 h*		*2 h*		*4 h*		*6 h*	
		A (%)	*HV10*	*A (%)*	*HV10*	*A (%)*	*HV10*	*A (%)*	*HV10*	*A (%)*	*HV10*
1.1÷1.5	-	25.1	297	29.3	306	40.4	320	23.1	313	20.0	321
2.1÷2.5	*TiCN+TiN*	26.7	314	26.8	304	27.1	322	12.0	309	7.9	318
3.1÷3.5	*TiN*	25.3	294	26.0	292	31.8	312	19.3	317	19.4	311
4.1÷4.5	*cBN*	26.2	298	27.9	305	30.2	308	19.3	314	13.6	301

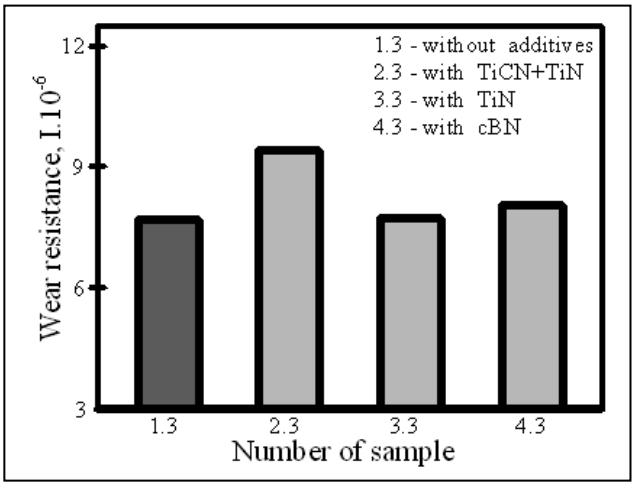


Fig. 5. Wear resistance I of austempered ductile iron samples without nanoadditives (1.3) and with nanoadditives of (*TiCN+TiN)* (2.3), *TiN* (3.3) and *cBN* (4.3) after isothermal retention for 2 h at 380 °C.

TABLE II. NANOADDITIVES, HARDNESS HV10, IMPACT STRENGHT KC, WEAR RESISTANCE I AND RETAINED AUSTENITE QUANTITY A OF AUSTEMPERED DUCTILE IRONS ADI.

№ of sample	Austem-pering mode	Impact strength *KC* (MJ/m^2)	Wear resistance *I*	Retained austenite *A* (%)	
				before wear test	*after wear test*
1.3	900°C, 1 h + 380°C, 2 h	1.137	$7.67.10^6$	40.4	31.3
2.3		1.442	$9.42.10^6$	27.1	11.6
3.3		1.190	$7.72.10^6$	31.8	31.2
4.3		1.387	$8.03.10^6$	30.2	25.3

The austempered ductile iron samples hardness with upper bainitic structure varies from 308 to 322 HV10 (Table1).

The standard measured properties (hardness, etc.) not always are reliable criteria for the steel and cast iron wear resistance. The tribological properties of the metallic materials significantly depend on the structural condition, forming on the contact surface during the friction. When the material is exposed to an intensive plastic deformation the structural transformation goes off in the frictional contact area in the metastable structures (retained austenite, martensite, bainite). They greatly influence on the effective surface strength and respectively on the material tribological properties. During friction the retained austenite transforms partly into strain-induced martensite with the same amount of carbon as in high carbon austenite. This strain-induced martensite is untempered martensite characterized by high hardness and ability for intensive strengthening by wear [7]. The quantity of the retained austenite decreases in all of the tested ADI samples after tribological test. The highest decrease in the quantity of the retained austenite is observed in the sample with nanoparticles of *(TiCN+TiN)* - from 27.1% to 11.6% (Table 2). This sample has highest wear resistance ($I = 9.42.10^6$). The formation of strain-induced martensite from the metastable retained austenite in the frictional area (microtrip-effect) probably is one of the reasons for the wear resistance increase of these materials.

IV. CONCLUSIONS

Nanosized particles in the spheroidal graphite cast irons have a modifying effect on the graphite phase. In the case of austempering of spheroidal graphite cast iron, nanosized additives influence on the structure formation in the temperature range of the bainitic area. They alter the kinetics of the bainitic transformation and accelerate the transformation of austenite into bainite.

Austempered ductile irons (ADI) with nanosized additives possess higher wear resistance (to 23%) and higher impact strength (to 27%) in comparison to the samples without nanoadditives. The observed effect of the nanoadditives on the graphite phase characteristics and on the extent of the transformation of the austenite to bainite explain higher abrasion wear resistance of the tested austempered ductile irons with nanoadditives compared to the same without nanoadditives. Partially transformation of the metastable retained austenite to strain-induced martensite during friction, which is observed in the austempered ductile irons, affects their wear resistance.

REFERENCES

[1] E. Dorasil, *High-strenght Bainitic Nodular Cast Iron*. Academia, Prague,1985.

[2] H.K.D.H. Bhadeshia, *Bainite in Steels*, 2st ed., Inst. of Materials, Cambridge, London, 2001.

[3] Y.Wang, Z. Pan, Z. Wang, X. Sun, and L. Wang, "Sliding wear behavior of Cr–Mo–Cu alloy cast irons with and without nanoadditives", in *Wear*, vol. 271, 2011, pp. 2953– 2962.

[4] J. Kaleicheva, V. Mishev, R. Lasarova, and Z. Karaguiozova, "Study of Spheroidal Graphite Cast Irons with Nanosized Additives", in *Nanoscience & Nanotechnology'15*, eds. E. Balabanova, E. Mileva, Sofia, vol. 15 (1), 2015, pp. 40-43.

[5] J. Li, M. Chen, H. Gao, and Y. Zhao, "Structures and Properties of Cast Irons Reinforced by Trace Addition of Modified SiC Nanopowders", in *Chinese Journal of Chemical Physics*, vol.20 (6), 2007, pp. 625- 631.

[6] J. Kaleicheva, M. Kandeva, V. Mishev, and Z. Karaguiozova, "Wear Behavior of Ductile Cast Irons with Nanoparticle Additives", in *J. Chem. Chem. Eng.*,vol.7 (11), 2013, pp. 1044–1049.

[7] J. Kaleicheva, V. Mishev, G. Avdeev, Z. Karaguiozova, and B. Dineva, (2014). "Influence of Nanoadditives on the Structure and Properties of Austempered Ductile Irons", in *Conference Proceedings: European Conference on Heat Treatment and 21st IFHTSE Congress*, Munich, May 2014, pp. 537-543.

[8] G. Gavrilov, and C. Nicolov, *Electroless Nickel and Composite Coatings.* Tehnika, Sofia, 1985.

[9] J. Birkenstock, R. X. Fischer, and T. Messner, BRASS 2003: The Bremen Rietveld Analysis and Structure Suite. Ber. DMG, Beih. z. Eur. J. Mineral. 15(1), 21, 2003.

Bidirectional travelling surface waves in a convecting fluid

Bidirektionale laufende Oberflächenwellen in einem Konvektionsfluid

Ognyan Y. Kamenov*, Magdalina V. Uzunova†

* Department of Applied Mathematics and Informatics, Technical University of Sofia
P.O. Box 384, 1000 Sofia, Bulgaria, E-mail: okam@abv.bg
† Faculty of Transportation Engineering, Department of Mathematics, UACG
1000 Sofia, Bulgaria, E-mail: magi.uzunova@abv.bg

Abstract — In the present paper have been obtained exact, solitary, generalized periodic and real periodic solutions of one nonlinear and nonintegrable equation, describing the dynamics of bidirectional surface waves in a convecting fluid. This equation was introduced by M. C. Depassier, and its exact solutions are presented below for the first time.

Zusammenfassung — Im vorliegenden Beitrag sind exakte solitäre und verallgemeinerte periodische Lösungen der nichtlinearen Gleichung erhalten, die die Dynamik der bidirektionalen Oberflächenwellen in einem Konvektionsfluid beschreibt. Diese Lösungen werden zum ersten Mal veröffentlicht.

I. Introduction

In 1990 H. Aspe and M. Depassier [1] introduced the evolution equation of unidirectional waves in a convecting fluid

$$u_t + \alpha_0 u u_x + \alpha_3 u_{xxx} + \varepsilon[\alpha_1 (u u_x)_x + \alpha_2 u_{xx} + \alpha_4 u_{xxxx}] = 0,$$

where $u = u(t,x)$ characterizes the surface elevation of the fluid, while $\varepsilon, \alpha_0, \ldots, \alpha_4$ are parameters associated with the dynamic state of the fluid, as $0 < \varepsilon \ll 1$. A similar assumption of unidirectionality is substantiated for waves in plasma, in a fluid flow along an inclined plane, but is not justified for a convecting fluid with long waves developing on its surface. That is why M. Depassier [2], based on the Oberbeck-Boussinesq system [3] and equipped with appropriate boundary conditions, presents the following equation:

$$\begin{aligned} u_{tt} - u_{xx} - \varepsilon^2 (a_0 u_{ttxx} + a_1 u_{xx}^2) + \\ + \varepsilon^3 (R_0 u_{txx} + b_0 u_{txxxx} + b_1 u_{txx}^2) = 0, \end{aligned} \tag{1}$$

where $a_0 = \frac{21+34Pr}{21 Pr^2 Ga}$, $a_1 = \frac{3}{2} + \frac{15}{PrGa}$, $R_0 = \frac{2(Ra-30)}{15 \varepsilon^2 Pr\ Ga}$

$$b_0 = \frac{B}{Pr^4 Ga^2},\ b_1 = \frac{8}{PrGa},$$

$B = \frac{2798}{2079}\ Pr + (GaPr^2 - 1)(\frac{1265}{12096} + \frac{3973}{60480\ Pr})$

where Pr – the Prandtl number, Ga – the Galilean number and Ra – the Rayleigh number.

While there have been no exact localized solutions of equation (1) published so far, a number of authors [4],[5],[6],[7],[8] have obtained various exact solutions of its unidirectional version.

In this work has been applied the mapping and singular deformations method to obtain solitary localized solutions of the bidirectional equation (1), as well as one spatial variation of the Hirota-Satsuma bilinear transformation method [9], described in detail in [8],[10] to obtain a generalized periodic solution.

Equation (1) gives grounds to be considered as third approximation for the small parameter ε of the classical one dimensional wave equation $u_{tt} = u_{xx}$, which is the zero approximation of (1).

II. Mathematical notes

Due to the presence of four terms with mixed partial derivatives in (1), the last one is not invariant with regard to the Galilean transformation, as it does not possess the status of an evolution equation. For convenience in the analysis, let us rescale the variables in equation (1) in the form:

$$t \to \left(\frac{b_1}{a_1}\right) t; \quad x \to \sqrt[4]{\frac{a_1}{b_1 b_0^3}}\, x; \quad u \to \frac{1}{b_1}\sqrt{\frac{a_1}{b_0 b_1}}\, u\,,$$

thus reducing equation (1) to the following

$$\begin{aligned} u_{tt} - \alpha u_{xx} - \varepsilon^2 [\beta u_{ttxx} + (u^2)_{xx}] + \\ + \varepsilon^3 (\gamma u_{txx} + u_{txxxx} + (u^2)_{txx}] = 0\,, \end{aligned} \tag{2}$$

where $\alpha = \sqrt{b_1/a_1 b_0}$; $\beta = a_0 \sqrt{a_1/b_0 b_1}$;
$\gamma = R_0 \sqrt{b_1/a_1 b_0}$. Further on, we will analyze the more compact form of equation (2).

III. Localized solitary solutions

Let us assume that the solution of (2) can be repre-sented in the form of a "traveling" wave $u(t,x) = f(\xi)$, where $\xi = kx + \omega t + \delta$, and $f(\xi)$ is an unknown, sufficiently smooth function, and the parameters of the phase k, ω, δ which in the general case can also be complex. But in order to avoid the trivial solution we will assume that $k \neq 0$ and $\omega \neq 0$. Applying this reduction and after double integration with respect to x, equation (2) is reduced to the following nonlinear ordinary differential equation

$$\begin{aligned} (\omega^2 - \alpha k^2) f - \varepsilon^2 k^2 (\beta \omega^2 f'' + f^2) + \\ + \varepsilon^3 k^2 \omega (\gamma f' + k^2 f'''' + 2 f f') = S, \end{aligned} \tag{3}$$

where $S = \text{const}$, but in the general case $S = S_0 \xi + S_1$ and on the assumption that $S_0 = 0$ then the integration constant is real.

We will assume that the correlation

$$f(\xi) = A\varphi(\xi) + E + B\varphi'(\xi)/[\varphi(\xi) - C]\,,$$

represents the solution of the Riccati equation

$$\varphi'(\xi) = \sqrt{4\varphi^3(\xi) - 12C^2\varphi(\xi) + 8C^3}\ ,$$

in the initial equation (3). We will assume that the unknown at this stage parameters A, B, C, E, S are such that $A^2 + B^2 \neq 0$, in order to avoid the trivial solution, and $C \in \mathbb{R}$ withal $C > 0$. Under the hypothesis for the Riccati equation, we have (See [11],[12])

$$\varphi(\xi) = \wp(\xi, 12C^2, -8C^3),$$

where $\wp(\xi, g_2, g_3)$ is the Weierstrass elliptic function [12] with real invariants $g_2 = 12C^2$, $g_3 = -8C^3$ for $C > 0$. Substituting $f(\xi)$ with its equal in the initial equation (3) we obtain the following algebraic system for the coefficients in front of $\wp$ and $\wp'$.

$$\begin{cases} A(A + 6k^2) = 0 \\ 2B(\beta\omega^2 + A) = \varepsilon\omega(\gamma A + 2AE + 4B^2) \\ B[(\omega^2 - \alpha k^2) - 2\varepsilon^2 k^2(AC + E)] = \\ = 18\varepsilon^3 k^2 \omega A C^2 (A + 6k^2) \\ A(A + 6\beta\omega^2) = 12\varepsilon\omega B(A + k^2) \\ A(\omega^2 - \alpha k^2) = 2\varepsilon^2 k^2[AE + 2B^2 - \varepsilon\omega B(2AC + 2E + \gamma)] \\ S = -2BC\{(\omega^2 - \alpha k^2) + 2\varepsilon^2 k^2[C(3\lambda\omega^2 + 4A) + E]\} + \\ + \varepsilon^2 k^2 \omega C^2\{6A(16k^2 C - \gamma) + 4[A(4AC - 3E) - 6B^2]\} \end{cases} \quad (4)$$

Interesting from the physical point of view are only those real solutions of system (4) for which $A^2 + B^2 \neq 0$, $\omega \neq 0$, $C > 0$ for a wave number $k > 0$, which we accept as a parameter. This system admits two families of such solutions – for $B \neq 0$ and for $B = 0$. If $B \neq 0$, then the system admits the solution:

$$\begin{cases} A = -6k^2; \quad B = \dfrac{3(\beta\omega^2 - k^2)}{5\varepsilon\omega} \\ E = \dfrac{(\beta\omega^2 - k^2)(\beta\omega^2 + 24k^2)}{5\varepsilon^2 k^2 \omega^2} - \dfrac{\gamma}{2} \\ C = \dfrac{[(\beta^2 - 25)\omega^4 + k^2(25\alpha + 23\beta - 25\gamma\varepsilon^2)\omega^2 - 24k^4]}{300\varepsilon^2\omega^2 k^2} \\ S = 72\varepsilon^3 k^2 \omega C^2[k^2(3\gamma + 4E) + 2B^2] - \\ -2BC\{(\omega^2 - \alpha^2 k^2) + 2\varepsilon^2 k^2[3C(\beta\omega^2 - 8k^2) + E]\} \end{cases} \quad (5)$$

and the characteristic equation for th frequency ω is

$$\lambda^3 - k^2\left(\frac{6}{\beta} + \alpha + \beta - \gamma\varepsilon^2\right)\lambda^2 + k^4(\alpha + \gamma\varepsilon^2)\lambda - \frac{6k^2}{\beta} = 0 \quad (6)$$

as $\lambda = \omega^2$. To be correct we need to study for which values of the apriori parameters $\varepsilon, \alpha, \beta, \gamma$ and $k > 0$, the parameter C from the solution (5) is positive, i.e. $C > 0$, under the condition that $0 < \varepsilon \ll 1$; $\alpha > 0$; $\beta > 0$; $\gamma > 0$. If we denote the function

$$\psi(\lambda) = (\beta^2 - 25)\lambda^2 + k^2(23\beta + 25\alpha - 25\gamma\varepsilon^2)\lambda - 24k^4$$

then for $\beta > 5$ this function is convex and has only one positive root λ_2, i.e. $C(\omega) > 0$. If $0 < \beta < 5$ then $\psi(\lambda)$ assumes positive values if $0 < \omega^2 < \lambda_2$ since $\psi(\lambda)$ is concave. The above conditions which hold for every $k > 0, \alpha > 0$ and $\varepsilon \in (0,1)$, provide $C(\omega) > 0$. Finally, if $\beta = 5$, then $\psi(\lambda) > 0$ only if

$$\omega^2 = \lambda > \frac{24k^2}{5(23 + 5\alpha - 5\gamma\varepsilon^2)}$$

The above conditions providing $C > 0$, have a meaning only if the characteristic equation (6) allows at least one positive root. That's really the case, since the signs of the variations of its coefficients is either one or three, no matter what the sign of its second coefficient is and according to the Descartes theorem, this equation has either three positive roots or one triple positive root. Taking into account the correlations for the elliptic functions with invariants $g_2 = 12C^2$, $g_3 = -8C^3$ for $C > 0$, defined by (5)

$$\wp(\xi, g_2, g_3) = 3C\coth^2(\xi\sqrt{3C}) - 2C$$
$$\wp'(\xi, g_2, g_3) = (\wp - C)\sqrt{3C}\coth(\xi\sqrt{3C});$$

then we obtain the following family of solitary-wave solutions

$$\begin{aligned} u(t,x) = -18k^2 C.\coth^2\left(\xi\sqrt{3C}\right) + \\ + \frac{3\sqrt{3C}(\beta\omega^2 - k^2)}{5\varepsilon\omega}\coth\left(\xi\sqrt{3C}\right) + \\ + \left[\frac{(\beta\omega^2 - k^2)(\beta\omega^2 + 24k^2)}{50\varepsilon^2 k^2\omega^2} + \frac{24k^2 C - \gamma}{2}\right] \end{aligned} \quad (7)$$

This family of wave solitary solutions, with free parameters $k > 0$ and δ, is structured by two pulses of the type $\coth^r(\xi\sqrt{3C})$, $r = 1, 2$ with equal phases and spatial displacement (in the square brackets). Such waves are called shock waves. If in the system (5) $B = 0$, then it allows the following solution

$$\begin{aligned} A = -6k^2; \quad B = 0; \quad E = -\frac{\gamma}{2} \\ C = \frac{\beta(\beta^2 - 2) + 25(\alpha - \beta - 1) - 24k^4}{300\varepsilon^2} \\ \varepsilon^2 = \frac{\alpha\beta - 1}{\gamma}; \quad \omega = \pm k\sqrt{\beta} \\ S = \varepsilon^2 k^2(\gamma^2 - 12k^4 g_2). \end{aligned} \quad (8)$$

If in the condition of the hypothesis about the real invariants $g_2 = 12C^2$, $g_3 = -8C^3$ with $C > 0$, we apply the phase modulation $\delta = iK_2/\sqrt{e_1 - e_3}$, where $e_1 = e_2 = C, e_3 = -2C$, then (See [11],[12])

$$\wp(\xi, 12C^2, -8C^3) = C - 3C\,\mathrm{sech}^2(\xi\sqrt{3C}),$$

where $\xi \to kx + \omega t$. The following solitary-wave solution is generated:

$$u(t,x) = 18k^2 C.\mathrm{sech}^2\left(\xi\sqrt{3C}\right) - (6k^2 C + \gamma/2) \quad (9)$$

and all the parameters in this solution are defined by the equality (8). The above solitary solution cannot be degenerated into a soliton one since for $k^2 = \gamma/12C$. The requirement $C > 0$ according to (8) means that

$$0 < k < \{[\beta(\beta^2 - 2) + 25(\alpha - \beta - 1)]/24\}^{1/4},$$

which is an additional restriction for the wave number.

IV. Periodic solutions

Let us apply the Hirota-Satsuma transformation [9] to the solution of equation (2)

$$u(t,x) = c + \mu(\ln\zeta)_{xx}, \quad (10)$$

where c, μ are unknown parameters at this stage, and let us assume that $\mu \neq 0$ in order to avoid the trivial solution, while $\zeta = \zeta(t,x)$ belongs to the class of functions $W_F^{2,6}(\Omega)$ containing all the real and complex functions defined in the stripe $\Omega = \{(t,x) \in \mathbb{R}^2, 0 < t < \infty, -\infty < x < \infty\}$ which have continuous derivatives with respect to t and x up to second and sixth order, and this space possesses a strong topology. After substituting $u(t,x)$ from (10) in (2), single integration with respect to x and representing the logarithmic derivatives by the Hirota bilinear operator

$$D_t^n D_x^m \varphi(t,x).\psi(t,x) =$$
$$= (\partial/\partial t - \partial/\partial t')^n(\partial/\partial x - \partial/\partial x')^m \varphi(t,x).\psi(t',x')|_{t=t', x=x'}$$

(See [11, 12]) we will obtain a bilinear reduction of (2)

$$\frac{1}{2\zeta^2}[D_t^2-\alpha D_x^2-\varepsilon^2\beta D_t^2D_x^2-8G]\zeta.\zeta+$$
$$+\frac{\varepsilon^2}{\zeta^2}\left(\frac{D_x^2\zeta.\zeta}{2\zeta^2}\right)[6\beta D_t^2\zeta.\zeta-\mu D_x^2\zeta.\zeta-2c\zeta^2]+$$
$$+\varepsilon^3\frac{\partial}{\partial t}\left\{\frac{3}{\zeta^2}[D_x^4+(\gamma+2c)D_x^2+8k^2H]\zeta.\zeta\right\}+$$
$$+(\mu-6)\left(\frac{D_x^2\zeta.\zeta}{2\zeta^2}\right)^2=0$$

where $8G$ is a resultant constant of integration, while $8k^2H$ is the so-called "differential" constant whose role will be specified later.

