RCITD 2015

Proceedings in

Research Conference In Technical Disciplines

The 3rd International Virtual Research Conference In Technical Disciplines

19. - 23. October 2015

Slovak Republic

- All published papers undergone single blind peer review.
- All published papers are in English language only. Each paper had assigned 2 reviewers and each paper went through two-tier approval process.
- Papers are published as delivered by authors without content modification. All accepted papers have been formally checked by the conference Technical Committee and regularly reviewed in single blind peer reviewing process by the conference Scientific Committee (Reviewers Committee).
- The publication contents are the sole responsibility of the publication author and it does not reflect the Publishing Society, Zilina, Slovakia.

Open Access Online archive is available at: **http://www.rcitd.com/archive**

(proceedings will be available online one month after the publication release).

In case of any questions, notes or complaints, please contact us at: **info(at)rcitd.com**.

Paper Citation Example: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, Proceedings of The 3rd International Virtual Research Conference In Technical Disciplines (RCITD-2015), Vol. 3, pg #-#. Zilina, Slovakia.

Section Chairman Committee

Mariana Iancu

• About:

Prof. PhD, Director of Specialized Department with Psychopedagogical Profile, Associate researcher at Institute of Education Sciences. International Jury member of OIB Bruxelles, prof. PhD with officiale invitation at Faculty of Biology, University of Latvia, Riga. Counselor in career of students at European Projects with European Structural Founds. Member of Scientific Committee-Reviewers and Chairman Committee at numerous Internationals Conferences from Slovakia, Macedonia, speaker at international conferences. Author of numerous books and scientific articles about scientific research published in Romania, U.S.A., U.K., Slovakia, Belgium, Latvia and so on.

Educational background: PhD in Education Sciences-University Bucharest (2003) with thesis in Didactical Biology, degree in Biology-University Bucharest (1988), MA in Public Policies and European Integration-S.N.S.P.A. Bucharest (2009)

• Affiliation:

BIOTERRA University of Bucharest, Romania

• Scientific specialization:

Database systems, biology teaching methodology, biology teaching, pedagogy, the psychology of education, psychopedagogy, didactical communication, biology education, bioethical education, ecological education, education for health, class management, educational manager inspector, trainer, mentor, policy and European Integration

• Country:

Romania

Dragan Peraković

• About:

Dragan received B.S degree at the University of Zagreb, Faculty of Transport and Traffic Sciences (Faculty), Croatia, 1995th, 2003. defended masters and 2005. Ph.D. dissertation named: A Model of Distribution of Information to Users of the Transport System. Dragan is Head of the Department for Information and Communication Traffic and Head of Chair of Information Communication Systems and Services Management, all at the Faculty, where he is currently associate professor.

Area of scientific interests and activities: modeling of information and communication systems and their applications; traffic simulation; forensic analysis of information and communication systems, services and terminal devices; design and development of new communications services and software modules; possible application of ICT as assistive technologies (for inclusion of people with disabilities) and e-Learning.

Dragan is co-author of 1 (one) CC paper and author/co-author more then 80 scientific papers in journals and proceedings of international conferences and 4 chapters in international scientific books.

He participated in the work of several international scientific projects and R&D studies (CiViTAS ELAN, Harmonized Inland Navigation Transport Through Education and Information Technology - HINT, Green Intermodal Freight Transport - GIFT, etc.). He is also the initiator and main researcher of many important national projects (Research of the Context of the use of Smartphones and Related Information and Communication Services, Research of Possibilities on Applying the IoT Concept to Improve the Safety Movement of Blind and Visually Impaired Persons along the Traffic Network, Information and Communication Services for the Movement along the Traffic Network by the Persons of Reduced and Difficult Mobility).

He is a full member of TM Forum, IEEE, DAAAM International and SDIWC, Croatian Chamber of Traffic and Transport Technology (board member and president of the professional grade engineers Postal and Information and Communication Traffic).

• Affiliation:

University of Zagreb, Faculty of transport and traffic sciences

• Scientific specialization:

Area of scientific interests and activities is modeling of information and communication systems and their applications in the environment of the transport system, traffic simulation, forensic analysis of information and communication systems, design and development of new communications services and software modules.

• Country:

Croatia

International Scientific Committee and Reviewers Committee

Saber Mohamed Abd-Allah
CAS, China

Mohammad Hossein Ahmadi
International Institute of Earthquake Engineering and Seismology

Abd Elmoniem Ahmed Elzain
Kassala University

Rashad A. Al-Jawfi
Ibb University

Cornelia ANGHEL DRUGĂRIN
Eftimie Murgu University of Resita, Romania

Mohamed Anis Dhuieb
Ecole Centrale de Nantes, France

Karina Cecilia Arredondo Soto
Autonomous University of Baja California, Mexico

Stefan Badura
Publishing Society, Slovakia

Nikolaos C. Kokkinos
Kavala Institute of Technology, Greece

Zdena Dobešová
Univerzita Palackého, Czech Republic

Michael Dossis
TEI of Western Macedonia, Greece

Ivan Grgurević
University of Zagreb, Croatia

Subhash Chander Dubey
Govt. College of Engineering and Technology, India

Mariana Iancu
BIOTERRA University of Bucharest, Romania

Mohammad Jannati
Universiti TeknologiMalaysia, MALAYSIA

George Katsadoros
University of the Aegean, Greece

Niyazi Ugur Kockal

Utku Köse
Usak University, Turkey

Maja Kostadinovska
National and University Library "St. Clement of Ohrid", Macedonia

Ibrahim Kucukkoc
University of Exeter

Binod Kumar
JSPM Jayawant Institute of Computer Applications, India

Igor Kuzmenko
Kyiv polytechnic institute, Ukraine

Maciej Laskowski
Lublin University of Technology, POLAND

Jolanta Latosińska
Kielce University of Technology, Poland

Artis Mednis
Institute of Electronics and Computer Science, Latvia

Farhad Mirzaei
National Dairy Research Institute(N.D.R.I.) Deemed University, India

Angelo Robert Nichola (J) Molson
Ministry of Interior Affairs, Greece

EDWARD MUNTEAN
University of Agricultural Sciences and Veterinary Medicine, Romania

Bruno Nobre
IADE-U (Instituto de Arte Design e Empresa Universitário), Portugal

Bernard OKELO NYAARE
JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY, Kenya

Artur Opalinski
Gdańsk University of Technology, Poland

Fernando Osuna Pérez
University of Granada, Spain

DRAGAN PERAKOVIĆ
University of Zagreb, Croatia

Marko Periša
Faculty of Transportation and Traffic Sciences, Croatia

Hynek Roubik
Czech University of Life Sciences, Czech Republic

Teodor Rusu
University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Romania

Daiva Seskauskaite
Kaunas Forestry and Environmental Engineering University of Applied Sciences, Lithuania

Yilun Shang
Tongji University, Israel

Elżbieta Szafranko
University of Warmia and Mazury in Olsztyn, Poland

Agnieszka Tłuczak
Opole University, Poland

Suyogkumar Vijay Taralkar
MIT Academy of Engineering, India