If in the last bilinear equation, we choose $\mu=6$, then it is represented as a conjunction of three residual equations

$$(D_t^2-\alpha D_x^2-\varepsilon^2\beta D_t^2D_x^2-8G)\zeta.\zeta=0 \quad (11)$$
$$[D_x^4+(\gamma+2c)D_x^2+8k^2H]\zeta.\zeta=0 \quad (12)$$
$$2c\zeta^2=3\beta D_t^2\zeta.\zeta-3D_x^2\zeta.\zeta \quad (13)$$

The presence of three residual equations, one of which (13) not being polynomial-bilinear with respect to the operators D_t and D_x, clearly demonstrates the non-integrable nature of the initial equation (2). If there exists a sufficiently smooth function $\zeta(t,x)$ of the class $W_F(\Omega)$, satisfying all the three residual equations, then the function $u(t,x)$, formed by the ratio (10), would be an exact localized solution of equation (2). For this purpose, we will assume that the Jacobi θ-function (See [12]) is a probable solution

$$\zeta(t,x)=\theta_3(\xi,q)=\sum_{n=-\infty}^{\infty}q^{n^2}e^{2in\xi},$$

where $\xi=kx+\omega_0t+\delta$ and $q=e^{i\pi\tau}$, $Im\tau>0$ i.e. $0<|q|<1$. This function is a well-defined biperiodic function for $|q|\in(0,1)$ with real period π and quasi period 2. In parallel, we will check the residual equations (11) and (12) as regards their solvability with the function $\theta_3(\xi,q)$ as they have a similar polynomial-bilinear structure. If we substitute $\theta_3(\xi,q)$ in equations (11) and (12) and use the properties of the bilinear operators $D_t^m\zeta.\zeta$, $D_x^n\zeta.\zeta$, $D_t^nD_x^m\zeta.\zeta$ (See [13]), we will obtain accordingly the infinite systems

$$\sum_{m=-\infty}^{\infty}F_1(m)e^{2im\xi}=0,\ \sum_{m=-\infty}^{\infty}F_2(m)e^{2im\xi}=0,$$

where

$$F_1(m)=\sum_{m=-\infty}^{\infty}[-4\omega_0^2(2n-m)^2+$$
$$+4ak^2(2n-m)^2-16\varepsilon^2\beta k^2(2n-m)^4-8G] \text{ for (11)},$$
$$F_2(m)=\sum_{m=-\infty}^{\infty}[16k^4(2n-m)^4-$$
$$-4k^2(\gamma+2c)(2n-m)^2+8k^2H]q^{n^2+(n-1)^2} \text{ for (12)}.$$

Let us apply the index parity principle [14] to the infinite systems above. This principle means that if in the systems $F_j(m)=0$, $m=0,\pm1,\pm2,\dots$, $j=1,2$ we rearrange the terms $n\to n+1$, then we will easily obtain the chains

$$F_j(m)=F_j(m-2)q^{2(m-1)}=\cdots=$$
$$=\begin{cases}F_j(0)q^{\frac{m^2}{2}}, & \text{if } m \text{ is even};\ j=1,2\\ F_j(1)q^{(m^2-1)/2}, & \text{if } m \text{ is odd};\ j=1,2\end{cases}$$

which means that if we group the terms in the infinite sums $F_j(m)=0$ into even and odd addends, we will obtain the two compact equalities:

$$F_j(0)\theta_3(2\xi,q^2)+F_j(1)q^{-1/2}\theta_2(2\xi,q^2)=0,\ j=1,2$$

where $\theta_2(z,q)$ is the second Jacobi θ-func-tion: $\theta_2(z,q)=\sum_{n=-\infty}^{\infty}q^{(n-1/2)^2}e^{i\xi(2n-1)}$, i.e. the infinite systems are reduced to the two pairs of equations

$$F_1(0)=0;\ F_1(1)=0 \text{ and } F_2(0)=0;\ F_2(1)=0.$$

If we apply the transformation formulas to these two equations, they will transform into the following two linear non-homogeneous algebraic systems

$$q[\theta_3'+8\varepsilon^2\beta k^2(\theta_3'+q\theta_3'')]\omega_0^2+\theta_3G=\alpha k^2q\theta_3';$$
$$q[\theta_2+8\varepsilon^2\beta k^2(\theta_2'+q\theta_2'')]\omega_0^2+\theta_2G=\alpha k^2q\theta_2';$$
$$8q(\theta_3'+q\theta_3'')]k^2+\theta_3H=q(\gamma+2c)\theta_3';$$
$$8q(\theta_2'+q\theta_2'')]k^2+\theta_2H=q(\gamma+2c)\theta_2',$$

as the first system is with relation to ω_0^2 and G, and refers to (11), while the second one is with relation to k^2 and H, and refers to (12). For convenience we have used the denotations $\theta_j=\theta_j(0,q^2)$ and $\theta_j'=d\theta_j/dq$, $j=2,3$. Both algebraic systems are compatible and determined, which is easily verified by the linear independence of θ_2 and θ_3. The only solutions of these systems are

$$\omega_0(q)=\pm\sqrt{\frac{\alpha k^2}{1+8\beta\varepsilon^2k^2W_0(\theta_2,\theta_3)}} \quad (14)$$
$$G(q)=\frac{8\alpha\beta\varepsilon^2qW(\theta_2',\theta_3')}{W(\theta_2,\theta_3)[1+8\beta\varepsilon^2k^2W_0(\theta_2,\theta_3)]}$$
$$k^2=\frac{\gamma+2c}{8W_0(\theta_2,\theta_3)};\ H(q)=\frac{q(\gamma+2c)W(\theta_2',\theta_3')}{W(\theta_2,\theta_3)W_0(\theta_2,\theta_3)} \quad (15)$$

where $W(\theta_2,\theta_3)=\theta_2\theta_3'-\theta_2'\theta_3$ is the Wronskian of the functions θ_2 and θ_3, $W'(\theta_2,\theta_3)=\theta_2\theta_3''-\theta_2''\theta_3$, $W_0(\theta_2,\theta_3)=1+qW'(\theta_2,\theta_3)/W(\theta_2,\theta_3)$.

At this stage of the solution, the active parameters are c,q,δ, which could be complex in the general case. It is not difficult to see from (14) that the integration constant $G(q)$ is different from zero for $|q|\in(0,1), k\neq0$ and for every admissible value of the apriori para-meters $\alpha,\beta,\gamma,\varepsilon$. This condition is necessary for the existence in general of a periodic solution to (2).

The suitability of the index parity principle is inapplicable in trying to satisfy the residual equation (13). Hence, we will use a spatial version of the bilinear-transformation method (See [8],[10]), representing the spatial displacement in the formal series

$$c=6(k^2-\beta\omega_0^2)\sum_{m=-\infty}^{\infty}c_m(q) \quad (16)$$

with unknown terms $c_m(q)$, while the wave number k and the phase frequency ω_0^2 are defined by means of (15) and (14), respectively. If we substitute (16) in equation (13) and apply the Cauchy formula

$$\sum_{\nu=-\infty}^{\infty}a_\nu\sum_{s=-\infty}^{\infty}b_s=\sum_{m,n=-\infty}^{\infty}a_mb_{n-m}=\sum_{m,n=-\infty}^{\infty}a_{n-m}b_m$$

we will obtain the following infinite system

$$c_m\sum_{n=-\infty}^{\infty}q^{(n-m)^2+(n-2m)^2}e^{2im\xi}=$$
$$=\sum_{n=-\infty}^{\infty}(2n-m)^2q^{n^2+(n-m)^2}e^{2im\xi}$$

(we suppose that $k^2\neq\beta\omega_0^2$) from which we can determine

$$c_m(q) = \frac{\sum_{n=-\infty}^{\infty}(2n-m)^2 q^{n^2+(n-m)^2}}{\sum_{n=-\infty}^{\infty} q^{(n-m)^2+(n-2m)^2}}, m = 0, \pm 1, \ldots \quad (17)$$

It is not difficult to show that the series $\sum_{n=-\infty}^{\infty} c_m(q)$, with terms defined by (17), is absolutely convergent if $0 < |q| < 1$. Indeed, if we rearrange the terms of this series applying the reduction $n \to n + 2m$, then we will obtain the representation

$$\sum_{m=-\infty}^{\infty} c_m(q) = [-R_0(z,q)/9 - 2imR_1(z,q) + \\ +9m^2 R_2(z,q)]q^{4m^2}, \text{where}$$

$$z = m\pi\tau; R_0(z,q) = \frac{\ddot{\theta}_3(3z,q^2)}{\theta_3(z,q^2)}; R_1(z,q) = \frac{\dot{\theta}_3(3z,q^2)}{\theta_3(z,q^2)}; \\ R_2(z,q) = \theta_3(3z,q^2)/\theta_3(z,q^2).$$

From the last equality follows the estimate

$$\sum_{m=-\infty}^{\infty} |c_m(q)| \le \frac{1}{9}\sum_{m=-\infty}^{\infty} |R_0||q|^{4m^2} + \\ +2\sum_{m=-\infty}^{\infty} |m||R_1||q|^{4m^2} + 9\sum_{m=-\infty}^{\infty} m^2|R_2||q|^{4m^2}$$

i.e. the series of the absolute values is bounded by the superposition of three absolutely convergent series, since $|R_j(z,q)| \le P_j$, $j = 0,1,2$; $P_j = \text{const}$. Furthermore eve-ry series $\sum_{m=-\infty}^{\infty} |m|^{\nu}|q|^{4m^2}$, $\nu \in \mathbb{N}$, $0 < |q| < 1$ is convergent according to the D'Alembert criterion.

We can already make the conclusion that the function $\zeta(t,x) = \theta_3(\xi,q)$, with parameters calculated from (14) and (15) and c, defined by (18), defines an exact localized complex solution of equation (2) in the for

$$u(t,x) = 6(k^2 - \beta\omega_0^2)\sum_{m=-\infty}^{\infty} c_m(q) + 6k^2\frac{\partial}{\partial\xi}\left[\frac{\dot{\theta}_3(\xi,q)}{\theta_3(\xi,q)}\right] \quad (18)$$

V. Real periodic solutions

The double poles of the complex meromorphic solution (19) being in the lattice

$$\xi_{mn} = (m + 1/2)\pi + \tau(n + 1/2)\pi, \quad m,n \in \mathbb{Z}$$

can be avoided if we limit the variation of the phase variable ξ. Within the condition of the hypothesis for q to be real, i.e. $\tau = ip, p > 0$, whereat $q = e^{i\pi\tau} = e^{-\pi p} \in (0,1)$, this restriction could be $|Im(\xi)| < \pi p$, i.e. if ξ varies in the horizontal stripe $-\pi p < Im(\xi) < \pi p$. This is also an analyticity condition of the localized exact solution (19).

Practical applicability of this solution within the hypothesis $q = e^{-\pi p}$, $p > 0$ can be achieved for two choices of the wave number k from (15). If $k \in \mathbb{R}$, we can assume that $k > 0$. Taking into account the Fourier representation (See [12]) we obtain the following real restriction of (18)

$$u(t,x) = 6(k^2 - \beta\omega_0^2)\sum_{m=-\infty}^{\infty}[c_m(q) + \\ + \frac{2k^2}{k^2 - \beta\omega_0^2}(-1)^m m\,\text{cosech}(\pi mp).\cos(2m\xi)], \quad (19)$$

which describes a family of sinusoidal harmonics with different amplitudes and individual spatial displacements, developing in zones with weak nonlinearity ($p \to 0$, i.e. $q \to 1$).

If k is an imaginary number (with positive imaginary part), then for suitable choice of the phase displace-ment $\delta \to i\delta$, we can achieve a completely imaginary phase $\xi \to i\xi + i\pi p$ and then, using the quasi-periodic modulation for $\theta_3(z + i\pi p/2, q)$, we will obtain from (18) the periodic solitary-wave solution of (2)

$$u(t,x) = 6(k^2 - \beta\omega_0^2)\sum_{m=-\infty}^{\infty}[c_m(q) + \\ + \frac{k^2}{k^2 - \beta\omega_0^2}\text{sech}^2(\xi - m\pi p)] \quad (20)$$

VI. Conclusions

In this periodic real solution is carried out a complex superposition of sech^2-profiles, each of which having its own spatial displacement $c_m(q)$. In this case we cannot talk about the existence of one solitary wave which would be an envelope of the solitary wave profiles (21), because the phase velocity ω/k of the solitary wave (7), cannot become equal to the phase velocity ω_0/k of the periodic wave (21), which is valid in the zones of strong nonlinearity ($p \to \infty$, i.e.$q \to 0$).

References

[1] H.Aspe, M.C.Depassier , "Evolution equation of surface waves in a convecting fluid", *Phys. Rev. A* 41, 1990, 3125-28.

[2] M.C.Depassier , "Evolution equation for bidirectional surface waves in a convecting fluid", *Phys. Fluids* 18, 2006, 107102.

[3] J.Boussinesq , *Théorie analytique de la chaleur t.2*. Paris: Gauthier-Villars.

[4] S.Lou, G.Huang G, H.Ruan, "Exact solitary waves in a convecting fluid", *J. Phys. A: Math. Gen.* 24, L587, 1991.

[5] A.V.Porubov, "Exact travelling wave solutions of nonlinear evolution equation of surface waves in a convecting fluid", J. *Phys. A: Math. Gen.* 26, 1993, 797.

[6] E.J.Parkes, B.R.Duffy, P.C.Abbott, "The Jacobi elliptic-function method for finding periodic – wave solutions to nonlinear evolution equations", *Phys. Lett. A* 295, 2002, 280.

[7] N.A.Kudryashov, "Nonlinear differential equations with exact solutions expressed via the Weierstrass function", *Intern. Science project no. 1379*-2 1, 2003.

[8] O.Y.Kamenov, "Periodic solutions of the non-integrable convective fluid equation", *J. Math. Phys.* 53, 2012, 063705.

[9] R.Hirota, J.Satsuma, "A variety of nonlinear network equations generated from the Bäcklund transformation for the Toda lattice", *Prog. Theoret. Phys. Suppl*. 59, 1976, 64.

[10] O.Y.Kamenov, "Exact periodic solutions of the sixth-order generalized Boussinesq equation ", *J. Phys. A: Math. Theor*. 42, 2009, 375501.

[11] G.Bateman, A.Erdelyi, *Higher transcendental functions* , 1955, New York: Mc Graw-Hill Company Inc.

[12] D.F.Lawden, *Elliptic Functions and Application*, 1989, Springer-Verlag New York

[13] R.K.Bullough, P.J.Caudrey, *Solitons*, 1980, Springer-Verlag Berlin

[14] A.Nakamura, "A direct method of calculating periodic wave solutions to nonlinear evolution equations. Exact two-periodic wave solutions", *J. Phys. Soc. Japan* 47, 1979, 1701.

[15] A.Parker, "On exact solutions of the regularized long – wave equation: A direct approach to partially integrable equations. II. Periodic solutions", *J. Math. Phys.* 36, 1995, 3506.

Solitary and Cnoidal Waves in the Equation of a Convecting Fluid

Solitäre und knoidale Wellen in der Gleichung des Konvektionsfluids

Magdalina Uzunova

Universität für Architektur, Bauwesen und Geodäsie
Sofia, Bulgarien, E-Mail: magi.uzunova@abv.bg

Abstract — In the present work several different models of exact solitary families of traveling waves of the evolutionary equation for a convective fluid are obtained. This is achieved on the basis of the generalized mapping and singular deformations method. Families of the difficult to find rational and cnoidal waves were also obtained in the general case of the studied equation.

Zusammenfassung — Im präsentierten Beitrag sind verschiedene Modelle von exakten solitären Scharen von „laufenden" Wellen für die Evolutionsgleichung des Konvektionsfluids erhalten. Die Ergebnisse werden auf der Basis der verallgemeinerten Methode der Abbildung und der singulären Verformungen erreicht. Im allgemeinen Fall der erforschten Gleichung sind auch schwer auffindbare Scharen von rationalen und knoidalen Wellen erhalten.

I. Einleitung

Die Evolutionsgleichung des Konvektionsfluids

$$u_t + \alpha_0 u u_x + \alpha_1 u_{xx} + \alpha_2 u_{xxx} + \\ +\alpha_3 u_{xxxx} + \alpha_4 (u u_x)_x = 0, u = u(t,x) \quad (1)$$

eingeführt von Aspe und Depassier [1], beschreibt die Dynamik der langen Wellen in einem Konvektionsfluid auf einer schiefen Ebene, der Driftwellen in Plasma oder der Konvektionsprozesse in Fluiden mit freier Oberfläche. Diese Gleichung enthält in sich selbst verschiedene Modellgleichungen in Abhängigkeit von den Parametern $\alpha_i, i = 0,1,2,3,4$. Für $\alpha_4 = 0$ stimmt die Gleichung (1) mit der Evolutionsgleichung von Kuramoto-Sivashinski [2], bekannt als verallgemeinerte Evolutionsgleichung der Wellendynamik, überein. Für $\alpha_3 = \alpha_4 = 0$ ergibt sich die klassische Korteweg-de Vries-Gleichung [3]. Die letzte Gleichung ist von mehreren Autoren in den 60-er Jahren des vorigen Jahrhunderts erforscht. Die Gleichung (1) beschreibt die Verbreitung von unidirektionalen langen Wellen. Sie hat eine komplizierte mathematische Struktur. Die Gleichung ist von vierter Ordnung mit zweitem Singularitätsgrad. Sie hat zwei nichtlineare Glieder und ist nicht integrierbar. Durch Anwendung der direkten Version der Methode der Abbildung haben zuerst Lou, Huang, Ruan [4] eine exakte solitäre Wellenlösung von der Art „laufender" Stoßwelle erhalten. Später hat Porubov [5] durch passende Phasenveränderungen einer elliptischen Lösung der Gleichung (1), dargestellt mit Hilfe der Basisfunktion von Weierstraß [6], [7], eine exakte solitäre und eine knoidale Lösung erhalten. Diese Lösungen gelten nur unter den Bedingungen $\alpha_2\alpha_4 = \alpha_0\alpha_3$ oder $\alpha_2\alpha_4 = 6\alpha_0\alpha_3$. Das heißt, dass diese Lösungen bedingt sind.

In dieser Forschung wird ein breiteres Spektrum von solitären, rationalen und knoidalen Lösungen der Gleichung (1) dargestellt, erhalten mit Hilfe einer Verallgemeinerung der Methode der Abbildung und der singulären Verformungen (Methode von Weiss-Tabor-Carnevale [8]). Diese Methode hat ermöglicht, nicht nur exakte analytische Lösungen zu lokalisieren, sondern auch neue Zweiparameterschare von bedingungslosen Lösungen der betrachteten Evolutionsgleichung zu entwickeln. Zu den exakten lokalisierten Lösungen ohne zusätzliche Bedingungen zwischen den Koeffizienten zählt auch die durch eine räumliche Version der bilinearen Transformationsmethode von Hirota [10]-Matsuno [11] erhaltene von Kamenov [9] periodische Lösung. Sie gilt nicht nur in Gebieten der schwachen, sondern auch in Gebieten der starken Nicht-linearität. In beiden Fällen ist das Prinzip der nichtlinearen Überlagerung von Toda [12] gültig.

II. Solitäre „laufende" Wellen

Die Koeffizienten $\alpha_i, i = 0,1,2,3,4$ sind mit den dynamischen Eigenschaften des Konvektionsfluids verbunden.

$$\begin{cases} \alpha_0 = \dfrac{3(10 + Pr.Ga)}{2Pr.Ga}\ , \alpha_4 = \dfrac{\varepsilon 8}{\sqrt{Ga}} \\ \alpha_1 = \dfrac{\varepsilon R_0 Pr}{15}\ , \alpha_2 = Pr\sqrt{Ga}\left(\dfrac{1}{6} + \dfrac{17}{21}Pr\right) \\ \alpha_3 = \dfrac{Pr}{2079}(682(Pr)^2 Ga + 717) \end{cases}$$

In diesen Formeln ist Pr die Prandtl-Zahl. Weiter ist Ga die Galilei-Zahl. ε ist ein kleiner Parameter $(0 < \varepsilon \ll 1)$, so dass die Rayleigh-Zahl (Ra) mit $Ra = Pr.Gr$ den kritischen Wert $\varepsilon^2 R_0$ übersteigt. R_0 ist für das Fluid ein kritischer Wert der Rayleigh-Zahl. Gr ist die Grashof-Zahl. Es wird eine nichttriviale lokalisierte glatte Lösung der Gestalt „laufender" Welle der Gleichung (1)

$$u(t,x) = \zeta(\xi) \quad (2)$$

mit $\xi = k(x + \omega t)$ im Gebiet
$\Omega = \{(t,x) \in \mathbb{R}^2, 0 < t < \infty, -\infty < x < \infty\}$ gesucht.
$\zeta(\xi)$ ist eine reelle Funktion von der Klasse $C^4(\Omega)$. Zum Vermeiden der Nulllösung sind die unbekannten reellen Parameter ohne Einschränkung der Gesamtheit als
$k > 0, \omega \neq 0$ vorausgesetzt. Nach Einsetzen von (2) in (1) und einmaligem Integrieren nach der Phasenveränderlichen ξ ergibt sich folgende nichtlineare Differentialgleichung

$$\omega\zeta + \frac{\alpha_0\zeta^2}{2} + \alpha_1 k\zeta' + \alpha_2 k^2\zeta'' + \alpha_3 k^3\zeta''' + \\ +\alpha_4 k\zeta\zeta' = E \tag{3}$$

Die kumulative Integrationskonstante E kann im Allgemeinen von t abhängen.

Die Methode der Abbildung und der singulären Verformungen beruht auf folgender Annahme. Ein Verhältnis $\zeta = L(\varphi)$ bildet die Lösung der Gleichung $\mu(\varphi) = 0$ auf die Anfangsgleichung $F(\zeta) = 0$ ab. Die Gleichung (3) hat den Singularitätsgrad 2. Deshalb wird angenommen, dass das Verhältnis

$$\zeta = A\varphi^2 + B\varphi + C \tag{4}$$

mit $\varphi = \varphi(\xi)$ eine beliebige nichttriviale Lösung der Riccati-Differentialgleichung

$$\varphi' = c_0\varphi^2 + c_1\varphi + c_2 \tag{5}$$

auf eine lokalisierte Lösung der Anfangsgleichung (3) abbildet. In (4) und (5) sind A, B, C, c_0, c_1, c_2 unbekannte Parameter. Zum Ausschließen der trivialen Lösung wird $A \neq 0, c_0 \neq 0$ vorausgesetzt. Nach Einsetzen von (4) und (5) in (3) ergibt sich das folgende nichtlineare Gleichungssystem

$$\begin{gathered} Ac_0k(\alpha_4 A + 12\alpha_3 k^2 c_0^2) = 0 \\ \frac{1}{2}\alpha_0 A^2 + 6\alpha_2 Ak^2 c_0^2 + 6\alpha_3 k^3(9Ac_0^2c_1 + Bc_0^3) + \\ +\alpha_4 kA(2Ac_1 + 3Bc_0) = 0 \\ \alpha_0 AB + 2\alpha_1 kAc_0 + 2\alpha_2 k^2 c_0(5Ac_1 + Bc_0) + \\ +2\alpha_3 k^3 c_0(19Ac_1^2 + 6Bc_0c_1 + 20Ac_0c_2) + \\ +2\alpha_4 k(3ABc_1 + 2ACc_0 + B^2c_0 + 2A^2c_2) = 0 \\ \omega A + \frac{1}{2}\alpha_0(B^2 + 2AC) + \alpha_1 k(2Ac_1 + Bc_0) + \\ +\alpha_2 k^2(4Ac_1^2 + 3Bc_0c_1 + 8Ac_0c_2) + \\ +\alpha_3 k^3\big(c_1^2(8Ac_1 + 7Bc_0) + c_0c_2(8Bc_0 + 52Ac_1)\big) + \\ +\alpha_4 k\big(A(2Cc_1 + 3Bc_2) + B(Bc_1 + Cc_0)\big) = 0 \\ \omega B + \alpha_0 BC + \alpha_1 k(2Ac_2 + Bc_1) + \\ +\alpha_2 k^2(c_1(6Ac_2 + Bc_1) + 2Bc_0c_1) + \\ +\alpha_3 k^3\big(Bc_1^3 + c_2(14Ac_1^2 + 16Ac_0c_2 + 8Bc_0c_1)\big) + \\ +\alpha_4 k\big(BCc_1 + c_2(2AC + B^2)\big) = 0 \\ E = \omega C + \frac{1}{2}\alpha_0 C^2 + \alpha_1 kBc_2 + \alpha_2 k^2 c_2(2Ac_2 + Bc_1) + \\ +\alpha_3 k^3 c_2\big(Bc_1^2 + 2c_2(3Ac_1 + Bc_0)\big) + \alpha_4 kBCc_2 \end{gathered} \tag{6}$$

Von der physikalischen Hinsicht gibt es folgende interessante Lösung:

$$\begin{gathered} A = -12\frac{\alpha_3}{\alpha_4}k^2c_0^2; B = 0 \\ C = \frac{(\alpha_2\alpha_4 - \alpha_0\alpha_3)(4\alpha_2\alpha_4 + 9\alpha_0\alpha_3) - 15\alpha_1\alpha_3\alpha_4}{25\alpha_3\alpha_4^2} \\ c_0 \in \mathbb{R}\ ist\ freier\ Parameter, c_1 = \frac{\alpha_0\alpha_3 - \alpha_2\alpha_4}{5\alpha_3\alpha_4 k} \\ \omega = \frac{1}{5\alpha_3}(2\alpha_2\alpha_4 - 7\alpha_0\alpha_3)C - \\ -\frac{1}{5\alpha_3}(5\alpha_3k^2c_1^2 - \alpha_2kc_1 - \alpha_1)(13\alpha_0\alpha_3 - 3\alpha_2\alpha_4) - \\ -\frac{2(\alpha_0\alpha_3 - \alpha_2\alpha_4)}{25\alpha_3^2\alpha_4^2}\big(5\alpha_1 + 2\alpha_2(\alpha_0\alpha_3 - \alpha_2\alpha_4)\big) \\ c_2 = \frac{5\alpha_3k^2c_1^2 - \alpha_2kc_1 - \alpha_1}{20k^2c_0\alpha_3} \end{gathered} \tag{7}$$

$$E = \omega C + \frac{1}{2}\alpha_0 C^2 - \\ -\frac{3}{50\alpha_4}(\alpha_2 + 3\alpha_3 kc_1)(5\alpha_3k^2c_1^2 - \alpha_2kc_1 - \alpha_1)^2$$

$$k > 0\ ist\ freier\ Parameter$$

Die erhaltene Lösung (7) ermöglicht die Bestimmung eines breiteren Spektrums von exakten lokalisierten Lösungen der Gleichung (1) unter der Annahme, dass die Koeffizienten $\alpha_i, i = 0,1,2,3,4$ positiv sind. Dann gibt es folgende Fälle für die Lösungen von (1):

Wenn $c_1^2 < 4c_0c_2$, d.h. $0 < k < \frac{5\alpha_1\alpha_3\alpha_4}{\alpha_2(\alpha_2\alpha_4 - \alpha_0\alpha_3)}$, dann gilt

$$u(t,x) = 3\frac{\alpha_3}{\alpha_4}k^2(c_1^2 - 4c_0c_2)tan^2\left(\frac{\xi}{2}\sqrt{4c_0c_2 - c_1^2}\right) + \\ +6k^2c_1^2\frac{\alpha_3}{\alpha_4}\sqrt{4c_0c_2 - c_1^2}\,tan\left(\frac{\xi}{2}\sqrt{4c_0c_2 - c_1^2}\right) + \\ +C - 3k^2c_1^2\frac{\alpha_3}{\alpha_4} \tag{8}$$

Wenn $c_1^2 > 4c_0c_2$ und

$$-\frac{1}{2c_0}\left(-c_1 - \sqrt{c_1^2 - 4c_0c_2}\right) < \varphi < \frac{1}{2c_0}\left(-c_1 + \sqrt{c_1^2 - 4c_0c_2}\right)$$

d.h. $k > max\left(0, \frac{5\alpha_1\alpha_3\alpha_4}{\alpha_2(\alpha_2\alpha_4 - \alpha_0\alpha_3)}\right)$, dann gilt

$$u(t,x) = 3\frac{\alpha_3}{\alpha_4}k^2(4c_0c_2 - c_1^2)tanh^2\left(\frac{\xi}{2}\sqrt{c_1^2 - 4c_0c_2}\right) - \\ -6k^2c_1^2\frac{\alpha_3}{\alpha_4}\sqrt{c_1^2 - 4c_0c_2}\,tanh\left(\frac{\xi}{2}\sqrt{c_1^2 - 4c_0c_2}\right) + \\ +C - 3k^2c_1^2\frac{\alpha_3}{\alpha_4} \tag{9}$$

Wenn $c_1^2 > 4c_0c_2$ und

$$\varphi \in \left(-\infty; -\frac{-c_1 - \sqrt{c_1^2 - 4c_0c_2}}{2c_0}\right) \cup \left(\frac{-c_1 + \sqrt{c_1^2 - 4c_0c_2}}{2c_0}; +\infty\right)$$

d.h. $k > max\left(0, \frac{5\alpha_1\alpha_3\alpha_4}{\alpha_2(\alpha_2\alpha_4 - \alpha_0\alpha_3)}\right)$, dann gilt

$$u(t,x) = 3\frac{\alpha_3}{\alpha_4}k^2(4c_0c_2 - c_1^2)cotanh^2\left(\frac{\xi}{2}\sqrt{c_1^2 - 4c_0c_2}\right) - \\ -6k^2c_1^2\frac{\alpha_3}{\alpha_4}\sqrt{c_1^2 - 4c_0c_2}\,cotanh\left(\frac{\xi}{2}\sqrt{c_1^2 - 4c_0c_2}\right) + \\ +C - 3k^2c_1^2\frac{\alpha_3}{\alpha_4} \tag{10}$$

Wenn $c_1^2 = 4c_0c_2$, d.h. $k = \frac{5\alpha_1\alpha_3\alpha_4}{\alpha_2(\alpha_2\alpha_4 - \alpha_0\alpha_3)}$, dann gilt

$$u(t,x) = \frac{\alpha_2\alpha_4 - \alpha_0\alpha_3}{10\alpha_3\alpha_4kc_0} - \frac{1}{c_0\xi} \tag{11}$$

Alle Parameter $(c_0, c_1, c_2, k, \omega, C)$ in den Formeln (8), (9), (10), (11) sind in (7) bestimmt. Die Lösungen (8), (9), (10) sind solitäre Stoßwellen, die Lösung (11) ist rational. Alle solitären Lösungen beschreiben Zweiparameterschare von den entsprechenden „laufenden“ Wellen mit freien Parametern k $(k > 0)$und c_0 $(c_0 \neq 0)$.Sie enthalten zwei gleichartige Impulse mit den gleichen Phasen und mit verschiedenen Amplituden. Ihre räumlichen Abweichungen sind gleich, obwohl die Wellen unter verschiedenen Umständen entstehen. Wegen der Gleichheit

$$C - 3k^2c_1^2\frac{\alpha_3}{\alpha_4} = \frac{3(\alpha_2\alpha_4 - \alpha_0\alpha_3)^2}{25\alpha_3\alpha_4^3}$$

hängen die Erhöhungen von den freien Parametern nicht ab. Die solitären Stoßwellen (9) und (10) sind alternativ, d.h. es wird nur eine von denen für alle zulässigen Werte der Parameter erzeugt. Die Wellen der Schar (8) haben die steilste Front. Sie entstehen nur im Fall $\alpha_2\alpha_4 > \alpha_0\alpha_3$. Die aus kurzlebenden und sehr instabilen Wellen bestehende rationale Schar (11) hängt nur von einem Parameter ab. Diese Wellen bilden sich nur für einen Wert der Wellenzahl k und unter der Bedingung $\alpha_2\alpha_4 > \alpha_0\alpha_3$. Ihre scharfen Gipfel befinden sich in den Polstellen, d.h. $\xi = x + \omega t = 0$,

$0 < t < \infty, -\infty < x < \infty$. Das passiert nur für eine negative Frequenz ω, definiert durch (7), d.h. nur bei nach rechts „laufenden" Wellen.

III. Knoidale Wellen

Die knoidalen Lösungen sind besondere Lösungen der Evolutionsgleichungen, die durch die zweite elliptische Funktion von Jakobi [7] dargestellt werden. Sie ist $cn(\theta, m)$ mit (m Modul der Funktion) $0 < m < 1$.

Zum Vermeiden der bedingten Lösung [13] wird die Lösung von (1) in der Form „laufender" Welle

$$u(t,x) = f(\theta) \tag{12}$$

gesucht. Die Phase θ ist $\theta = k(x + \gamma t)$ mit $k > 0, \gamma \neq 0$.

Die Funktion $f(\theta) \in C^4(\Omega)$ ist eine unbekannte glatte reelle Funktion. Nach Einsetzen von (12) in (1) und nach Integrieren nach θ wird folgende nichtlineare Differentialgleichung dritter Ordnung erhalten:

$$\gamma f + \frac{\alpha_0 f^2}{2} + \alpha_1 k f' + \alpha_2 k^2 f'' + \alpha_3 k^3 f''' + \alpha_4 k f f' = G \tag{13}$$

G ist die kumulative Integrationskonstante. Im Sinne der Methode der Abbildung und der singulären Verformungen wird die Lösung von (13) durch das Verhältnis

$$f(\theta) = S\Psi^2(\theta) + T \tag{14}$$

dargestellt. S und T sind unbekannte Parameter mit $S \neq 0$ zum Vermeiden der trivialen Lösung. $\Psi(\theta)$ ist eine unbekannte reelle Funktion, die der Riccati- Gleichung

$$\Psi' = \sqrt{-m^2\Psi^4 + (2m^2 - 1)\Psi^2 + 1 - m^2} \tag{15}$$

mit $m \in \mathbb{R}$ genügt. Aus der Theorie der elliptischen Funktionen [6], [7] ist es bekannt, dass für $0 \leq m \leq 1$ die Lösung der Gleichung (15) die zweite elliptische Funktion von Jacobi $\Psi(\theta) = cn(\theta, m)$ ist. Nach Einsetzen von (14) und (15) in (13) ergibt sich das folgende Gleichungssystem:

$$\begin{gathered} \alpha_0 S^2 - 12 S \alpha_2 m^2 k^2 = 0 \\ \gamma S + \alpha_0 S T + 6 S \alpha_2 k^2 (2m^2 - 1) = 0 \\ \gamma T + \frac{1}{2}\alpha_0 T^2 + 2S(1 - m^2)\alpha_2 k^2 = G \\ S^2 \alpha_4 k - 12 S \alpha_3 m^2 k^3 = 0 \\ 2 S \alpha_1 k + 12 S \alpha_3 k^3 (2m^2 - 1) + 2 S T \alpha_4 k = 0 \end{gathered} \tag{16}$$

Das System hat eine Lösung nur wenn

$$\alpha_0\alpha_3 = \alpha_2\alpha_4 \tag{17}$$

Es folgt: wenn eine knoidale Lösung existiert, ist sie immer bedingt. Diese Lösung ist:

$$S = 12k^2m^2\frac{\alpha_3}{\alpha_4}$$

$$T = -\frac{\alpha_1 + 6\alpha_3 k^2(2m^2 - 1)}{\alpha_4}$$

$$G = -\frac{\alpha_0\alpha_1^2}{\alpha_4^2} + 6\frac{\alpha_3 k^2\left(4\alpha_2\alpha_4(1 - m^2) - \alpha_3(2m^2 - 1)\right)}{\alpha_4^2}$$

$$\gamma = \frac{\alpha_0\alpha_1}{\alpha_4}$$

mit $0 < m \leq 1, k > 0$. Die erhaltene lokalisierte bedingte knoidale Lösung hat die Form:

$$u(t,x) = 12m^2k^2\frac{\alpha_3}{\alpha_4}cn^2(\theta; m) - \\ - \frac{\alpha_1 + 6\alpha_3 k^2(2m^2 - 1)}{\alpha_4} \tag{18}$$

Für $m = 1$ gilt $cn(\theta, 1) = sech\theta$. In diesem Sonderfall ist die Lösung:

$$u(t,x) = 12k^2\frac{\alpha_3}{\alpha_4}sech^2(\theta) - \frac{\alpha_1 + 6\alpha_3 k^2}{\alpha_4} \tag{19}$$

eine von einem Parameter abhängige Schar von solitären Wellen.

Die bedingten knoidalen und solitären Lösungsscharen der Gleichung des Konvektiosfluids sind nicht überraschend, sondern- erwartet. Bei der Auswahl in (12)

$$\gamma = \frac{\alpha_2\alpha_3}{\alpha_4}$$

gilt das Verhältnis

$$u_t + \alpha_0 u u_x + \alpha_1 u_{xx} + \alpha_2 u_{xxx} + \alpha_3 u_{xxxx} + \alpha_4 (u u_x)_x = \\ = k\alpha_2\left(\frac{\alpha_3}{\alpha_4} + k\frac{d}{d\xi}\right)\left(\frac{\alpha_4}{\alpha_2}f''' + \frac{\alpha_0\alpha_3}{\alpha_1\alpha_2}ff' + f'\right)$$

Die Bedeutung ist: Jede lokalisierte Lösung der reduzierten Korteweg-de Vries-Gleichung

$$\frac{\alpha_4}{\alpha_2}f''' + \frac{\alpha_0\alpha_3}{\alpha_1\alpha_2}ff' + f' = 0$$

wird mit Hilfe von (12) auf eine exakte Lösung der Gleichung (1) abgebildet. Zum Beispiel: Diese Gleichung hat eine Ein-Soliton-Lösung der Form

$f(\theta) = k^2 sech^2\left(\frac{k}{2\sqrt{3}}\sqrt{\frac{\alpha_1}{\alpha_4}}\left(x + \frac{\alpha_2\alpha_3}{\alpha_4}t\right)\right)$, die auch eine Lösung der Gleichung des Konvektionsfluids (1) unter der Bedingung (17) ist.

IV. Abschließende Bemerkungen

Es ist logisch, die Anwesenheit der in [4] erhaltenen einzelnen solitären Wellen in der Zweiparameterschar der solitären Stoßwellen (9) zu erwarten. Nach dem Ansatz $c_0 = -1, c_1 = 0, c_2 = 1, k = \pm\frac{\alpha_2\alpha_4 - \alpha_0\alpha_3}{10\alpha_3\alpha_4}$ stimmen beide Scharen überein. Dieser Umstand bestimmt die Notwendigkeit, die verallgemeinerte Version der Methode der Abbildung und der singulären Verformungen zu verwenden. Einige Autoren wie Kudryashov[14], Ryabov-Sinelshchikov-Kochanov [15], Parkes-Duffy-Abbott [13]schlagen vor, an der Stelle der Riccati-Gleichung $\varphi' = c_0\varphi^2 + c_1\varphi + c_2$ die Gleichung $\varphi' = -\varphi^2 + 1$ $(c_0 = -1, c_1 = 0, c_2 = 1)$ zu nehmen. Für manche nichtlineare Evolutionsgleichungen ist die letzte Gleichung nicht passend. Sie überträgt die Abhängigkeiten zwischen den apriorischen Parametern auf Abhängigkeiten zwischen den aposteriorischen Parametern. Dies macht die lokalisierten Lösungen bedingt. Sie haben

beschränkte praktische Anwendung. Die verallgemeinerte Version der Methode der Abbildung und der singulären Verformungen erlaubt, sowie solitäre als auch rationale Wellen in der Gleichung des Konvektionsfluids zu erhalten. Diese Wellen sind mit der direkten Methode nicht erkennbar. Obwohl die verallgemeinerte Version der Methode der Abbildung und der singulären Verformungen eine höhere Empfindlichkeit aufweist, reicht es nicht aus, im allgemeinen Fall die knoidalen Wellen zu erfassen. Im allgemeinen Fall wurde eine analytische knoidale Lösung von Kamenov [16] erhalten. Dafür wurde die räumliche Version der bilinearen Transformationsmethode angewandt.

LITERATURVERZEICHNIS

[1] H.Aspe, M. C. Depassier, „Evolution equation of surface waves in a convectlng fluid“, *Phys. Review A*, Vol. 41, № 6, pp. 3125-3128, 1990

[2] Y.Kuramoto, T. Tsuzuki, „Persistent propagation of concentration waves in dissipative media far from thermal equilibrium“, *Prog. Theor. Phys.*, Vol. 55, Issue 2, pp.356–369, February 1976

[3] C.S.Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura, „Method for solving the Korteweg- de Vries equation“, *Phys. Rev. Lett.* Vol.19, №19, pp. 1095-1097, 1967

[4] Sen-yue Lou, Guo-xiang Huang, Hang-yu Ruan, „Exact solitary waves in a convecting fluid“, *J. Phys. A Math. Gen.* Vol. 24, №11, pp. L587-L590, 1991

[5] A.V.Porubov, „Exact travelling wave solutions of nonlinear evolution equation of surface waves in a convecting fluid“, *J. Phys. A: Math. Gen.*, Vol. 26, №17, pp. L797-L800, 1993

[6] N.I.Ahieser, *Elemente der Theorie der elliptischen Funktionen*, (auf Russisch), Nauka, Moskau, 1970

[7] D.F.Lawden D., *Elliptic Functions and Applications*, Springer-Verlag New York, 1989

[8] J.Weiss, M. Tabor, G. Carnevale, „The Painlevé property for partial differential equations”, *J. Math. Phys.*, Vol. 24,issue 3, pp.522-526, 1982

[9] O.Y.Kamenov, „Periodic solutions of the non-integrable convective fluid equation, *J.Math. Phys.,Vol.* 53,issue 6, 063705, 2012

[10] R.Hirota, „Exact N-soliton solutions of the wave equation of long waves in shallow-water and in nonlinear lattices“, *J. Math. Phys.*, Vol. 14, pp. 810-814, 1973

[11] Y.Matsuno, *Bilinear Transformation Method*, Academic Press, New York 1984

[12] M.Toda, *Theory of Nonlinear Lattices***,** Springer, Berlin, 1981

[13] E.J.Parkes, Duffy, B. R. & Abbott, P. C., „The Jacobi elliptic-function method for finding periodic-wave solutions to nonlinear evolution equations“, *Physics Letters A*, 295, pp. 280-286, 2002

[14] N.A.Kudryashov, „Simplest equation method to look for exact solutions of nonlinear differential equations“, *Chaos, Solitons & Fractals, Elsevier*, Vol.24, №5, pp. 1217-1231, 2005

[15] P.N.Ryabov, D. I. Sinelshchikov, M. B. Kochanov, „Application of the Kudryashov method for finding exact solutions of the high order nonlinear evolution equations“, *Appl. Math. and Comp., Elsevier*, Vol.218, №7, pp. 3965-3972, 2011

[16] O.Y.Kamenov, Doktorarbeit (DSc) „Räumliche Abweichungen in den periodischen Lösungen der nichtintegrierbaren Evolutionsgleichungen”,(auf Bulgarisch) TU-Sofia, Sofia, 2014

Transparent Governance for Corruption Prevention

Transparente Verwaltung für Korruptionsprävention

Zwetelina Gankova – Ivanova[1], Svetlozar Todorov[2]

1: Technical University of Gabrovo, Faculty of Economics
Gabrovo, Bulgaria, e-mail address: zwetelina_gankova_7@yahoo.de

2: e-mail address: z.todorov@abv.bg

Zusammenfassung — Das Ziel dieses Beitrags besteht darin**,** das Korruptionsphänomen in Bulgarien als ein ernsthaftes soziales und wirtschaftliches Problem zu behandeln. Ohne auf die detailierten begrifflichen Definitionen und Diskussionen einzugehen, wird die Korruption von der Perspektive der Europäischen Union behandelt. Die wichtigsten Aktivitäten der Antikorruptionspolitiken werden dargestellt und in diesem Zusammenhang die Notwendigkeit einer effizient funktionierenden und transparenten Verwaltung als Voraussetzung der Korruptionsbekämpfung in Bulgarien begründet.

Abstract — The aim of this paper is to deal with the corruption phenomenon in Bulgaria as a serious social and economic problem. Without addressing the detailed conceptual definitions and discussions, corruption is dealt with by the European Union's perspective. The main activities of the anti-corruption policy are presented and in this context the need for efficient and transparent governance administration as a precondition for corruption prevention in Bulgaria is justified.

I. Einführung

Mit dem Beitritt Bulgariens in die Europäische Union (EU) wurde die Verwaltungsintegration zunehmend die größte Herausforderung für die öffentlichen Verwaltungsstrukturen. Zu den strategischen Zielen der Verwaltungskapazität gehören folgende Elemente: Verbesserung der Arbeit der Staatsverwaltung zur Umsetzung wirksamer Politiken, Bereitstellung qualitativ hochwertiger Dienstleistungen für Bürger und Unternehmen und Schaffung von Bedingungen für ein nachhaltiges Wirtschaftswachstum und Beschäftigung einerseits und Erhöhung der Professionalität, Transparenz und Rechenschaftspflicht im Justizsystem, die Korruptionsbekämpfung andererseits.

Im Hinblick auf die EU-Mitgliedschaft wächst die Bedeutung der Wettbewerbsfähigkeit bulgarischer Volkswirtschaft wesentlich. Sie ist direkt mit der Bekämpfung der Korruption verbunden, da nicht nur die Optimierung im Bereich des privaten und öffentlichen Sektors beeinträchtigt, sondern die vollwertige Mitgliedschaft des Landes in der EU durch Verzerrung der Marktmechanismen gestört werden. Laut dem internationalen Korruptionswahrnehmungs-Index (CPI) von *Transparency International*[1] nimmt Bulgarien 2011 den ersten Platz in der EU mit einem Index von 3.3 ein. [13] Ein schlechteres Ergebnis schnitt das Land nur im Jahr 1998 (2.9) ab. Die Angaben von *Transparency International* zeigen, dass die Verschlechterung der Kennzahlen immer stärker wird – von einer Bewertung 3.8 im Jahr 2009, über 3.6 im Jahr 2010 auf die neue 3.3 im Jahr 2011.

Die Korruption auf den oberen Verwaltungsebenen, der Handel mit Einfluss sowie das Defizit an Rechenschaft und Transparenz bei der Vergabe öffentlicher Aufträge üben einen negativen Einfluss auf die Wirtschaftsentwicklung als ganzes und insbesondere auf die Inanspruchnahme der EU-Fördermittel auf den Beitritt des Landes in den Schengen-Raum und auf die vollwertige Mitgliedschaft in der EU aus.

Die Korruption hindert das Wachstum und die erfolgreiche Umsetzung der Stabilisierungsstrategien im wirtschaftlichen Bereich dort, indem diese Situation zur Erhöhung der Staatsausgaben führt und den unstabilen Staatshaushalt dadurch beansprucht und die ausländischen Investoren abschreckt.

Zur Verbesserung der Transparenz sowie der Rechenschaftspflicht und der Integrität in der Tätigkeit der Staatsverwaltung sind Maßnahmen in der Strategie für transparente Regierungsführung, Prävention und Bekämpfung der Korruption 2006-2008 festgelegt. Sie finden Ausdruck im Programm für Transparenz in den Aktivitäten der Staatsverwaltung und wurden 2006 entwickelt. Im Operationellen Programm *Gute Verwaltung* in der Förderperiode 2014-2020 und in der Strategie zur Entwicklung der E-Verwaltung in der Republik Bulgarien 2014-2020 sind weitere Maßnahmen vorgesehen

Es gibt bereits ein umfassendes Konzept für die Entwicklung von E-Government in Bulgarien. Allerdings ist eine relative Verzögerung im Prozess der E-Government-Implementierung im Vergleich zu den anderen EU-Mitgliedstaaten zu beobachten. Bisher konnten vier Hauptgründe für die unzureichende Umsetzung von E-Government in Bulgarien identifiziert werden: das Fehlen einer angemessenen regulatorischen Grundlage; die mangelnde Interoperabilität der Informationssysteme in der Verwaltung; der Mangel an einem vollständigen elektronischen Austausch auf abteilungsübergreifender Ebene und das ungelöste Problem der Datenvereinigung.

[1] Der Index umfasst 183 Länder, die nach einem geographischen und politisch ökonomischen Prinzip zugeordnet sind. Bulgarien nimmt mit 3.3 den 86. Platz im Jahr 2011 weltweit ein (je niedriger der Index ist, desto höher ist die Korruption im Land).

II. Die Korruption in Bulgarien von der Perspektive der Europäischen Union

Im Hinblick auf die EU-Mitgliedschaft wächst die Bedeutung der Wettbewerbsfähigkeit bulgarischer Volkswirtschaft wesentlich. Das Erreichen des erforderlichen Niveaus ist direkt mit der Bekämpfung der Korruption verbunden, die durch Verzerrung der Marktmechanismen nicht nur die Optimierung im Bereich des privaten und öffentlichen Sektors beeinträchtigt, sondern die vollwertige Mitgliedschaft des Landes in der EU stört. Die Korruption auf den oberen Verwaltungsebenen, der Handel mit Einfluss sowie das Defizit an Rechenschaft und Transparenz bei der Vergabe öffentlicher Aufträge üben einen negativen Einfluss auf die Wirtschaftsentwicklung Bulgariens als ganzes und insbesondere auf die Inanspruchnahme der EU-Fördermittel.

Das Hauptinstrument zur Messung der Korruptionsumgebung ist der *jährlich durch die Gemeinschaft Transparenz ohne Grenzen erfasste Index***.** Im Zeitraum 1998 – 2006 zeigen die Untersuchungen in Bulgarien einen langsamen, aber kontinuierlichen Anstieg dessen Werte.

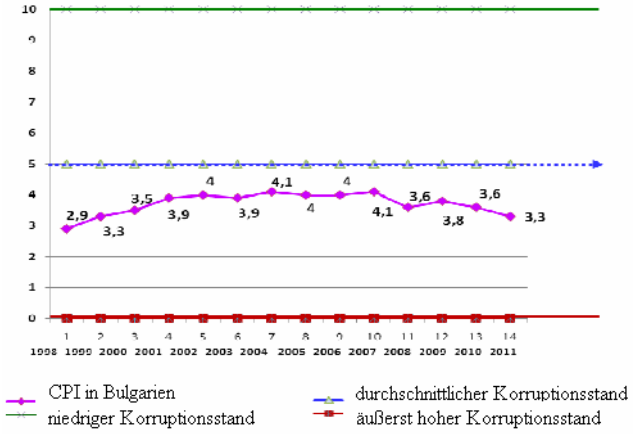


Abb. 1. Korruptionsindex in Bulgarien (Quelle: Transparency International, Bulgaria, CPI 2011)

In der darauf folgenden Zeitperiode 2011-2016 ist ein Anstieg dieser Werte zu beobachten und 2016 erreichen sie 4,1. (http://transparency.bg/)

Der Beitritt Bulgariens in die EU ist von einer Reihe spezifischer Maßnahmen begleitet, die gewisse Mängel in einigen Bereichen wie Flugsicherheit, Lebensmittelsicherheit, Agrarfonds und Justizreform, Bekämpfung der Korruption und der organisierten Kriminalität verhindern und beseitigen müssen. Hinsichtlich der *Justizreform* und der *Bekämpfung der Korruption und der organisierten Kriminalität* ist ein *Mechanismus für Zusammenarbeit und Überprüfung des Fortschritts* entwickelt [2]**.**

Die Korruption wird durch die europäischen Institutionen ständig überwacht, und Bulgarien ist als EU-Land weiterhin ein Teil des Mechanismus für Zusammenarbeit und Überwachung im Bereich Justiz und Inneres. Die Umsetzung dieses Mechanismus wird weiter fortgesetzt, bis jeder der für Bulgarien anwendbaren Indikatoren zufriedenstellend umgesetzt wird. Die Indikatoren sind zum Zeitpunkt des EU-Beitritts bestimmt und beziehen sich auf wesentlich bedeutende Angelegenheiten für das Funktionieren der Mitgliedländer, d.h. auf Unabhängigkeit, Effizienz und Ethik der Justiz und auf die Bekämpfung der Korruption und der organisierten Kriminalität.

Der vierte und der fünfte Indikator des Mechanismus beziehen sich auf die Notwendigkeit wirksamer Maßnahmen zur Korruptionsbekämpfung landesweit, einschließlich gegen die Korruption auf hoher Ebene und in den öffentlichen Einrichtungen.

Der Bericht der Europäischen Kommission an das Europäische Parlament und den Rat der Europäischen Union, der im Januar 2017 veröffentlicht wurde, besagt, dass die von der bulgarischen Regierung entwickelte aktualisierte nationale Anti-Korruptions-Strategie ein wichtiger Schritt sei, jedoch bleibe die Herausforderung, dass diese Strategie umgesetzt wird, indem der notwendige ordnungspolitische Rahmen einzuführen und die geplanten Institutionen einzurichten sind. Die Kommission stellt fest, dass die in dieser Hinsicht erzielten Fortschritte ein Jahr nach der Veröffentlichung des Berichts im Januar 2016 noch eingeschränkt seien. In diesem Sinne wird auf die Maßnahmen gegen die Korruption auf hoher Ebene und auf den abgelehnten Gesetzentwurf und damit auf das fehlende im Voraus geplante neue einheitliche Korruptionsbekämpfungsorgan hingewiesen. Nach Angaben der Kommission wäre diese Behörde für die Kontrolle der Interessenkonflikte und für die Eigentumserklärungen von leitenden Beamten sowie für die Durchführung von Verwaltungsermittlungen in Bezug auf mögliche Korruptionsfälle und rechtswidrige Bereicherung zuständig. (https://ec.europa.eu/info/sites/info/files/com-2017-43_en.pdf)

Die Institutionen und Strukturen, die die Grundfunktionen des Staates ausüben, sind von der Korruption am stärksten betroffen. Ein schwerwiegender Grund für Korruptionshandlungen im öffentlichen Bereich ist die Verflechtung des privaten und des öffentlichen Interesses der befugten Amtsträger. Zu den durch die Korruption meist bedrohten Bereichen in der öffentlichen Verwaltung zählen die öffentlichen Ausschreibungen, die Erteilung von Zulassungen und Lizenzrechten, die Privatisierung, die Einziehung von Steuern, Zollabgaben und Gebühren, die Anstellung von Beamten u. a.

Durch die Mechanismen der öffentlichen Aufträge stellt die Transformation öffentlicher Fonds in private Finanzmittel eine potenzielle Voraussetzung für Korruption und Betrug dar. Wegen des Korruptionsniveaus und der Schwächen in der Verwaltungskapazität stößt Bulgarien auf Schwierigkeiten bei der Verwirklichung der EU-Programme. Nach Ermittlungen des Europäischen Amts für Betrugsbekämpfung, kurz OLAF[2], wurden Mängel und Unregelmäßigkeiten in der Verwaltung der durch EU bereitgestellten Mittel festgestellt und deswegen sind einige Fonds und Finanzinstrumente ausgesetzt. [12] Die Korruptionspraktiken erweisen sich aber als Haupthindernis für die effiziente Nutzung der EU-Finanzhilfen und als Grund für das Aussetzen der Mittel in den nächsten Jahren, was seinerseits die Wirtschaftsentwicklung negativ beeinflusst.

Die Korruption und die organisierte Kriminalität sind auch Gründe für den ausbleibenden Beitritt Bulgariens in den Schengen-Raum. Die entgangenen wirtschaftlichen Vorteile aus diesem Ausbleiben sind enorm.

III. Antikorruptionspolitiken

Der Entwurf eines Gesetzes zur Bekämfung der Korruption und zur Beschlagnahme des illegal erworbenen Vermögens

[2] Das Europäische Amt für Betrugsbekämpfung (franz.: Office Européen de Lutte Anti-Fraude, OLAF) ist ein Amt der Europäischen Kommission, dessen Aufgabe der Schutz der Finanzinteressen der EU ist. Dazu gehört die effiziente, verantwortungsvolle und transparente Bekämpfung des Finanzbetrugs, der Korruption und sonstiger gesetzwidriger Handlungen, einschließlich der schlechten Verwaltung der europäischen Institutionen, die die EU betreffen.

wurde 2016 abgelehnt, 2017 wird es gebilligt. Auf dieser Grundlage wird die Gründung eines Antikorruptionsorgans in Bulgarien vorgesehen, welches die Vorschriften der Kommision zur Korruptionsbekämpfung und zur Beschlagnahme des illegal erworbenen Vermögens im Sinne dieses Gestzes erfüllen wird.
(http://www.justice.government.bg/Files/.pdf)

In Bulgarien erweist sich die Korruption als eines der Haupthindernisse im Prozess der Etablierung der demokratischen Werte und der europäischen Integration und in der Schaffung eines günstigen Umfelds für ausländische Investitionen und für beschleunigte Wirtschaftsentwicklung.

Das *erfolgreiche Entgegenwirken der Korruption* erfordert: *politischen Willen und proaktive Handlungen; Wirtschaftsumfeld mit gültigen und klaren Regeln*, objektiv angemessene Kriterien und funktionierende Mechanismen für deren Umsetzung in die Praxis; *strenge Finanzkontrolle und Rechenschaftspflicht; institutionelle Reform; Bürgerwillen und Ausübung von Zivilkontrolle über die Verwaltung; unabhängige Justiz; engagierte Medien*.

Die Antikorruptionspolitiken gehören zu den Hauptkriterien für die erfolgreiche Wirtschaftspolitik Bulgariens. Das Entgegenwirken der Korruption wird zu neuen ausländischen Investionen und dadurch zur Entwicklung der Volkswirtschaft beitragen. Im Interesse der westlichen Länder und der EU liegt, dass Bulgarien zu einem Rechtsstaat mit stabilen politischen Institutionen und wettbewerbsfähigem Marktumfeld wird.

Durch die 2001 konzipierte *Nationale Strategie zum Entgegenwirken der Korruption* sind Institutionen zum Entgegenwirken der Korruption gegründet, neue gesetzliche Maßnahmen ergriffen und das Funktionieren der staatlichen Verwaltung vervollkommnet [1].

Die nächste *Strategie zur transparenten Verwaltung, Prävention und zum Entgegenwirken der Korruption 2006-2008* entwickelt die gesammelten Erfahrungen weiter und konkretisiert die Prioritätsbereiche für Entgegenwirken und Prävention gegen die Korruption auf den oberen Verwaltungsebenen.

Auf der Zentralebene funktioniert der *Rat zur Koordinierung der Antikorruptionsausschusse* seit April 2006, der die Verbindung zwischen den Ausschussen für Bekämpfung der Korruption im Oberen Justizrat (OJR), im Parlament und im Ministerrat pflegt.

In allen regionalen Verwaltungen funktionieren *öffentliche Regionalräte zur Bekämpfung der Korruption, die Programme zur Durchführung der Strategie zur transparenten Verwaltung und Prävention gegen die Korruption* verabschieden. Seit 2004 werden diverse *Weiterbildungen in Prävention und Entgegenwirken der Korruption* für Vertreter der staatlichen Verwaltung auf allen Verwaltungsebenen durchgeführt [4]. Während der Durchführung der *Strategie zur transparenten Verwaltung, Prävention und zum Entgegenwirken der Korruption* wurde ein Projekt des Ministeriums der staatlichen Verwaltung und der Verwaltungsreform im Jahr 2006 unter dem Titel *„Entgegenwirken der Korruption in der staatlichen Verwaltung durch Weiterbildung der darin tätigen Beamten“* realisiert.

Hinsichtlich der Popularisierung des *Ethik-Kodexes für die Beamten in der staatlichen Verwaltung* sind *Standards für Verwaltungsethik* [3] in allen Verwaltungsstrukturen im Jahr 2006 zur Verfügung gestellt**.**

Ein gutes Beispiel für das Entgegenwirken und die Prävention gegen die Korruption sind die *Maßnahmen an den Grenzkontroll- und Zollstellen* [5], wodurch ein wesentlicher Fortschritt zu bemerken ist. Dazu zählen Kontrolle und Strafmaßnahmen, keine Toleranz, Kontrollen nach bestimmten Angaben, bzw. Unerwartete Kontrollen, Videoüberwachung, Informationsbroschüren, regelmäßige Aus- und Weiterbildung der Angestellten, physiologische Untersuchungen und Umfragen, Bekanntgabe von aufgedeckten Korruptionsfällen zwecks Prävention, Einführung eines Systems zur einheitlichen Zahlung und eines Schichtdienstsystems. [5]

Die Zivilgesellschaft ist an der Beurteilung der Antikorruptionspolitik aktiv beteiligt. Durch die Unterstützung der Medien werden zahlreiche öffentliche Diskussionen über die Korruption veranstaltet. Koalition 2000 ist eine solche Initiative. Zu deren Aktivitäten zählen: die Erarbeitung eines Aktionsplans gegen die Korruption, ein Monitoringsystem, eine Antikorruptionskampagne zur Information und Aufklärung und jährliche Berichte zur Beurteilung des Korruptionsniveaus im Land.

Durch das *PHARE-Programm* zur Entwicklung der Zivilgesellschaft ist ein *Projekt über die wachsende Bedeutung der Zivilgesellschaft in der Bekämpfung der Korruption* verwirklicht. Im Zeitraum 2008-2010 ist die Durchführung des Projekts im Zusammenhang mit dem *Übergangsfinanzinstrument* (Transition facility) *“Entwicklung der Zivilgesellschaft”* verwirklicht**,** wodurch eine Etablierung der Zivilkontrolle und eine aktive Partnerschaft zwischen der Zivilgesellschaft und der Verwaltung im Rahmen der Erarbeitung und Durchführung einer effizienten Antikorruptionspolitik und Instrumente zu erzielen sind.

Die Kontrolle über die Tätigkeit der Verwaltung ist besonders wichtig sowohl für die Sicherstellung der Effizienz, als auch für das Entgegenwirken und die Prävention gegen die Korruption. In diesem Zusammenhang werden unterschiedliche Maßnahmen *zur internen und externen Aufsicht* unternommen.

In Bezug auf *die interne Aufsicht* werden in allen Ministerien *Aufsichts- und Revisionsstellen* errichtet, die dem entsprechenden Minister in Übereinstimmung mit dem Gesetz über die Verwaltung unterstellt werden**.** Ihre Unabhängigkeit ist im Gesetz festgelegt. Zu deren Befugnissen gehören:

- *Analyse der Effizienz in der Verwaltung*, indem diese Analyse bei der Optimierung der Verwaltungsstruktur dienen muss;
- Kontrolle über die *Einhaltung der internen Ordnungsregeln* hinsichtlich der Tätigkeitsorganisation in der Verwaltung;
- *Überprüfung der Beschwerden, Anträge und Meinungen* gegen gesetzwidrige und/oder fehlerhafte Handlungen bzw. Nichthandlungen der Beamten.

Nur 24% der Aufsichtsstellen geben ihre Tätigkeit durch Informationen auf einer Internetseite, durch regelmäßige Pressekonferenzen und Berichte öffentlich bekannt.

In Übereinstimmung mit dem Gesetz über die Verwaltung ist eine *Hauptaufsichtsstelle* im Rahmen des Ministerrates gegründet, die dem Ministerpräsidenten direkt untergeordnet ist. Zu den *Hauptfunktionen dieser Stelle* gehören:

- Unterstützung der Tätigkeitsorganisation bei der Durchführung der staatlichen Antikorruptionspolitik;
- Analyse der Effizient der staatlichen Verwaltung, die mit der Antikorruptionspolitik verbunden ist;
- Koordinierung und Unterstützung der Tätigkeit in den lokalen Aufsichtsstellen;
- Überprüfung der Berichte über Interessenkonflikten, Korruptionsfälle und sonstige Verstoße der Exekutive und der Beamten im höheren Dienst und Erstellung von Ergebnisberichten;
- Analyse der Ursachen und Bedingungen für die festgestellten Verstoße und Ergreifen von Maßnahmen für deren Behebung;

- Unterstützung bei der Durchführung der Beschlüsse des Ausschusses für Prävention und Entgegenwirken der Korruption.

Zur *Erfüllung dieser Funktionen* übt die Hauptaufsichtsstelle Kontrolle über die Gesamttätigkeit der Verwaltung aus, indem sie allgemeine und spezifische Überprüfungen im Auftrag des Ministerpräsidenten sowie Prüfungen konkreter Fälle aufgrund von einer Anordnung des Ministerpräsidenten durch.

Eine andere Methode zur Ausübung interner Aufsicht über die Tätigkeit der Verwaltung ist *das interne Audit***.** Durch Bewertung und Verbesserung der Effizienz des Risikomanagements, der Kontrolle und der Führung hat dieses Audit eine wesentliche Rolle zum Erreichen einer transparenten und offenen Verwaltung. Das interne Audit erfolgt durch eine *spezialisierte Stelle in jeder Verwaltungsstruktur.* Diese Stelle führt Untersuchungsverfahren aller Strukturen, Programme, Aktivitäten und Prozesse durch, wobei auch die Anweisungsbefugten für die EU-Fördermittel auf allen Ebenen kontrolliert wird. [7].

Die externe Aufsicht wird durch den Rechnungshof mithilfe dessen regionalen und lokalen Filialen (6 territoriale Niederlassungen und 28 Büros landesweit) ausgeübt. Der Rechnungshof übt Finanzkontrolle über alle öffentlichen Behörden aus, einschließlich über die kommunale Selbstverwaltung sowie über die Finanzmittel aus EU-Fonds und Programmen, überprüft jedes Vergabeverfahren von öffentlichen Aufträgen und führt anschließend Aufsicht darüber [8]. Die Ergebnisse werden regelmäßig im Internet veröffentlicht.

Ein weiteres Organ zur Ausübung von externer Aufsicht ist die *Agentur für staatliche Finanzaufsicht* **(ASFA).** Ihre Hauptaufgabe besteht im Schutz der Finanzinteressen durch Finanzprüfungen hinsichtlich der Einhaltung der Vorschriften, im Aufdecken von Vermögensschäden und im Sanktionieren der Schuldigen.

IV. Schlussfolgerungen

Zum Schluss sei es noch einmal zu erwähnen, dass die Korruption immer als ein reproduzierbares nachhaltiges Phänomen bestimmt wird.

In Bulgarien ist das Korruptionsniveau hoch und darunter leiden die Marktreformen und das qualitative Wachstum des Landes, benachteiligt ist auch jedes Individuum. In diesem Sinne sind die Ausarbeitung und die Umsetzung eines wirksamen, effizient funktionierenden Antikorruptions-programms eine obligatorische Voraussetzung für das langfristige Wirtschaftswachstum.

Eine dauerhafte Verbesserung der Korruptionslage ist nur durch Maßnahmen möglich, die Bestandteil eines radikalen Verwaltungsreforms mit effizient funktionierenden Strafverfolgungsbehörden sind.

In Bezug auf die noch hohe Zahl der Korruptionsfälle sind die Mechanismen für Bekanntgabe von Korruptionsfällen und Feedback zu stärken, indem die Öffentlichkeit darüber besser informiert werden soll. Genauso wichtig ist die Durchführung eines regelmäßigen Monitorings der Strategie für transparente Verwaltung, Prävention und Entgegenwirken der Korruption.

Die Politik des Entgegenwirkens der Korruption beinhaltet neben der Politik für Transparenz und Offenheit auch die Kontrolle der Verwaltung und noch weitere Tätigkeiten in anderen Bereichen wie das Auswahlverfahren bei der Anstellung der Beamten, Forschungen auf dem Gebiet der Antikorruptionsmaßnahmen, das gute Management des staatlichen Dienstes, die Entwicklung des E-Governments, die Einführung eines Systems an der Grenzkontrollstellen zur einheitlichen Zahlung aller Gebühren etc.

Es gibt kein einheitliches System zur Bekanntgabe der Ergebnisse im Zusammenhang mit der Tätigkeit aller Aufsichtsstellen. Es mangelt auch an einheitlichen Standards zur Regelung deren Tätigkeit. Für den vollständigen Erfolg der Antikorruptionspolitik ist es von außerordentlicher Bedeutung, dass diese integrative Vorgehensweise, die sich nicht nur auf die Bekämpfung der Korruption, sondern auf deren Prävention bezieht, fortgesetzt wird.

Die Rechtsvorschriften, die die Tätigkeit der Aufsichtsstellen regeln, sind zu vervollkommnen. Dafür ist eine vollständige Analyse zur Identifizierung der dabei entstehenden Probleme erforderlich. Es ist erforderlich, dass einheitliche Standards eingeführt werden, und die Bekanntgabe der Ergebnisse ist zu verbessern, dies bedeutet, dass ein einheitliches Veröffentlichungssystem entwickelt werden muss.

Zur Verbesserung der Transparenz, der Rechenschafts-pflicht und der Kontrolle in der Tätigkeit der Aufsichtsstellen ist es besonders wichtig, dass die koordinierende Funktion der Hauptaufsichtsstelle verstärkt wird.

Bezug nehmend auf die besondere Rolle der internen Audit-Stellen für eine effizientere und transparentere Tätigkeit der Verwaltung ist es auch notwenidg, den Bedarf zur Stärkung deren Kapazität vollständig zu analysieren.

Aufgrund der Wichtigkeit von externer Aufsicht zum Erzielen optimaler Ergebnisse in der Verwaltungstätigkeit und zum Gewähren einer hohen Transparenz und einer wirksamen Bestechungsbekämpfung ist es besonders wichtig den Bedarf zur Stärkung der Kapazitäten entsprechender Aufsichtsstrukturen und –institutionen zu analysieren und die erforderlichen Maßnahmen zu ergreifen.

Literaturverzeichnis

[1] Bericht über die Durchführung der Nationalen Strategie zum Entgegenwirken der Korruption im Zeitraum 2004 – 2005 http://www.acc.government.bg/documents/otchet-2004-1.doc

[2] Beschluss des Ausschusses vom 13.12.2006 zur Entwicklung eines Mechanismus für Zusammenarbeit und Überprüfung des Fortschritts Bulgariens in besonders wichtigen Bereichen der Justizreform und in der Bekämpfung der Korruption und der organisierten Kriminalität, EG (2006/929), OB L 354/56, 14.12.2006.

[3] Bericht über das Programm für Transparenz in der Tätigkeit der staatlichen Verwalung und der Personen im höheren Dienst, zum Dezember 2006, S. 14, http://www.mdaar.government.bg/lucidity.php

[4] Bericht über den Zustand der Verwaltung, 2006, S. 156 http://www.mdaar.government.bg/docs/Annual%20Report%20%202006_26.07.2007.pdf

[5] Berciht der Europäischen Komission über den Fortschritt Bulgariens in Begleitmaßnahmen nach dem Beitritt in die EU, 27.06.2007, S. 19 http://www.evropa.bg/bg/del/info-pad/press-releases/doklad.html

[6] Gesetz über Änderung und Ergänzung der Verfassung Bulgariens (Amtsblatt, Nr.. 27 vom 31.03. 2006), 3. Änderung (§ 3)

[7] Gesetz über das interne Audit im öffentlichen Sektor (Art. 9)

[8] Gesetz über den Rechnungshof (Art. 5)

[9] Gesetz über die kommunale Selbstverwaltung und die kommunalen Behörden (Art. 21a)

[10] Jahresbericht des Ombudsmanns der Republik Bulgarien, Juni 2005 – März 2006 (S. 37) http://www.ombudsman.bg/annual_report.pdf

[11] Programm für Transparenz in der Staatsverwaltung und in der Tätigkeit der Personen im höheren Dienst http://www.mdaar.government.bg/programs.php

[12] Report from the Commission to the European Parliament and the Council on the Management of EU-funds in Bulgaria, Commission of the European Communities, Brussels, 23.7.2008, COM (2008) 496 final.

[13] Transparency International Bulgaria, CPI 2011

Enhancement of surface properties of steel samples using new technique for electroless nickel plating

Verbesserung der Oberflächeneigenschaften von Stahlproben durch neue Technik für stromlose Vernickelung

Julieta Kaleicheva[*], Zdravka Karaguiozova[†], Valentin Mishev[*],
[*]Faculty of Industrial Technology, Technical University of Sofia,
Sofia, Bulgaria, e-mail: jkaleich@tu-sofia.bg, v_mishev@tu-sofia.bg
[†]Space Researche and Technology Institute, Bulgarian Academy of Sciences
Sofia, Bulgaria, e-mail: karazuzi@yahoo.com

Abstract **—** This study is focused on the investigation of the possibilities for materials surface properties enhancement by electrolees nickel plating method. Substrates of steel 17CrNiMo6 are used for the manufacture of the specimens. Electroless method EFTTOM-NICKEL is applied for the development of an innovative technology for production of as plated nickel and composite Nickel coatings. Detonation nanodiamond particles (ND) are used as a strengthening material for production of composite nickel coatings. Suspension of ND is added directly to the electroless bath using a suitable surfactant to achieve well-dispersed particles in the bath and to facilitate their embodiment and equal distribution in the coating. The influence of ND particles on the coatings mechanical and physical properties is confirmed by the results achieved carrying out as morphology and microstructure observations by optical metallography and scanning electron microscopy (SEM) and also wear and nano indentation tests. The influence of the thermal treatment of the obtained coatings on the above mentioned properties is examined.

Zusammenfassung **—** Die vorgestellte Arbeit untersucht die Möglichkeiten zur Verbesserung der Oberflächeneigenschaften der Materialien mittels der Methode der stromlosen Vernickelung. Die Testproben sind aus Stahl 17CrNiMo6 gefertigt. Eine innovative Technologie zur Herstellung von Nickel- und Verbund-Nickel-Beschichtungen ist mit der EFTTOM-NICKEL-Methode entwickelt. Nano-Diamantpartikel (ND), die durch das Detonationsverfahren erhalten werden, sind als Verstärkungsmaterial für die Herstellung von Verbund-Nickel-Beschichtungen verwendet. ND-Salzgel wird direkt zu der Beschichtungslösung unter Verwendung eines geeigneten Tensids zugemischt, um die Partikel in der Wanne besser zu dispergieren und ihre Infiltrierung und gleichmäßige Verteilung in der Beschichtung zu erleichtern. Der Einfluss von ND-Partikeln auf die mechanischen und physikalischen Eigenschaften der Beschichtungen wird von den erhaltenen Ergebnissen bei der Beobachtung der Morphologie und der Mikrostruktur der Beschichtungen durch optische Metallographie und Rasterelektronen-mikroskopie (SEM), sowie von den durchgeführten Verschleiß- und Nanoindentationsversuchen, bestätigt. Ebenfalls ist auch der Einfluss der thermischen Behandlung der erhaltenen Beschichtungen auf die vorgenannten Eigenschaften untersucht.

I. Introduction

Surface finishing of the materials is an opportunity for obtaining new improved properties by surface modification. One preferred, simple and ecologically friendly method is electroless plating of different materials. Many advantages between all of these methods possess electroless nickel plating. Electroless nickel plating is a form of alloy treatment designed to produce consistent, even thickness across component surfaces with zero or compressive stress without an electrical current during the process, characterized by increase resistance and hardness in a metal or plastic, superior corrosion resistance and solder ability [1]. Incorporation of distinct strengthening particles into the deposit to impart a specific property in the electroless nickel coatings (EN) forms composite coatings. Theoretically, almost any type of particle could be co-deposited, as long as it could withstand the conditions within an EN bath, and if it is of the appropriate size [2]. Electroless nanometer composite coatings where nano-particulates used as reinforcing phase are important due to their excellent performance. The incorporation of nano-sized particles within Ni–P autocatalytic coatings greatly improved their properties and added entirely new features to the coatings performance, which enhance their applicability in different industries such as electronic components and computers, general mechanics, automobile, paper mills, textile and food [3]. Many authors prove the improved coatings properties in the presence of nanodiamond in the bath [4 – 8]. But the opinions about the action mechanism of nanoparticles are controversial. Some of the researches believe the properties improvement is due to the nanodiamond incorporation into the Ni-P matrix, while others attribute this improvement to the change of the reaction conditions, which results in the coatings morphology [9].

The mechanical properties mainly depend on the phosphorus/boron content in the deposit. The coatings normally have high strength, limited ductility, and high modulus of elasticity. The ultimate tensile strength of commercial coatings exceeds 700 MPa and allows the coatings to withstand loading without damage [10]. The structural changes during heat treatment at temperatures above 220° C cause a volumetric shrinkage of electroless nickel deposits of up to 6%. This increases tensile stress and reduces compressive stress in the coating. High levels of internal stress also reduce

the ductility of the coating [11]. However, heat treatment has little effect on the ductility [12].

The aim of the present work is to investigate the influence of nano additives of ND to the plating bath and the heat treatment of the obtained coatings on the microstructure and physical and mechanical properties of electroless nickel coatings, plated on the 17CrNiMo6 steel samples. It is supposed the ND shape and particles size (4÷6 nm), as their specific nature and high surface activity ensure higher density of the coatings and others functional qualities.

II. Materials and Investigation Methods

17CrNiMo6 is used for a production of the samples. The composition of the 17CrNiMo6 steel is shown in Table 1. The form of the samples is consistent with the requirement of the respective test method.

Nano sized diamond particles (ND) obtained by detonation method are used as strengthening particles. ND production is performed by an optimized technology in order to achieve high quality product, suitable for use in the electroless deposition of composite coatings [13].

The physical properties of ND are shown in Table 2.

The coatings are produced by advanced technology based on EFTTOM-NICKEL method [14], developed at SRTI-BAS. The strengthening particles are used in a suspension. The formation of the particles aggregates is avoided by a deactivation of the particles surface. The suspension preparation is performed in an environment of suitable surfactants as follows: to 10 ml distilled water the proper quantity of the nano sized powder is added and mixed for 20 minutes with 0,02g of surfactant (in this case sodium lauryl sulphate) before the plating process. The plating process is performed in an electroless nickel plating bath without and with addition of strengthening particles. Two types of coatings are made:

- Electroless nickel coating (Ni)
- Composite nickel with nanodiamond coating (Ni/Ni+ND) in the following technological mode:
  - pH = 4.6 ÷ 4.7
  - T = (92 ÷ 95) ºC
  - Coating time = 15 min (for Ni coating); 10 min (for Ni+ND coating)
  - Optimal uses of the solution 7 ÷ 8 times per hour
  - Concentration of the strengthening particles - 5 g/l .

The preparation of nickel coating takes place through a one-step process, whereas the composite coatings of Ni/Ni+ND are obtained by a two-stage process in which two layers are formed:

- Electroless nickel layer, obtained in a plating bath by "EFTTOM-NICKEL" Method
- Composite nanostructured layer Ni with strengthening particles, obtained in the same bath with nano sized additives concentration of 5 g/l.

Chemical process of nickelization is:

$$3NaH_2PO_2+3H_2O+NiSO4 \rightarrow 3NaH_2PO_3+H_2SO_4+2H_2+Ni$$

The coating composition is shown in Table 3.

TABLE I. Chemical composition of 17CrNiMo6 steel

Composition	C	Si	Mn	Cr	Mo	V	Ni
(%)	0,18	0,20	0,70	1,65	0,30	-	1,55

TABLE II. Physical properties of ND

Strengthening particles	Average size	Specific surface area (m^2/g)
ND	2-4 nm	330-400

TABLE III. Composition, heat treatment, thickness and wear of the coatings

Number of the sample	Coating		Thick-ness, δ (μm)	Wear (g)
	Composition	*Heat treatment*		
1	Ni	-	8.7	7.5
2	Ni	290ºC,6h	8.7	5.8
3	Ni/Ni+ND	-	9.3	8.0
4	Ni/Ni+ND	290ºC,6h	7.3	4.8

Part of the sodium hypophosphite is undergone to reduction process with phosphorous formation:

$$NaH_2PO_2+H \rightarrow H_2O+NaOH+P$$

Hereby the coating obtained is amorphous alloy consists of nickel and phosphorous.

This layer works as a snubber in a two stage process, enhancing the coating ability to take a contact load and improves the adhesion between the sample base and the coating.

After plating the coatings 2 and 4 are undergone to a heat treatment at 290ºC for 6 hours (Table 3) for the coatings' adhesion enhancement and hardness increase.

The below described methods are used to characterize the coatings.

The microstructure is observed by means of an optical metallographic microscope Neophot 32. The samples are treated with 3 % HNO_3 - C_2H_5OH solution before testing. The data for the coatings thickness are also received (Table 3).

The wear of the coatings are tested by friction wear tests performed under 50MPa loading conditions – in accordance with the Polish Standard PN-83/H-04302 (Table 3).

Investigation of the hardness and elasticity modulus of the produced coatings and of the substrate is performed by nanotester FISCHERSCOPE® H100 (Table 4). The data presented for each of the samples are received by defining of the average value of 5 measurements with 95% convergence. The tests are carried out under loading of F=50mN (the penetrating of the indenters is less than 10% of the coating thickness to avoid the substrate influence on the results. The steel substrate is measured under loading F=1000mN.

The results for HU universal hardness (the indenter is in the maximum loaded position taking into account the elastic and plastic deformation at the measuring location (Table 4).

H_{plast}. – plastic hardness (indenter is in the final unloaded position taking into account the residual plastic deformation);

h – maximum depth of penetration of the indenter top under maximum loading during test;

E* – approximate module of elasticity (Jung module). The real E is evaluated by the formula E*= E/(1-v2),

where v is a sliding coefficient (v=0,30, for the steel substrate, but for Ni coatings - 0,29);

W_{total} – total energy spent for an elastic and plastic deformation of the coating during indenter penetration;

W_{total}l = Wr + We , where Wr % – plastic part of W total ; We % – elastic part of W_{total};

HV - recalculated plastic hardness H_{plast}. into Vickers hardness.

TABLE IV. MECHANICAL CHARACTERISTICS OF THE COATINGS

Number of the sample	HU (MPa)	$H_{plast.}$ (MPa)	h (μm)	E* (GPa)	E (GPa)	W_{total} (nJ)	W_r (%)	HV (kg/mm^2)
Coatings data, Load F = 50 mN								
1	6625	7814	0.534	146	134	10.80	66.18	738
2	10267	13117	0.429	185	169	8.90	54.00	1238
3	6670	7968	0.533	143	131	10.80	66.00	752
4	11253	14453	0.410	198	181	8.31	53.59	1364
Substrate data F = 1000 mN								
8	4479	6159	2.478	170	155	10.45	77.25	581

III. RESULTS AND DISCUSSIONS

Steel 17CrNiMo6 samples are normalized before plating. Optical microscope observations reveal all of the tested samples as a white strip, following surface relief and filling the defects such as micro pores and micro cracks (Fig. 1, 2). Samples 2 and 4 are undergone to heat treatment at 290°C for 6 hours after plating. No changes in the substrate microstructure are found. The coatings thickness measurement shows results in the range of 5, 3 ÷ 9, 3 μm (Table 3).

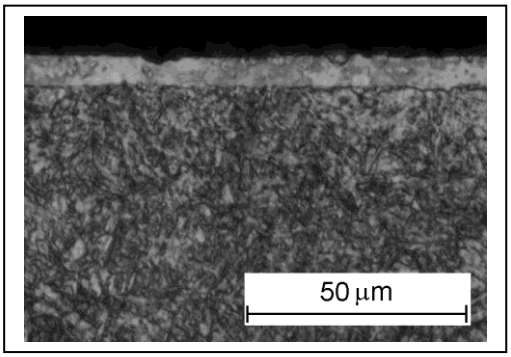


Fig. 1. Ni- coating microstructure after heat treatment at 290°C, 6h (sample 2)

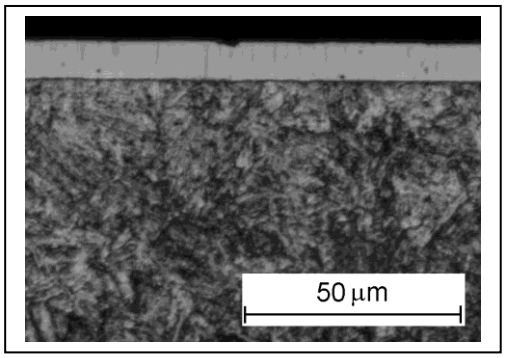


Fig. 2. Ni/Ni+ND coating microstructure after heat treatment at 290°C, 6h (sample 4)

X-Ray analysis shows an amorphous structure for Ni and Ni/Ni+ND coatings in as-plated condition (samples 1 and 3). Heat treatment at 290°C, 6h of the coated samples (samples 2 and 4) initiates a crystallization process in the coating and the qualitative X-Ray analysis of the last mentioned samples proves the presence of the crystalline phase of Ni_3P and Ni in the coatings' structure (Fig. 3, 4) [15].

The tests performed by nano tester FISCHERSCOPE H100 allow receiving data for the hardness and elasticity module of the substrate and of the coatings (Table 4).

The increase of the universal hardness HU from 6625 for nickel coatings to 6670 MPa for composite nickel coatings is received.

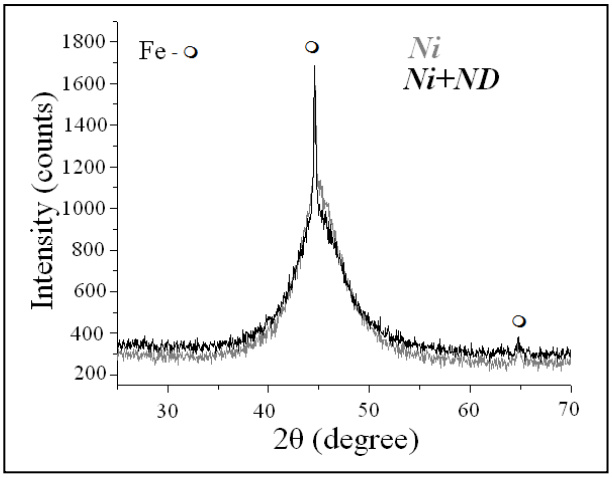


Fig. 3. X– ray diffraction pattern of samples 1 (Ni coating) and 3 (Ni/Ni+ ND coating) in an as-plated condition.

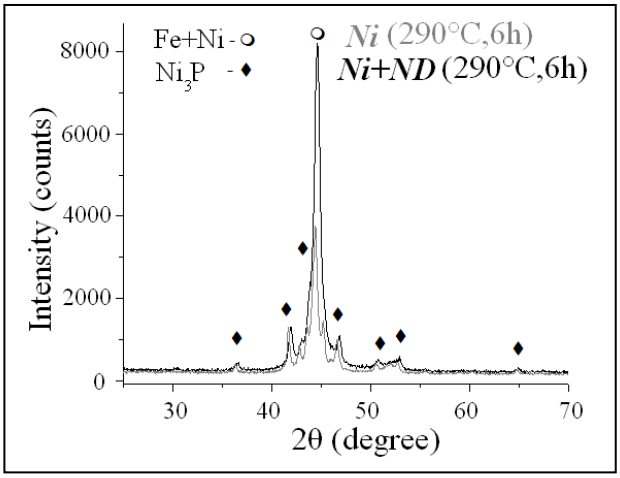


Fig. 4. X– ray diffraction pattern of samples 2 (Ni coating) and 4 (Ni/Ni + ND coating) after heat treatment at 290°C, 6h.

The heat treatment of the coatings at 290°C, 6h leads to the increase of the universal hardness HU to 10267 MPa for Ni coatings and to 11253 MPa for composite Ni/Ni+ND coatings (Fig. 5).

$H_{plast.}$ defined on the base of the residual plastic deformation shows higher value for the composite coatings with addition of ND respectively 7968 compared to the Ni coatings (namely 7814). The heat treatment of the coatings at290°C, 6h leads to the increase of the $H_{plast.}$ to 13117 for Ni coatings and to 14453 for composite Ni/Ni+ND coatings.

The higher universal and H_{plast} hardness achieved after heat treatment at 290°C, 6h could be explained with the structure change and with the dispersed crystalline phase formation of Ni_3P [15].

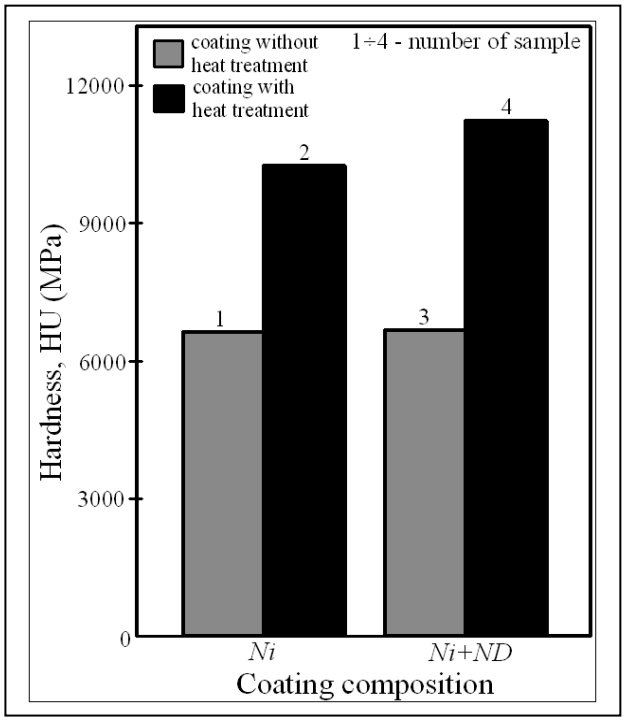


Fig. 5. Universal coatings harness *HU*

The higher hardness of Ni / Ni + ND coating compared to the Ni coating is due to addition of the high hard nano sized particles of ND.

The universal and H_{plast} hardness are higher for all of the coatings compared to this one of the substrates.

The substrate material possess higher approximate module of elasticity E* than this one of the coatings in as plated conditions. While the coatings heat treatment results in higher module values than those of the substrate material.

For the both coatings of Ni and Ni / Ni + ND the heat treatment leads to a significant increase in E*.

The low ductility is due to the amorphous structures of the coatings in as plated condition that essentially precludes plastic deformation. The deformation is therefore mostly elastic until fracture occurs [16].

The W_{total} – total energy spent for an elastic and plastic deformation of the coating during indenter penetration and W_r % – plastic part of W total are in correlation with the results achieved for the approximate module of elasticity E*.

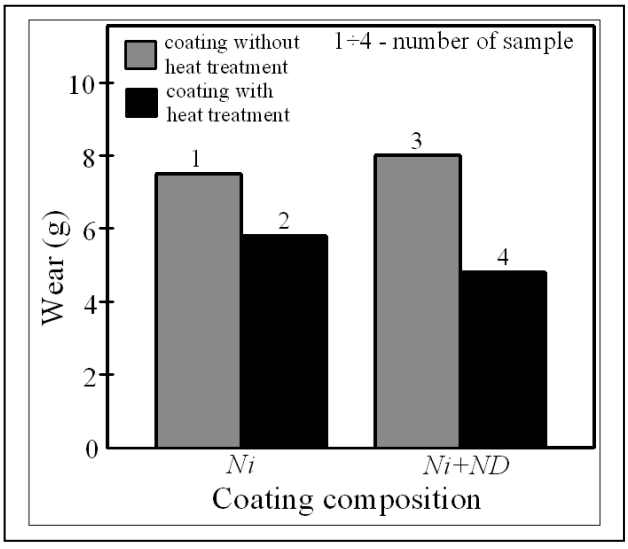


Fig. 6. Wear of the coatings Ni и Ni/Ni+ND

The performed wear test determines the influence of the added particles of nano sized ND on the coatings wear (Table 1, Fig. 6). The correlation between hardness and wear of the samples with coating of Ni and Ni/Ni+ND after heat treatment at 290°C, 6h is established (samples 2 and 4).

IV. Conclusions

Nanosized diamond additions to the plating bath for electroless nickel plating possess a modifying effect on the coatings microstructure, which influences on their physico-mechanical properties. An additional impact has also the coatings heat treatment at 290°C, 6 h:

X-Ray analysis reveals amorphous structure of the coatings in as-plated condition which is converted to a crystalline formation of Ni_3P and Ni phases after heat treatment.

Slight increase in the micro hardness is defined after addition of ND strengthening particles and about 1,7 times increase after heat treatment. Higher hardness of the heat treated samples correlates with the lower wear;

References

[1] Ron Parkins, "Properties and applications of electroless nickel " Nickel Development Institute, 2001.

[2] Lloyd Ploof, "Electroless nickel composite coatings," *Advanced Materials and Process*, Sirius Technology, Inc. Oriskany, New York, pp.36-38, May 2008.

[3] O.A. Leon, M.H. Staia and H.E. Hintermann, "Wear mechanism of Ni–P–BN(h) composite autocatalytic coatings, " *Surf. Coat. Technol*., vol. 200, pp. 1825-1829, 2005.

[4] H. Xu, Z. Yang, M.K. Li, Y.L. Shi, Y. Huang and H.L. Li, "Synthesis and properties of electroless Ni–P–nanometer diamond composite coatings," *Surf. Coat. Tech*., vol. 191, pp.161-165, 2005.

[5] H. Matsubara, M. Kobayashi, H. Nishiyama, N. Saito, Y.Inoue and M. Mayuzumi, "Co-deposition Characteristics of Nanodiamond Particles in Electrolessly Plated Nickel Films, " *Electrochemistry*, vol. 72, pp. 446-448, 2004.

[6] H. Matsubara, Y. Abe, Y. Chiba, H. Nishiyama, N. Saito, K. Hodouchi and Y. Inoue, "Co-deposition mechanism of nanodiamond with electrolessly plated nickel films, " *Electrochimica Acta*, vol. 52, pp. 3047–3052, 2007.

[7] A. Miteva, "Functionally graded materials in tribology, " *Trib. Journal Bultrib*, vol. III, no 3, pp. 371-375, 2013.

[8] A. Miteva, "On the Microstructure and Mechanical Properties of Nanocomposites, " *Proceedings of the eighth scientific conference with International Participation S E S 2012*, Sofia, Bulgaria, pp. 220-225, 2012.

[9] A. Gurga, V. Mochalin, D. Pepe, C. Picardi and Y. Gogotsi, "Nanoindentation Study of the Effect of Nanodiamond Additives on Electroless Deposition Nickel-Boride Coating, " *Advances in Technology of Materials and Materials Processing Journal*, vol 10, no 1, pp. 47-52, 2008.

[10] Jothi Sudagar, Jianshe Lian and Wei Sha, "Electroless nickel, alloy, composite and nano coatings – A critical review, " *Journal of Alloys and Compounds*, vol. 571, pp. 183–204, 2013.

[11] K.G. Keong and W. Sha, "Crystallization and Phase Transformation Behaviour of Electroless Nickel-Phosphorus Deposits and Their Engineering Properties, " *Surf. Eng*., vol.18, pp. 329–343, 2002.

[12] K. Parker, "Effects of heat treatment on the properties of electroless nickel deposits, " *Plat. Surf. Finish*., vol.68, pp.71–76, 1981.

[13] S. Stavrev S., S. Lazarov, K. Stoev, L. Markov and Valeri I. Ivanov . US Patent No. 5353708 ,1994.

[14] G. Gavrilov and C. Nicolov, *Electroless Nickel and Composite Coatings*, Sofia, 1985.

[15] Z. Karaguiozova, J. Kaleicheva, V. Mishev, G. Avdeev and S. Stavrev, "Microstructure and Properties of Electroless Composite Nickel Coatings with Nanodiamond, *Nanoscience and Nanotechnology*, vol.14, pp. 74-77, 2014.

[16] R. Weil, J. Lee and K. Parker, " Comparison of Some Mechanical and Corrosion Properties of Electroless and Electroplated Nickel Phosphorus Alloys, " *Plat. and Surf. Fin*., vol. 76, pp. 62-66, 1989.

Shadow Economy, transparent E-Governance

Schattenwirtschaft, transparente E-Verwaltung

Zwetelina Gankova-Ivanova

Technical University of Gabrovo, Faculty of Economics
Gabrovo, Bulgaria, e-mail: zwetelina_gankova_7@yahoo.de

Zusammenfassung — Das Ziel dieses Beitrags besteht darin**,** das Wesen der Schattenwirtschaft näher zu erläutern, wobei deren typische Erscheinungsformen und Ausprägungen in Bulgarien und die daraus resultierenden Effekte in verschiedenen Sektoren der Wirtschaft im Zusammenhang mit den Anforderungen für eine transparente und faire Staatsverwaltung dargestellt werden. Auf dieser Grundlage sind die bereits bestehenden Erfahrungen und Fortschritte bei der Einführung von E-Government als Instrument der modernen Varwaltung in Bulgarien zu beurteilen.

Abstract — The aim of this paper is to clarify the nature of the shadow economy by presenting its typical manifestations and activities in Bulgaria and the resulting effects in different sectors of the economy in relation to the requirements for transparent and fair state administration. On this basis, the already existing experiences and progress in the introduction of e-government as an instrument of the modern governance in Bulgaria have been discussed and assessed.

I. EINFÜHRUNG

Der Beitritt Bulgariens in die Europäische Union (EU) hat die Schattenwirtschaft und die nicht angemeldete Arbeit, die sog. Schwarzarbeit in einen neuen Kontext gebracht. Die Faktoren, die ihre Verbreitung und die Konsequenzen davon bestimmen, gewinnen internationale Dimensionen zunehmend und überschreiten die nationalen Rahmenbedingungen. Die Effizienz der nationalen Initiativen zur Überwindung dieser Erscheinungen hängen immer mehr von den koordinierten internationalen Bemühungen ab. Dies erfordert auch einen neuen Ansatz zur Messung und Beurteilung dieser Phänomene, was auf einem gemeinsamen europaweiten Standard beruht und zuverlässige internationale Vergleiche ermöglicht.

Ein anderer, nicht wenig wichtiger Aspekt der Regierungspolitik, konzentriert sich auf die Herausforderung, eine transparente und praktikable Verwaltung zu schaffen, deshalb muss der Staat der zunehmenden Bereitschaft zur Ausweichung in die Schattenwirtschaft und zur Schwarzarbeit durch bessere staatliche Leistungen entgegenwirken. Es wird festgestellt, dass je besser die staatlichen Institutionen arbeiten und je positiver die Einstellung gegenüber dem Staat ist, desto mehr ist die Bevölkerung bereit, für die staatlichen Aufgaben und die geplanten staatlichen Gegenleistungen Steuern und Abgaben zu zahlen.

Die Hauptdimensionen in der Entwicklung der öffentlichen Verwaltung stehen im Zusammenhang mit der Forderung der Grundsätze für Transparenz und Rechenschaftspflicht als Voraussetzung für eine verantwortungsvolle Staatsverwaltung. In der Strategie für transparente Regierungsverwaltung, Prävention und Bekämpfung der Korruption, 2006-2008, im Programm für die Umsetzung der Strategie, im Operationellen Programm „Gute Verwaltung“ für die Förderperiode 2014-2020 und in der Strategie für die Entwicklung der E-Verwaltung in der Republik Bulgarien 2014-2020 sind Maßnahmen zur Verbesserung der Transparenz, Rechenschaftspflicht und Integrität in der Tätigkeit der Staatsverwaltung vorgesehen.

Mit dem Beitritt der Republik Bulgarien in die Europäische Union wurde die Verwaltungsintegration die größte Herausforderung für die heutigen öffentlichen Verwaltungsstrukturen. Daher sind die Verbesserung der Arbeit der Staatsverwaltung zur Umsetzung wirksamer Politiken, Bereitstellung qualitativ hochwertiger Dienstleistungen für Bürger und Unternehmen und die Schaffung von Bedingungen für ein nachhaltiges Wirtschaftswachstum und Beschäftigung einerseits und Erhöhung der Professionalität, Transparenz und Rechenschaftspflicht im Justizsystem andererseits strategische Ziele der Verwaltungskapazität.

II. WIRTSCHAFTLICHE RAHMENBEDINGUNGEN

Nach dem Beitritt Bulgariens in die EU erfuhr das Land einen allgemeinen Aufwärtstrend. Auf Grundlage der makroökonomischen Kennzahlen weckt die wirtschaftliche Entwicklung Bulgariens ein besonderes Interesse und stellt einige Fragen zur Diskussion. Einerseits weist Bulgarien selbst das niedrigste BIP pro Kopf (6.630 €; EU: 28.000 €, das BIP-Wachstum für 2016 lag bei 3,4%) sowie eine der höchsten Armutsquoten (21,8%) innerhalb der EU auf. Andererseits beträgt 2016 der Bruttoschuldenstand des Staates 29,5% des BIP, das Defizit des Staatshaushalts – 0%, die Inflation liegt bei -1,3%, die Arbeitslosenquote beträgt 7,6%. Gestützt ist die bulgarische Wirtschaft vor allem durch den Dienstleistungssektor (67,3% des BIP). [6]

Hohe Einkommens-Disparitäten prägen das Bild. Bulgarien lag laut Eurostat mit 3,80 Euro Lohn- und Lohnnebenkosten pro Stunde im Jahr 2014 an letzter Stelle in der EU. Der durchschnittliche Bruttomonatslohn wurde den statistischen Angaben zufolge Anfang 2017 auf 527 Euro berechnet, der Mindestlohn – rund 235 Euro. Der bulgarische Arbeitsmarkt generiert eine niedrige Arbeits- und Innovationskraft. Die Auswanderung und die schrumpfenden Bevölkerungszahlen im erwerbsfähigen Alter stellen ein langfristiges Risiko für das Wachstumspotenzial der Wirtschaft dar (IWF: 2015: 7,2 Mio. Einwohner, davon 20% im Alter über 65; Prognose 2060: 5,5 Mio. Einwohner, davon 32% im Alter über 65).

Eine niedrige Erwerbsbeteiligung und Fehlqualifizierungen deuten laut EU-Kommission weiterhin auf beträchtliche strukturelle Herausforderungen für den Arbeitsmarkt und die Bildungspolitik hin. Die guten Fremdsprachenkenntnisse der Bevölkerung kombiniert mit dem geringen Lohnniveau

machen Bulgarien zu einem attraktiven Outsourcing-Standort für Kundenservicecenter, Call Center und IT–Dienstleistungen (*weitere Schwerpunkte: Automobilzulieferindustrie, Nahrungsmittelindustrie, Metallindustrie, Maschinenbau, Bergbau, Tourismus, Software-Entwicklung, Pharmaindustrie, Textil- und Bekleidungsindustrie, Landwirtschaft*). Für viele Investoren bleibt das Land jedoch weitgehend unbekannt. Eines der wirtschaftspolitischen Ziele ist es, das Angebot von diversen Dienstleistungen durch Verfahrensvereinfachung zu beschleunigen. Damit sollen die Hindernisse für Investoren überwunden werden.

Gleichzeitig ist der Anteil der Schattenwirtschaft am BIP im Jahr 2016 auf 28% berechnet und ist vorwiegend in den Bereichen Tourismus, Bauwirtschaft, Handel, Gesundheitswesen präsent. Hinsichtlich der Korruption in der Verwaltung (laut dem globalen Korruptionsindex von Transparency International-World Ranking 2016) belegte Bulgarien den 75. Platz von insgesamt 176 Ländern (Deutschland: Platz 10). Das schwache Justizsystem und die oligarchischen Strukturen sind weiterhin ein Problem und beeinträchtigen das Investitionsklima und die ökonomischen Möglichkeiten.

III. SCHATTENWIRTSCHAFT – ERSCHEINUNGSFORMEN UND AUSPRÄGUNGEN

Die Schattenwirtschaft zieht die Aufmerksamkeit und das Interesse eines breiten Kreises von Forschern, Experten, Politikern und Bürgern auf sich. [7], [8], [9], [10], [12], [14] Verwendet werden diverse Begriffe, um dieses Phänomen genauer zu definieren wie z. B. informelle oder parallele Wirtschaft, Wrtschaft im Schatten, graue, schwarze, illegale oder Untergrundwirtschaft, unregistrierte oder nicht angemeldete Geschäftsaktivitäten etc., die mehr oder weniger bedeutungsverwandt bzw. -ähnlich sind.

Der Begriff *Schattenwirtschaft* ist mit vielen wissenschaftlich-theoretischen Diskussionen verbunden, die in den wissenschaftlichen Kreisen durchgeführt werden, jedoch beruht diese Erscheinung der modernen Wirtschaft auf zwei führenden zusammenhängenden Merkmalen: **erstens,** inwieweit eine wirtschaftliche Tätigkeit legal ist und von legalen juristischen Personen durchgeführt wird; und **zweitens,** inwieweit eine wirtschaftliche Tätigkeit und ihre Ergebnisse von den offiziellen Statistiken und den Informationssystemen staatlicher Institutionen erfasst und berücksichtigt worden sind.

Das erste Merkmal basiert auf der Verwendung von Begriffen wie gesetzwidrige, illegale, schwarze Untergrundwirtschaft. In diesem Sinne bedeutet der Begriff *Schattenwirtschaft* Herstellung von Waren und Angebot von Dienstleistungen, die durch das Gesetz nicht zulässig sind und/oder von illegalen Wirtschaftssubjekten ausgeführt werden. Dazu gehören z. B. Drogen- und Warenschmuggel, Menschenhandel, Verkauf von gestohlenen Waren und andere Arten krimineller Wirtschaftsaktivitäten.

Das zweite Merkmal bezieht sich auf Definitionen wie nicht angemeldete, unregistrierte, nicht anerkannte, informelle Wirtschaft. Nach diesem Merkmal deckt die Schattenwirtschaft die Geschäftstätigkeit und ihre Ergebnisse ab, die von den dafür befugten Institutionen wie statistische Ämter, Steuer- und Zollverwaltung, Kommunen etc. nicht erfasst bzw. registriert und berücksichtigt werden. Dazu gehören z. B. die Schwarzarbeit, bei der die Abeitskräfte nicht angemeldet sind und/oder keine Vergütung bekommen (illegale Beschäftigung), Umsatz- und Steuerhinterziehung, keine Entrichtung von Beiträgen für Sozial- und Krankenversicherung, Arbeit in privaten Haushalten etc.

Aufgrund dieser Ausprägungen lassen sich vier Wirtschaftsbereiche unterscheiden:

- *die offizielle Wirtschaft* (sie ist legal und rechenschaftspflichtig);
- *die informelle Wirtschaft* (sie ist legal und nicht rechenschaftspflichtig), dazu gehören die Ausübung handwerklicher und landwirtschaftlicher Tätigkeiten ohne Meisterbrief oder Gewerbeanmeldung, Arbeit in privaten Haushalten, freiwillige ehrenamtliche Arbeit und nachbarschaftliche Gegenleistungen, die keine Pflichtanmeldung erfordern;
- *die illegale (schwarze) Wirtschaft* (illegal und unerfasst), in der entweder gesetzlich verbotene Waren hergestellt oder illegale Tätigkeiten durchgeführt werden, oder die wirtschaftliche Tätigkeit wird von illegalen Wirtschaftssubjekten durchgeführt;
- *nicht angemeldete (graue) Wirtschaft,* in der legale juristische Personen einen Teil deren Geschäftstätigkeit, ihre Einkünfte und die Zahl der Beschäftigten rechtswidrig verbergen. Es handelt sich auch um Steuerhinterziehung sowie Nichtzahlung von Sozial- und Krankenversicherungen etc.

Die Tatsache, dass die Aktivitäten, die zum Begriff *Schattenwirtschaft* gehören, nicht offiziell gemeldet werden, bedeutet nicht, dass sie keinen Forschungsgegenstand sein sollten. Innerhalb jedes der drei Bereiche der Schattenwirtschaft sind Methoden entwickelt, um verschiedene Aspekte der nicht gemeldeten Unternehmen zu überwachen und zu beurteilen. Es ist zweckmäßig, dass die systematische Überwachung der Schattenwirtschaft alle ihre Kernsphären (informelle, nicht angemeldete, illegale) sowie die einzelnen Aspekte so weit wie möglich umfasst. Dies würde die Gesamtbeurteilung der Wirtschaftsaktivitäten in einem Land, den Umfang und den Anteil der Schattenwirtschaft und ihrer einzelnen Bereiche ermöglichen. Die Messung oder die Bewertung der Schattenwirtschaft kann durch diverse Methoden erfolgen, die die individuellen Merkmale der Schattenwirtschaft in unterschiedlichem Maße widerspiegeln. Das Endergebnis dieser Bewertung wird in der Regel als relativer Anteil am BIP dargestellt. Diese Methoden der Erfassung der schattenwirtschaftlichen Aktivitäten sind aber kein Gegenstand des vorliegenden Beitrages, da sie seinen Rahmen sprengen würden.

Das Nationale Statistische Institut in Bulgarien führt weitere Berechnungen in Bezug auf die Schattenwirtschaft durch, indem die Gründe für die unausreichende Vollständigkeit der offiziellen Berichte berücksichtigt und dem BIP hinzugefügt werden.

Unter Berücksichtigung der nicht angemeldeten (grauen) Wirtschaft fällt die Gruppe der Wirtschaftssubjekte, bei denen eine mutwillige Steuerhinterziehung und keine Anmeldung der tatsächlichen Beschäftigung und der real erbrachten Leistungen vorhanden sind. [7], [8] Dieser Umstand ist von großer Bedeutung für die Nachberechnungen des Anteils der Schattenwirtschaft am BIP.

Die Einschränkung der Schattenwirtschaft wird zum zentralen Thema der Diskussionen im öffentlichen Bereich bei Veränderungen in der Wirtschaftsgesetzgebung und vor allem vor Wahlen sowie bei der Billigung des Staatshaushalts.

Doch außerhalb des jährlichen Hidden Economy Index, der vom Zentrum für das Studium der Demokratie und sporadischen Analysen anderer nichtstaatlicher, akademischer und Unternehmensorganisationen präsentiert wird, gibt es kein adäquates Regierungssystem für Ex-ante und Ex-post-Bewertung der Auswirkungen der vorgeschlagenen Maßnahmen auf die verborgene Wirtschaft. Die kontinuierlich

öffentlich erklärten Bemühungen, der verborgenen Wirtschaft durch alle Kontrollstellen entgegenzuwirken, und die jährliche Offenlegung von Zehntausenden von Verletzungen der Verordnungen im Land im Zusammenhang mit der Geschäftsverbergung sind ein markantes Beispiel für das Ausmaß des Problems der Wirtschaftspolitik in Bulgarien. Die aufdringliche Wiederholung von Kontrollhandlungen und der Mangel an langwieriger Wirkung von ihnen deuten darauf hin, dass alternative Lösungen gesucht werden müssen.

Schätzungen des Bruttoinlandsprodukts des Nationalen Statistischen Instituts zum Zweck der Entdeckung der verborgenen Wirtschaft im Rahmen des Volkswirtschaftlichen Gesamtrechnungssystems werden der Öffentlichkeit nicht systematisch offengelegt und dienen leider nicht als Ausgangspunkt für die Entscheidungsfindung in der öffentlichen Ordnung. Die Situation der versteckten Wirtschaft ist in der gesamten Europäischen Union ähnlich. Eurostat unterstützt keine vergleichenden Informationen über die Neubewertung der verborgenen Wirtschaft im BIP. Obwohl die Methodik für die Berechnung des BIP in der Europäischen union standardisiert ist, gibt es noch relativ große Unterschiede in der Art und Weise, in der eine nicht nachweisbare Wirtschaft als versteckte, bzw. Schattenwirtschaft bewertet wird.

Nach dem Beitritt Bulgariens in die EU begannen mehrere Großprojekte zur Forschung und Prävention der verborgenen Wirtschaft in Bulgarien, die von den national anerkannten Arbeitgeberverbänden und Gewerkschaften verwaltet wurden. Sie haben das Potenzial, die Information über die Erscheinungsformen und die Auswirkungen der versteckten Wirtschaft auf die Unternehmen und die anderen Wirtschaftssubjekte deutlich zu verbessern. Um jedoch ihre Erscheinungsformen zu reduzieren und zu begrenzen, wird eine viel effektivere Zusammenarbeit mit staatlichen Strukturen und neuen Ideen erforderlich sein, um auf nationaler Ebene angemessene Gegenmaßnahmen verabschiedet werden können.

Die Schattenwirtschaft wird im Zuge der Zeit immer wieder als ein wichtiger Einflussfaktor für die wirtschaftliche und die gesellschaftliche Entwicklung Bulgariens angesehen. Gründe für das Ausmaß und Auswirkungen der Schattenwirtschaft sind unter anderem auch die Qualität der staatlichen Verwaltung in den Institutionen, sowie das Ausmaß der Korruption und der Grad an ökonomischer Freiheit. Die vordringlichste Aufgabe der Regierungspolitik ist die Bekämpfung der Korruption, die am besten durch mehr Transparenz und Liberalisierung erreicht werden kann.

IV. Das E-Government – eine Voraussetzung für transparente und faire Staatsverwaltung

Das strategische Dokument zur Modernisierung der bulgarischen staatlichen Verwaltung und zum besseren Funktionieren der Justiz und der Strukturen der Zivilgesellschaft im Zeitraum 2007-2013 war das Operative Programm *Verwaltungskapazität*. Es wurde durch den Europäischen Sozialfonds (ESF) und durch den nationalen Haushalt finanziert. An seiner Stelle wurden das operationelle Programm *Gute Verwaltung* für die Förderperiode 2014-2020, sowie die Strategie für die Entwicklung der E-Verwaltung in der Republik Bulgarien 2014-2020 konzipiert.

Die Grundtendenzen in der Entwicklung der staatlichen Verwaltung sind mit der Etablierung von den *Prinzipien der Transparenz, der Offenheit und der Rechenschaftspflicht* als Voraussetzungen für eine gute Verwaltung verbunden. Die Maßnahmen in dieser Richtung liegen der *Strategie für transparente Verwaltung, Prävention und Bekämpfung der Korruption,* 2006-2008 und dem *Programm für Transparenz in der staatlichen Verwaltung und in der Tätigkeit der Personen im höheren Dienst,* erarbeitet im Jahr 2006, zugrunde.

Unterschiedlich sind die Mittel und die Instrumente, wodurch sich eine höhere Transparenz und eine bessere Rechenschaft in der Verwaltung erreichen lassen. Dazu zählen das *Verwaltungsregister* und die *Jahresberichte über den aktuellen Stand der Verwaltung*, andere sind im Bereich der *besseren Kundenbedienung*, von *E-Government oder der Verwaltung der Humanressourcen.*

Die Veröffentlichung von *Erklärungen über die Einkommens- und Vermögesverhältnisse der Personen im höheren Dienst* im Internet ist eine andere Möglichkeit zur besseren Transparenz und Offenheit. Die fehlerhaften Erklärungen werden an die Nationale Einkommensagentur zur Überprüfung weitergeleitet. Die Erklärungen und alle relevanten Dokumente werden auf der Internetseite des Rechnungshofs bekannt gegeben. Jeder Staatsbeamte bzw. jede Beamtin ist verpflichtet, bis zum 31. März jedes Jahres seine/ihre Vermögensverhältnisse und den Interessenkonflikt bei der Anstellungsbehörde offenzulegen.

Zum Dezember 2006 zeigt der Bericht bezüglich der Durchführung des *Programms für Transparenz in der staatlichen Verwaltung und in der Tätigkeit der Personen im höheren Dienst,* dass die Mechanismen für Rückinformation und Bekanntgabe von Korruptionsfällen zu den wesentlichen *Instrumenten für Transparenz und Offenheit* in den einzelnen Verwaltungsstrukturen gehören bespielsweise *E-Mail-Adressen und Hotlines* (82%), *Feedbackkasten/Beschwerdekasten (78%), Ethikkodizes* (78%). Relativ wenig populärer sind die *Kundenbefragungen hinsichtlich der Verwaltungsdienst-leistungen* (46%). [11]

Zu mehr Transparenz und Offenheit trägt das *Gesetz über den Zugang zur öffentlichen Information* (GZöI) bei. Die Verwaltung ergreift immer mehr Maßnahmen zur Sicherstellung von maximaler Offenheit und Publizität in deren Tätigkeit (z. B. Work-Shops mit Journalisten, regelmäßige Pressekonferenzen, aktuelle Information auf den Internetseiten der Ministerien und Behörden, Veranstaltung von Aufklärungskampagnen, Erarbeitung von Kommunika-tionsstrategien).

Trotz der schon zur Verfügung gestellten Mechanismen für Rückinformationen führt die unzureichende Informiertheit der Öffentlichkeit über ihr Vorhandensein zu deren Ineffizienz. Eine Schwäche bleibt auch die oberflächliche Analyse der erhaltenen Beschwerden, Meinungen und Empfehlungen.

Besonders wichtig ist es, dass die Staatsbürger Möglichkeiten für einen besseren Zugagng zur öffentlichen Information bekommen, wofür eine Stärkung der Informationskapazitäten erforderlich ist, d. h. Überprüfung der Angemessenheit der gültigen Dienstordnung und deren Einführung in mehrere Verwaltungsstrukturen, intensivere Aus- und Weiterbildungen der Beamten hinsichtlich der Arbeit laut GZöI, Verbesserung der Möglichkeiten zur Antragsstellung auf elektronischem Wege. Neben dem Zur-Verfügung-Stellen von öffentlicher Information laut GZöI ist die Tendenz zur Sicherstellung von maximaler Offenheit über die Tätigkeit der Verwaltung fortzusetzen.

In Zusammenhang mit den eingegangenen Engagements für die Modernisierung der öffentlichen Verwaltung wird in Bulgarien eine konsequente Politik geführt. Die Grundlage dafür ist der gesetzliche Rahmen, der schon vorhanden ist. Trotzdem sei es zu erwähnen, dass Bulgarien bei der Umsetzung des E-Governments im Vergleich zu den anderen EU-Staaten *relativ zurückgeblieben* ist. Die Analyse der Gründe dafür erlaubt das Ergreifen einiger Maßnahmen, die

gewährleisten werden, dass die europäischen Anforderungen erfüllt werden.

Bis jetzt sind vier Hauptgründe für die Mängel bei der Umsetzung des E-Governments in Bulgarien festgestellt:

- Keine ausreichend effektiv funktionierende gesetzliche Grundlage;
- Fehlen einer operativen Kompatibilität der Informationssysteme in der Verwaltung;
- Mangel an einem vollwertigen elektronischen Austausch unter den einzelnen Verwaltungsbehörden;
- Ungelöste Probleme mit der Datenvereinheitlichung.

V. SCHLUSSFOLGERUNGEN

1. Die schattenwirtschaftlichen Aktivitäten sind nicht nur bei Privatpersonen und im Kleingewerbe üblich, sondern in praktisch allen Unternehmen verschiedener Größe und Eigentumsform gang und gäbe. Der tatsächliche Warenumsatz wird dabei nicht angegeben, bzw. seine Zusammensetzung wird durch gefälschte Angaben der Kostenproportionen verzerrt dargestellt. Dadurch wird der verborgene Bargeldumlauf ausgeweitet.

2. Die wichtigsten Gründe, weswegen einige Unternehmen schattenwirtschaftlich tätig werden, sind: administrativ-organisatorische Hindernisse vor der Entwicklung des Unternehmertums; unzureichender Schutz der Eigentumsrechte und der *Corporate-governance*-Mechanismen. Das informelle Verhalten der Unternehmen stützt sich auf eine in der Wirtschaft vorhandene Form von Wirtschaftsoperationen: Geschäfte zur Umwandlung von Buchgeld in Bargeld und umgekehrt. Bei diesen Operationen werden fiktive Geschäfte getätigt, an denen Scheinfirmen beteiligt sind, die das Risiko illegaler Tätigkeit auf sich nehmen. Auf diese Weise nimmt die legale Geschäftswelt am verborgenen Bargeldkreislauf teil.

3. Die praktischen Anreize, Wirtschaftsaktivitäten zu verbergen, liegen in den Vorteilen, das Einkommensniveau der Beschäftigten auf diese Weise zu wahren und die Organisation und Betriebsfähigkeit des eigenen Unternehmens dadurch zu gewährleisten, dass die Beziehungen zu Geschäftspartnern und Staatsorganen aufrechterhalten werden. Die hauptsächlichen Probleme, auf die die schattenwirtschaftlich tätigen Unternehmen stoßen, sind das Risiko, die übermäßige Aufmerksamkeit der Kontrollorgane auf sich zu ziehen, Beschränkungen beim Zugang zur staatlichen Unterstützung zu erfahren sowie in Schwierigkeiten bei der firmeninternen Buchführung und des Controllings zu geraten.

4. Die Folgen, die die Verbreitung schattenwirtschaftlicher Aktivitäten selbst im legalen Bereich für Unternehmen, Branchen und Wirtschaft insgesamt hat, sind vielschichtig. Die Symbiose legaler und informeller Aktivitäten innerhalb einer Firma wirkt sich auf die Prozesse auf der Mikroebene aus: Das wirtschaftliche Anreiz- und Verhaltenssystem der Unternehmen, der Beschäftigten und der Eigentümer wird ebenso verzerrt wie das Beziehungssystem zwischen den Marktsubjekten. Eigentumsrechte werden verletzt, die Neigung zum Sparen und Investieren sinkt.

5. Die Bilanz der Vor- und Nachteile informeller Aktivitäten legaler Unternehmen ist für die bulgarische Wirtschaft nicht eindeutig. Die damit einhergehende institutionelle Entwicklung ist ebenfalls ungünstig. Die schattenwirtschaftliche Aktivität fördert Korruption und Kriminalisierung der Wirtschaft. Das System der Geschäftsbeziehungen stützt sich vorwiegend auf persönliche Beziehungen und informelle Vertragsabschlüsse. Die Geschäftsethik läßt nichtstaatliche, auch gewaltsame Mechanismen des Schutzes von Eigentumsrechten und Verträgen zu. Eine Folge davon ist, dass die Steuerhinterziehung zu einer festen Verhaltensnorm und damit zu einer instutionellen Falle wird.

6. Es gibt bereits ein umfassendes Konzept für die Entwicklung von E-Government in Bulgarien. Allerdings gibt es im Prozess der E-Government-Implementierung im Vergleich zu den anderen EU-Länder eine relative Verzögerung. Eine Analyse der Ursachen ermöglicht die Identifizierung von Maßnahmen, um sie auf eine Ebene zu entwickeln, die den europäischen Anforderungen entspricht. Bisher konnten vier Hauptgründe für die unzureichende Umsetzung von E-Government in Bulgarien identifiziert werden: das Fehlen einer angemessenen regulatorischen Basis; die mangelnde Interoperabilität der Informationssysteme in der Verwaltung; der Mangel an vollem elektronischen Austausch auf abteilungsübergreifender Ebene und das ungelöstes Problem der Datenvereinigung.

LITERATURVERZEICHNIS

[1] Bericht über die Durchführung der Nationalen Strategie zum Entgegenwirken der Korruption im Zeitraum 2004 – 2005 http://www.acc.government.bg/documents/otchet-2004-1.doc

[2] Beschluss des Ausschusses vom 13.12.2006 zur Entwicklung eines Mechanismus für Zusammenarbeit und Überprüfung des Fortschritts Bulgariens in besonders wichtigen Bereichen der Justizreform und in der Bekämpfung der Korruption und der organisierten Kriminalität, EG (2006/929), OB L 354/56, 14.12.2006.

[3] Bericht über das Programm für Transparenz in der Tätigkeit der staatlichen Verwalung und der Personen im höheren Dienst, zum Dezember 2006, S. 14, http://www.mdaar.government.bg/lucidity.php

[4] Bericht über den Zustand der Verwaltung, 2006, S. 156 http://www.mdaar.government.bg/docs/Annual%20Report%20%202006_26.07.2007.pdf

[5] Berciht der Europäischen Komission über den Fortschritt Bulgariens in Begleitmaßnahmen nach dem Beitritt in die EU, 27.06.2007, S. 19 http://www.evropa.bg/bg/del/info-pad/press-releases/doklad.html

[6] Datenblatt Bulgarien, Botschaft der Bundesrepublik Deutschland, Stand 09.08.2017

[7] Enste, Dominik H., „Schwarzarbeit und Schattenwirtschaft“, *Wirtschaftsdienst*, 2012/2, S. 136-138;

[8] Enste, Dominik E., „Ursachen der Schattenwirtschaft in den OECD-Staaten“, *iw-trends*, 4/2003, S. 1-19

[9] Enste, Dominik H., Schneider, F., „Welchen Umfang haben Schattenwirtschaft und Schwarzarbeit? Ein Versuch zur Lösung des Rätsels“, *Wirtschaftsdienst* 2006/3, S. 185-191;

[10] Gancheva, A, u.a., *Die verborgene Wirtschaft in Bulgarien* (auf Bulgarisch), Zentrum für Demokratieforschung, Sofia, 2004

[11] Programm für Transparenz in der Staatsverwaltung und in der Tätigkeit der Personen im höheren Dienst http://www.mdaar.government.bg/programs.php

[12] Schneider, F., Enste, D.H., “Shadow Economies:Siye, Causes and Consequences”, *Journal of Economic Literature*, Vol.XXXVIII (March, 2000), pp.77-114;

[13] Strategie für transparente Verwaltung und Korruptionsprävenz und –bekämpfung für die Zeitperiode 2006-2008; http://www.strategy.bg/StrategicDocuments/View.aspx?lang=bg-BG&Id=414

[14] Thießen, Ulrich, „Schattenwirtschaft: Vorsicht vor hohen Marktschätzungen“, *Wirtschaftsdienst,* 2011/3, S. 194-201;

Using Multilink Type Suspension as a Front Suspension For a Motorcycle

Verwendung Eines Multilink als Vorderradaufhängung für ein Motorrad

Georgi Yanachkov*, Simona Hesapchieva†
Department of internal combustion engines, automobiles and transport, Technical University of Sofia,
Sofia, Bulgaria, *) gyanachkov@tu-sofia.bg; †) Simona_Hesapchieva@abv.bg

Abstract **—** In the paper are investigated the possibilities to use a Multilink suspension as a motorcycle front suspension and how the common geometric parameters of the front wheel suspension are represented. One of the possibile variants is shown.

Zusammenfassung **—** In der Artikel werden die Möglichkeiten zur Verwendung einer Multilink-Aufhängung als Motorrad-Vorderradaufhängung untersucht und wie die gemeinsamen geometrischen Parameter der Vorderradaufhängung dargestellt sind. Eine der möglichen Varianten wird gezeigt.

I. INTRODUCTION

In modern motorcycles the most widely used type of suspension is the telescopic fork. Despite its advantages such as simplicity of construction, low cost and easy maintenance, this front suspension also has drawbacks: change of the trail when the suspension travels and increased tendency to "dive" the motorcycle when braking. Some companies offer different types of front suspension designs that somewhat compensate for the above-mentioned drawbacks. Examples include the BMW Motorrad with Telelever and Duolever suspension similar to the trapezoidal suspension for a car and the Yamaha with a one-arm front suspension similar to the McPherson suspension

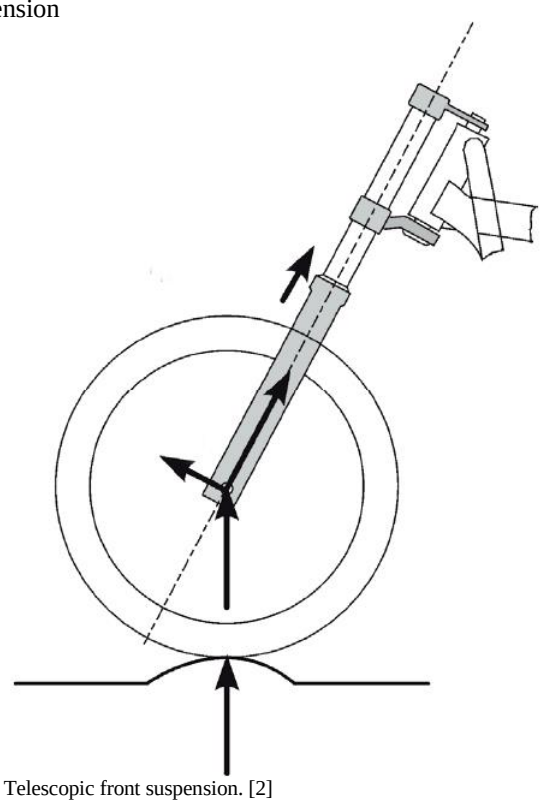


Fig. 1. Telescopic front suspension. [2]

Fig. 2. Telelever suspension type. (BMW Motorrad)

Fig. 3. Duolever suspension. (BMW Motorrad)

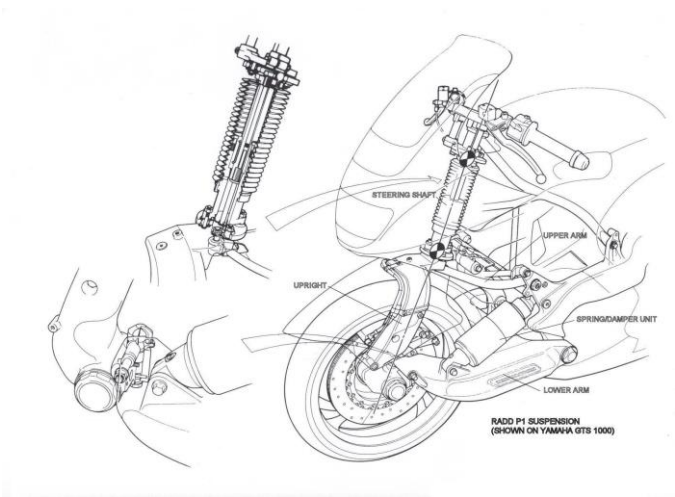


Fig. 4. Front suspension Yamaha GTS 1000.

II. Basic Geometric Characteristics of the Front Suspension of the Motorcycle

In Fig. 5 are shown the basic parameters of the front suspension of a motorcycle. The inscriptions are as follows:

- wheelbase L - 1350 - 1600 mm.
- rake angle α - 61 – 64°.
- trail c - 80 - 125 mm.
- height of the center of masses h – around 600 mm.

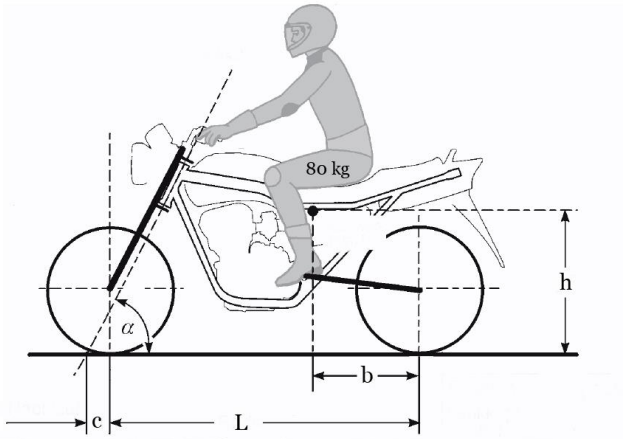


Fig. 5. Basic parameters of the motorcycle suspension [2].

While the motorcycle travels on a road the suspension moves and his geometric characteristics change. When driving on a flat road, the change is not very important due to the small shift only from the road bumps, but in the case of intensive braking, the changes have a significant influence on the suspension parameters, hence on the stability and controllability of the motorcycle.

Fig. 6 shows the change in the geometric parameters of a front suspension of a motorcycle with a telescopic fork while braking. It can be seen that the wheelbasebase and the trail are reduced and the rike angle increases. This reduces the stability of the motorcycle. In some extreme cases, for example in the case of intensive braking, this effect allows the driver to perform an extreme turn to overcome an obstacle, but overall the change in geometric parameters is considered a drawback. [4,5]

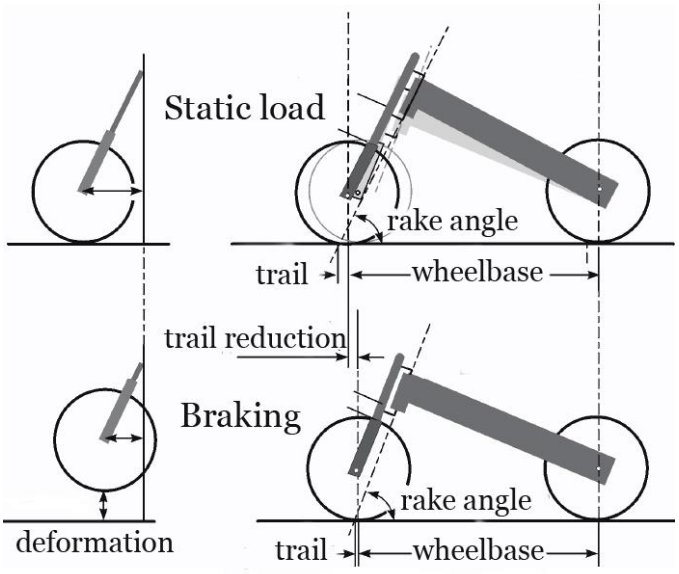


Fig. 6. Changes in the geometric characteristics of a motorcycle suspension with a telescopic fork.

III. Multilink Front Suspension for Motorcycle.

Multilink suspension is common in cars, primarily for high-end and sport cars. This type of suspension has not been seen so far as a front suspension for a motorcycle and one of the possible options for its application is shown below.

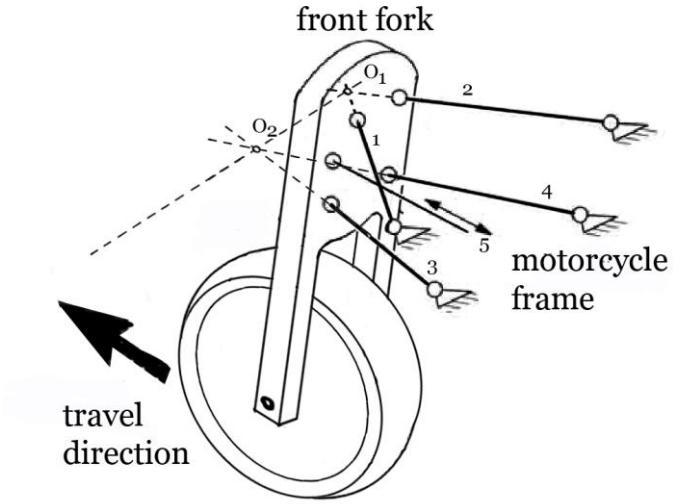


Fig. 7. Basic front Multilink suspension scheme for motorcycle.

Fig. 7 shows a general front Multilink suspension scheme for a motorcycle. This suspension is similar to trapezoidal suspension. The front wheel is hinged on the front fork. It is connected to the frame of the motorcycle through four carriers. The shoulders 1 and 2 form the upper carrier and, according to the four-seater rule, the instant center of rotation of the fork against the two arms is at point O1. Shoulders 3 and 4 form the lower carrier and the momentary center of rotation is point O2. Thus points O1 and O2 form the virtual axis of rotation of the fork. The arm 5 is a control element for the suspension and is connected to the steering handlebar of the motorcycle. In the absence of movement, the whole system has a degree of freedom and the only movement in it is in the vertical direction. Then there is no turning in the fork and the motorcycle moves in a straight line. When the arm 5 moves along the axis, then the fork rotates with respect to the virtual

axis of rotation and at the same time moves along the horizontal axis. The control arm 5 may be either directly connected to the motorcycle's handlebars or be made as two triangular jaws, one for the fork and one for the motorcycle handle (Fig. 8).

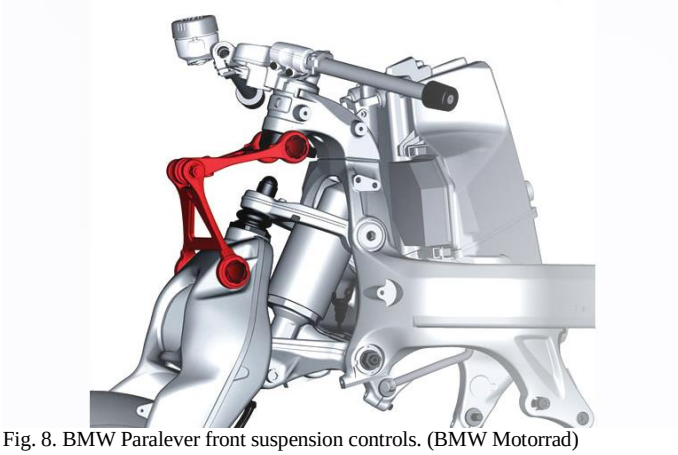


Fig. 8. BMW Paralever front suspension controls. (BMW Motorrad)

IV. Advantages and Disadvantages.

Advantages:

- By choosing the size of the levers in the suspension, it is possible to achieve constant geometrical characteristics of the suspension during its compression and relaxation.
- When changing the longitudinal slope of the upper and lower levers, can be achieved a desired anti-dive effect during braking.
- When changing the transverse slope of the levers, it is possible to achieve different momentary points of rotation of the fork, hence the inclination of the fork rotation axis and its stabbing in the road, which will give the trail to the front suspension.

Disadvantages:

- More complex construction.
- Multiple hinged connections (10 and more).
- Higher mass of unsprung masses.

V. Conclusions.

Multi-arm suspension meets the geometric characteristics of the front suspension of a motorcycle can be used as such. Detailed kinematic and dynamic research will be done in subsequent publications.

References

[1] M. Trzesniowski, *Rennwagentechnik*, Vieweg+Teubner Verlag Wiesbaden , Deutschland 2010.

[2] J. Stoffregen, *Motorradtechnik*, Vieweg&Sohn verlag Deutcshland 2006.

[3] K. Code, *A Twist of the Wrist Vol. 2*, E-Book, California, USA 1993

[4] L. Kunchev, K. Kosev, *Motorcycles, sport- and racecars*, TU-Sofia, Sofia 2017

[5] I. Evtimov, R. Ivanov, *Motorcycles*, Technical University Ruse, Ruse 2011

Grouping of technological features using a clustering algorithm

Gruppierung von technologischen Merkmalen unter Verwendung eines Clustering-Algorithmus

Galina Nikolcheva, Julieta Kaleicheva
Faculty of Industrial Technology, Technical University of Sofia, Sofia, Bulgaria, e-mail: ginic@tu-sofia.bg, jkaleich@tu-sofia.bg

Abstract — The work presents a clustering algorithm that groups technological features for the generation of setups. The features processed in each setup have a similar technological process. The clustering algorithm - COBWEB is used for their grouping. It is an incremental conceptual clustering algorithm that groups clusters into interacting features at each setup. A variant method is used to create the technological process.

Zusammenfassung **—** Die Arbeit präsentiert einen Clustering-Algorithmus, der technologische Features für die Generierung von Setups gruppiert. Die in jedem Setup verarbeiteten Funktionen haben einen ähnlichen technologischen Prozess. Der Clustering- Algorithmus - COBWEB wird für ihre Gruppierung verwendet. Es ist ein inkrementeller konzeptioneller Clustering-Algorithmus, der Cluster in interagierende Features bei jedem Setup gruppiert. Eine Variantenmethode wird verwendet, um den technologischen Prozess zu schaffen.

I. Introduction

In the past decades manufacturing industries are undergoing a transition from traditional methods to advanced manufacturing technologies, a many of which are computer based. The manufacturing research community has focused on developing and improving technologies such as CAD/CAM and Computer- aided process planning (CAPP). The features are the link between design and manufacturing in a CIMS environment. In recent years have seen many works on automated process planning and connected with that determining the number of setups, the sequence of making features in each setup [8,11, 14,17].

The sequence of making the features depends on feature interacting such as the geometric relationships between them are the reason for theirs interacting. The feature types are also used to identify the type of interaction and thus the applicable rules that govern their removal. Technological features are two types: non-intersect (simple) and intersect (complicated) that are interacted to one another.

The goal of the paper is group the interacting technological features based on their geometrical relationships by using clustering algorithm so that the groups obtained to have similar process plan.

II. GROUPING THE FEATURES

The geometrical relationships between features influences the process planning. They are studied and defined by different approaches: fuzzy, neural network and expert systems [14, 15,16]. Most researchers have used the idea of feature interactions to express the problems that occur in the planning process when some elements are arranged in a given geometric orientation relative to another [3].The feature interactions express Hayes [7] looks for feature interactions when generating the fixturing, and the feature sequencing portions of the process plan. It uses special rules aimed at particular interactions, and avoids those interactions by reordering sequence of features, or putting features into different setups. The geometric relationships between features are used [10] to identify intersected features and the sequences in making them. To determine the sequence of making the features it is necessary to identify intersected features and to organize them into regular groups of patterns. These groups of features have a similar process plan. Chang [2] grouped feature clusters based on tool approach direction. He defined approach and feed directions used for setup generation. An approach direction is a straight path that gives a tool an unobstructed access to the feature in the workpiece. The features may have more than one approach direction. Based on approach direction of the tools, similar features are grouped for a setup. Different methods are used for grouping of various objects and the results are best by using clustering algorithms [2,3].

A. CLUSTERING ALGORITHMS TO GROUPING OF FEATURES

Traditional clustering algorithms adopt one of two primary approaches: agglomerative and divisive. Agglomerative methods repeatedly group objects and clusters of objects together, based on a changing similarity requirement, to form larger and larger clusters. To achieve this a distance metric which allows comparisons between objects and clusters is needed. Divisive methods do the reverse, subdividing large clusters into smaller and smaller ones. The results of both methods can be viewed as a hierarchical tree.

Most methods produce an exclusive partitioning, simultaneously taking a set of instances and placing those instances into disjoint clusters (solving the clustering problem). A separate task is the formulation of a description for each cluster (the categorization problem) [1]. This is the assignment of category labels to the clusters. A large body of older research solves the clustering problem using statistical methods, but leaves categorization unaddressed.

Concept of nodes representing descriptions of concepts and instances is constructed in this approach. The learning system typically adds one instance at a time, following a path deeper down the hierarchy in a decision tree manner until the new instance has been classified under an existing concept or a new clustering methods provide a natural approach to solving both clustering and categorization problems. These methods include the formation of a concept hierarchy using an incremental top

down classification scheme and thus perform unsupervised learning. A knowledge based structure consisting concept cluster has been formed. This approach is divisive as instances classified under previously formed concepts are further segregated when sub-generalizations form. These sub-generalizations may include instances previously classified under a more general concept.

In this work we have considered incremental conceptual clustering algorithms such as UNIMEM [9], ERAM [4], COBWEB[5], STAGGER[12], ID4 [13].

B. APPLICATIONS TO FEATURE GROUPING

The COBWEB system is an incremental conceptual clustering algorithm. The program is written by Raymond Joseph Mooney in 1991. The system forms classification trees that are intended to yield "good" prediction among many attributes and can be used for a wide variety of purposes.

An important difference between COBWEB and earlier conceptual clustering systems is that it is incremental – COBWEB integrates an observation into an existing classification tree by classifying the observation along a path of "best" matching nodes. Like ID4, probabilistic summaries of previous observations are stored at each node, but the matching functions and the criteria used for subtree revision differ considerably. COBWEB uses the category utility function [6] to guide classification and tree formation.

Category utility bases its evaluation on all of the observation's attribute – values rather than a single one, making COBWEB a polythetic classifier as opposed to a monothetic classifier e.g. ID4. Subtree revisions in COBWEB are triggered by considering prediction ability over all attributes, but concern for multiple attributes complicates subtree revision. In ID4 a subtree is simply deleted, but in COBWEB a deletion that benefits one attribute may be inappropriate for others. In response, the system identifies points in the tree for cost-effective prediction of individual attributes. These points are marked by default values that COBWEB dynamically maintains during incremental clustering.

C. COBWEB

We use COBWEB – to group interacting features of each setup in clusters, so that they all have a similar technological process achieved by variant machining process.

The following lists of structures of data for the use of COBWEB must be create:

- A list of attributes-names;
- A list of domains – value corresponding to attributes;
- A list of raw – examples, appropriate with upper structures.

The attributes-names are included in the first structure (Figure 1). These names are obtained in different way. The type of the model (model) and the type of the base (type_of_the_base_solid) are known. The type of interacting (nested and tangent), the geometry (type_of_the_second_feature) and the identical geometry form of the features (type) are obtained by analysis. The approach directions were described in the previous sections. Two additional attributes have been introduced showing the passing (type_of_the_interaction_1) in a direction perpendicular to the plane, where the approach direction lies and the passing (type_of_the_interaction_2) in a direction parallel to or coincidental with the plane where the approach direction lies.

Some values of attributes in second list are obtained from the feature classification. They are:

- For the attribute model – protrusion or depression;

```
(setf *feature-names* '(
      model
      nested
      tagent
      type_of_the_second_feature
      approach_direction
      type_of_the_base_solid
      type_of_the_interaction_1
      type_of_the_interaction_2
      type ))
```

Fig. 1. First structure: attributes-names

```
(setf *domains*
  '(  (protrusion depression )
  (no_nested  nested_simple  nested_complex)
  (no_tangent tangent_one_side
     tangent_two_sides tangent_three_sides)
  (cylindrical_hole prism_hole conical_hole)
  (on_one_side on_different_sides )
  (cylindric prismatic sphere )
  (through_horizontal blind_horizontal)
  (through_vertical   blind_vertical)
  (identical          different))
```

Fig. 2. Second structure: domains

- For the attribute base – geometrical type of the base – cylindric, prismatic and sphere;
- For the attribute geometrical type of feature - cylindrical_hole (cylindrical hole), prism_hole (pocket), conical_hole (hole with a cone-shaped bottom).
- The values of the remaining attributes such as nested, tangent,type_of_the_interaction_1,type_of_the_interacti on_2 and type - Figure 2 are obvious.

Graphic figures of the interacting features have been created for carrying out an experiment - Figure 3. These are the most commonly used features in machining the details. Feature interactions are used to reason about geometric relations between features [3]. The interacting features are grouped in clusters after implementation of COBWEB. Each cluster contains features with similar technological process. The procedure starts by function train, as its argument is the third list – Figure 4 (train raw-examples).

The third list is a result of the first two, as its structure includes the feature names and their values.

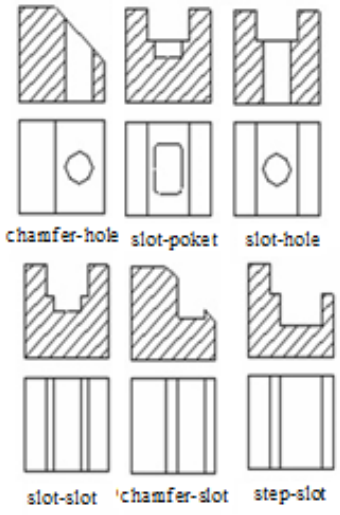


Fig. 3. Graphic figures of the interacting features

```
        setf *raw-examples* '((chamfer_hole  (depression   nested_simple  no_tangent
cylindrical_hole on_one_side  prismatic  blind_horizontal    through_vertical different   ))
                                (slot_pocket (depression  nested_simple     no_tangent
prismatic_hole on_one_side prismatic blind_horizontal    blind_vertical identical    ))
                              (slot_hole  (depression   nested_simple     no_tangent
  cylindrical_hole on_one_side prismatic  blind_horizontal    through_vertical different  ))
                             (slot_slot (depression   nested_simple     no_tangent
 prism_hole on_one_side prismatic through_horisontal  blind_vertical identical   ))
                             (chamfer_slot (depression  nested_simple  no_tangent  prism_hole
on_one_side prismatic through_horisontal  blind_vertical identical   ))

                                (step_slot (depression nested_simple no_tangent prism_hole
             on_one_side prismatic through_horisontal  blind_vertical identical    ))
```

Fig. 4. Train raw-examples

The result of COBWEB is two clusters, containing features with a similar technological process – the first cluster is

- C-37{STEP_SLOT CHAMFER_SLOT SLOT_SLOT SLOT_POCKET}

and the second one is

- C-38{SLOT_HOLE CHAMFER_HOLE}.

III. CONCLUSIONS

The utilized COBWEB system gives very good results for grouping of details that contain interacting features with a similar technological process. An increase in the number of instances leads to better precision in cluster forming. It may serve The classification and hierarchy structures may serve to generate setups based on machining operations.

REFERENCES

[1] J.R. Anderson, M. Matessa, "Exploration of an incremental Bayesian algorithm for categorization", *Machine Learning*; (9):1992, pp.275-308.

[2] T.C. Chang, *Expert Process Planning for Manufacturing, Addison-Welsey*, Reading, MA,1990.

[3] C.L.P. Chen, "Setup Generation and Feature Sequencing using an Unsupervised learning Algorithm", Proceedings 199, *NSF Design and Manufacturing Systems* 1993; 1,pp. 981-986.

[4] E. Feigenbaum, H. Simon,"EPAM-like models of recognition and learning", *Cognitive Science* 1984; 8, 305-336.

[5] D.H. Fisher, "Knowledge acquisition via incremental conceptual clustering", *Machine Lerning*, 1987; (2): 139- 172

[6] M.A Gluck, J.E. Corter, "Information, uncertainty, and the utility of categories*". Proceeding of the Seventh Annual Conference of Cognitive Science Society Irvine CA:Lawrence Elbaum* 1985; pp.283-287

[7] C.Hayes, ***"****Machineng planning: A Model of an Expert Level Planning Process*", Ph.D. Dissertation, Carnegie Mellon University, 1990.

[8] M.Hazarika, U. Dixit, "*Setup Planning for Machining*", Springer Chain Heidelberg, 2015

[9] M.Lebowitz, "Experiments with incremental concept formation: UNIMEM", *Machine Lerning* 1987; (2), pp.103-1138.

[10] S.K.Ong, A.Y.C. Nee, "Application of Fuzzy set Theory to setup planning", *Annals of the CIRP*, 1994, 43,1, pp.137-143.

[11] Y.H. Pao, K. Komeyli, D. Shei, S. LeClair, A.Winn, "The Episodal Associative Memory: Managing Manufacturing Information on the Basis of Similarity and Associativity", *Journal of Intelligent Manifacturing*,1993.

[12] J.C.Schlimmer, "*Concept acquisition through representational adjustment*", Unpublished doctoral dissertation. Irvine: University of California, Department of Information and Computer Science, 1987.

[13] J.C. Schlimmer, D. Fisher, "A case study of incremental adjustment of representations of learning". *Proceeding of the Fifth National Conference on Artificial Intelligence* 1986; Philadelphia, PA, 496-501.

[14] T.E. Westhoven, C.L.P.Chen,Y.H. Pao, S.R. LeClair, "The Episodal Associative Memory Approach for Sequencing Interactive Features in Process Planning", *AI for Eng., Design, Analy. And manufacturing*, 1992, 6, 3, pp.177-197.

[15] W. Wang, B.Subagdja, Ah-Hwee Tan, "A Self- Organizing Approach to Episodic Memory Modeling", 2010 IEEE *World Congress on Computational Intelligence* July, pp.18-23, CCIB, Barcelona, Spain, 2010.

[16] N. Xu, S.Huang, Y.Rong, "Automatic setup planning: current state-of-the-art and future perspective". *Int J Manuf Technol Manag* 2007,11(2), pp.193–208.

[17] Zhang G. Z, J. Han, Y. Y. Liu, "Approach to Machining Setup Process Planning of NC Machine Tools Based on Graph Theory", *Advanced Materials Research*, 2012,Vols. 542-543, pp. 1190-1195.

Application of UniSim Design to study the static and dynamic properties of a chemical technology equipment

Erforschung der statischen und dynamischen Eigenschaften einer chemischen Technologieausrüstung unter Anwendung von UniSim Design

Stefan Kiranov

Department of power engineering and automation, College of energetics and electronics at Technical university - Sofia, Sofia, Bulgaria, skiranov@tu-sofia.bg

***Abstract* —** The static and dynamic simulation of a crude oil distillation column was presented. The aim of this work is to study the properties of this object from control point of view. Investigations of the column model behavior were carried out to obtain the transitional processes and the static characteristics. The specialized software package *UniSim Design* was applied. The achieved results of simulation are presented and discussed.

Zusammenfassung - Die statische und dynamische Simulation einer Rohöldestillation ist presentiert. Das Ziel der Arbeit besteht darin, die Eigenschaften dieses Objektes, unter Steuerungsgesichtspunkt zu erforschen. Das Rohuntersuchungsmodell wurde erforscht um die Übergangsprozesse und die statische Eigenschaften des Objektes zu erhalten. Die spezialisierte Software *UniSim Design* wurde implementiert für die Simulation. Die erreichten Simulationsergebnisse sind presentiert und diskutiert.

I. Introduction

The development of the necessary automatic control systems for technological equipment is based on the knowledge of the dynamic and static properties of the controlled technological objects. For the purpose of analyzing the technological objects (TO), their description is done by mathematical models of the processes in them [5].

Modern software tools for computer modeling allow the compilation of sufficiently precise models. Some of the best known products in this area are Matlab of Mathworks, AspenTech of Aspen [6], [7], UniSim Design of Honeywell, Wolfram Mathematica. By them, the TO study is performed in a simulation mode instead of experimenting with the actual objects.

In the present work are offered results of modeling the statics and dynamics of oil atmospheric distillation column. The study is conducted with respect to its main input and output technological variables.

It is used Honeywell's specialized software package for modeling and simulation UniSim Design [2], [3]. It allows the simulation of different chemical technology units and installations in both static and dynamic mode.

II. Simulation of the dynamics of the object

The technological scheme (Fig. 1) of the surveyed object consists of one main column for atmospheric distillation of oil with 29 trays, condenser in its upper part and three small distillation columns (strippers). The input streams are petroleum (crude oil) and working steam, and the output streams include non-condensed gases, gasoline, kerosene, diesel, atmospheric gas oil (AGO) and water.

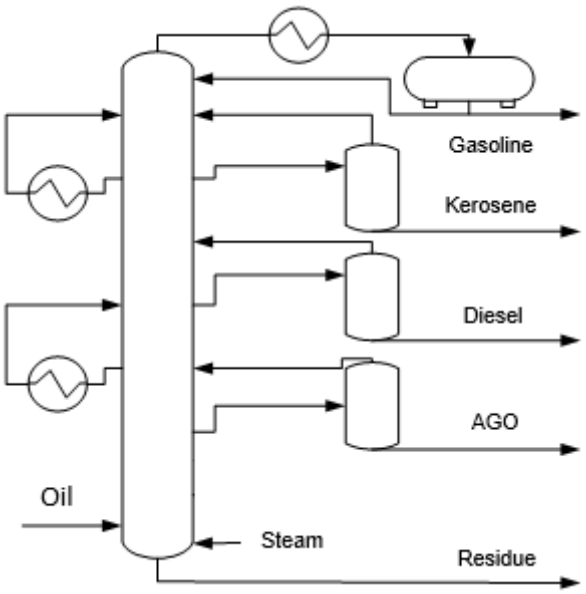


Fig. 1. Technological scheme of distillation column.

The block diagram of the object under consideration as part of the entire installation is shown in Fig. 2.

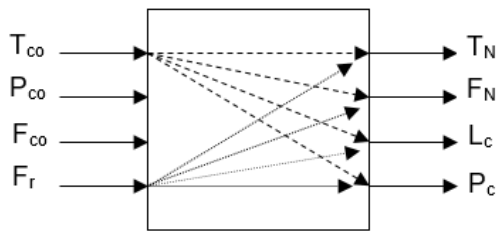


Fig. 2. Block diagram of the object.

The inputs covered by the study are: the temperature of the column feed T_{co}, its molar flow rate F_{co}, the reflux rate at the top of the column F_r and the pressure of the input raw material P_{co}. The outputs are: the temperature of the distillate (gasoline) T_N, the molar flow rate of the distillate F_N, the level of the condenser L_c and the pressure of the uncondensed gases in the condenser at the top of the column P_c.

Two groups of studies were performed in simulation mode.

The first group is studied the reaction of the output variables molar flow rate F_n and the temperature T_n of the gasoline and the level L_c in the condenser after applying a unit step change of the incoming oil temperature T_{co} at a constant value of the reflux rate F_r.

The other group is related to the reaction of the indicated output variables after a unit step change of the flow rate F_r at a constant temperature T_{co}. In both cases, the flow rate F_{co} and the pressure P_{co} are fixed values.

The first group of studies is carried out by a unit step change of T_{co} inlet temperature of 5 °C at a constant value of the reflux F_r. The investigations were conducted in the direction of increasing temperature in the range 322.6 - 347.6 °C. In the simulation, five transitional processes were captured for each of the channels considered in that range and were obtained six points of their static characteristics.

The resulting transitional processes are shown in Fig. 3, Fig. 4 and Fig. 5.

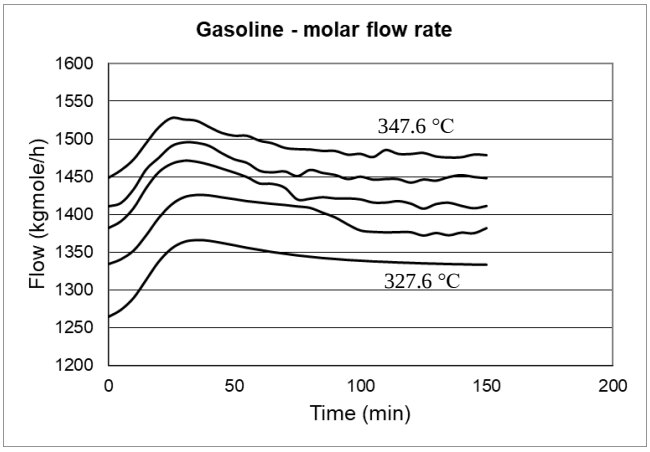


Fig. 3. Transitional processes at the channel inlet temperature – gasoline molar flow.

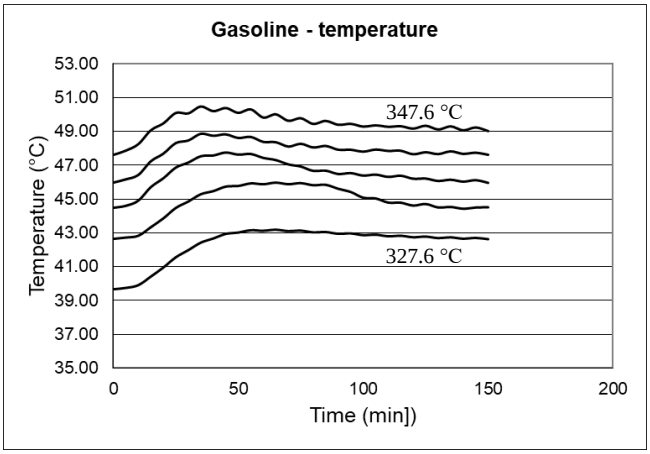


Fig. 4. Transitional processes at the channel inlet temperature – gasoline temperature.

The static characteristics of the channels under consideration are shown in Fig. 6, Fig. 7 and Fig. 8.

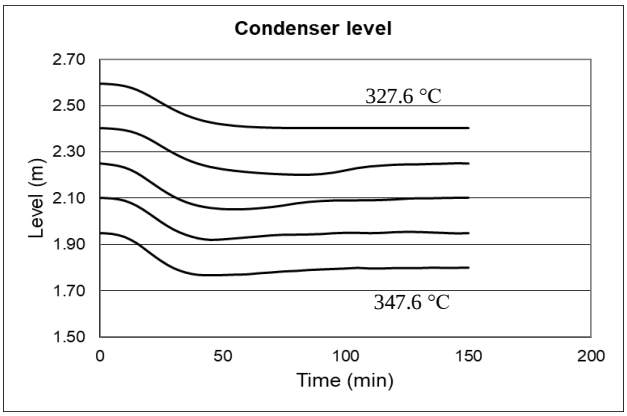


Fig. 5. Transitional processes at the channel inlet temperature – condenser level.

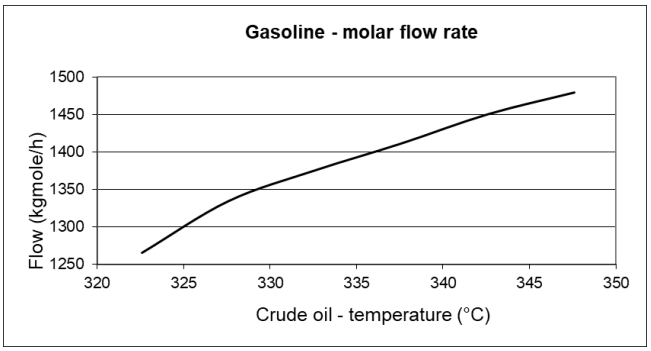


Fig. 6. Static characteristic at the channel crude oil temperature – gasoline molar flow.

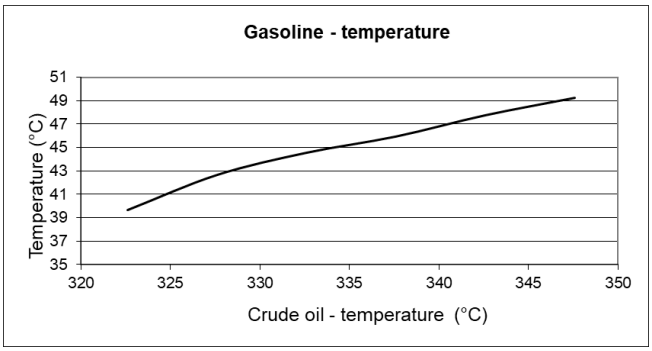


Fig. 7. Static characteristic at the channel crude oil temperature – gasoline temperature.

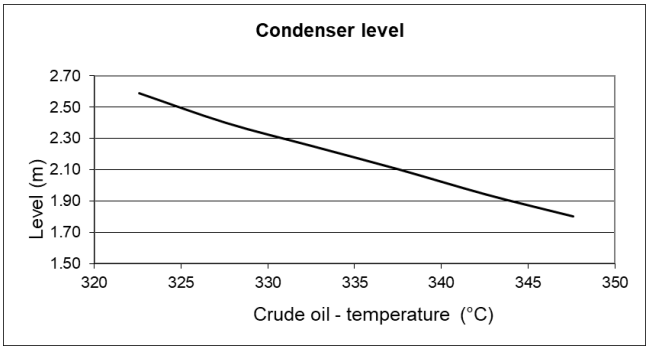


Fig. 8. Static characteristic at the channel crude oil temperature – condenser level.

In the second group of studies, a unit step change of reflux F_r of 100 kgmole / h in the direction of its increase in the range 900 - 1500 kgmole / h is applied. Six transitional processes were captured for each channel and were obtained seven points of their static characteristics. The transient characteristics are given in Fig. 9, Fig. 10 and Fig. 11.

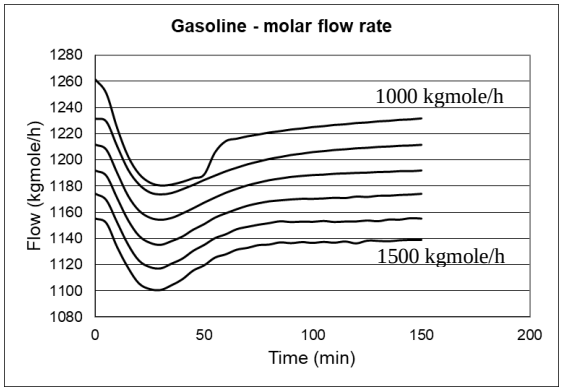


Fig. 9. Transitional processes at the channel reflux flow rate – gasoline flow rate.

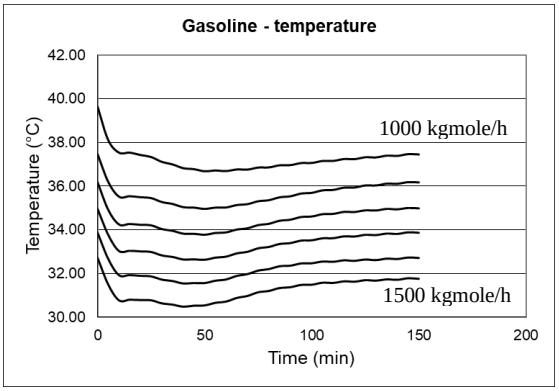


Fig. 10. Transitional processes at the channel reflux flow rate – gasoline temperature.

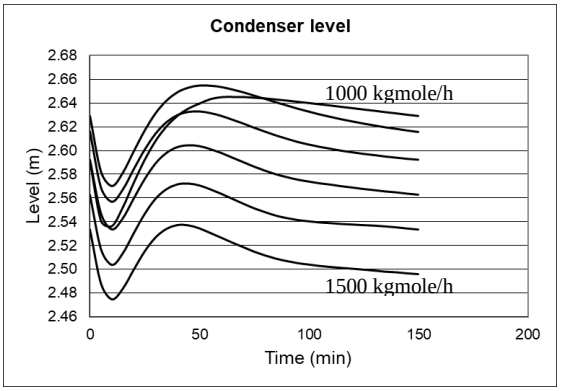


Fig. 11. Transitional processes at the channel reflux flow rate – condenser level.

The static characteristics are presented in Fig. 12, Fig. 13 and Fig. 14.

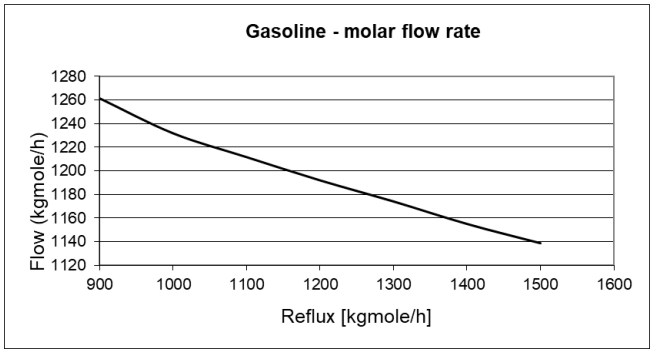


Fig. 12. Static characteristic at the channel reflux flow rate – gasoline flow rate.

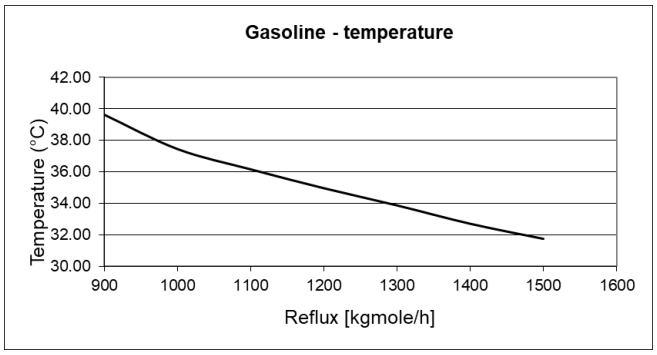


Fig. 13. Static characteristic at the channel reflux flow rate – gasoline temperature.

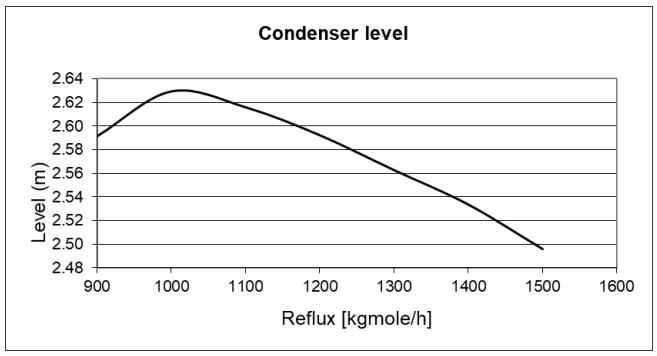


Fig. 14. Static characteristic at the channel reflux flow rate – condenser level.

III. Conclusion

The simulations performed with the model of real column show that the input variables - feedstock temperature and distillate flow rate in the recirculation loop (reflux) significantly affect all output variables - quantity and temperature of the distilled gasoline and the level of the distillate in the condenser.

Analysis of the data from the simulations gives grounds for the column to be considered as an inertial and non-linear object with significant interactions between its channels.

The dynamics of the object on the studied channels shows that the transition time is high (100-150 min).

The resulting static characteristics can be considered as linear in the small range around the work point. The graphs of these characteristics shown above expose that there is a static non-linearity in a wider operating range.

The study presents that the temperature of the inlet stream and the reflux flow significantly affect the amount of distillate removed. From the control point of view, the input temperature can be considered as a disturbing effect, and the flow in the recirculation circuit can be used for control purposes as a controlling action.

References

[1] M. Fahim, T. Sahhaf, A. Elkilani, *Fundamentals of petroleum refining*, Elsevier, 2010.

[2] Honeywell, *UniSim Design tutorials and applications*, Honeywell, 2010.

[3] Honeywell, *Dynamic modeling*, Honeywell, 2010.

[4] B. Liptak, *Instrument engineers' handbook: Process control and optimization*, CRC Press, 2006.

[5] D. Seborg, T. Edgar, D. Mellichamp, F. Doyle, *Process dynamics and control*, Wiley, 2011.

[6] W. Luyben, *Distillation Design and Control using Aspen Simulation*, Willey, 2006.

[7] J. Haydary, T. Pavlík, "Steady-state and dynamic simulation of crude oil distillation using Aspen plus and Aspen dynamics", *Petroleum and coal* 51(2) 100 - 109, 2009.

Authors Index

Achanov, Blagovest, 19

Demiröz, Serap, 57
Dimitrova, Konstantina, 23
Dimitrova, Ralitza, 53
Dimitrova, Virginia, 33
Dontschewa, Miglena, 81

Galabov, Vassil, 11, 27
Gankova–Ivanova, Zwetelina, 109, 117
Garabitov, Angel, 29
Georgieva, Vanesa, 15
Georgieva, Veska, 7
Greulich, Peter, 51
Gunkel, Gero Peter, 61

Hawlitschek, Anja, 3
Henrich, Dirk, 65
Hensel, Stefan, 41, 93
Hesapchieva, Simona, 121
Heupel, Thomas, 47, 57, 65, 69

Kaleicheva, Julieta, 97, 113, 124
Kambosev, Nikolay, 73
Kamenov, Ognyan, 101
Karaguiozova, Zdravka, 97, 113
Kiranov, Stefan, 127
Kirov, Boris, 27
Krenz, Till, 3

Mücke, Thomas, 77

Marinov, Marin, 41, 81, 93
Markova, Vanya, 37
Mihaylova, Antoniya, 7
Mishev, Valentin, 97, 113

Nikolcheva, Galina, 124
Nikolov, Kristian, 27
Nikolov, Petar, 11

Ovcharova, Jivka, 1

Petrov, Plamen, 7
Pfeiffer, Lucas, 45
Popowa, Julia, 45

Roth, Christian, 69

Sasho, Guergov, 85, 89
Schönauer, Sonja, 69
Scheicher, Thomas, 47
Shopov, Ventseslav, 37
Stamatov, Denise, 81

Todorov, Svetlozar, 109
Tsenov, Aleksandar, 85, 89

Uzunova, Magdalina, 101, 105

Yanachkov, Georgi, 121

Zug, Sebastian, 3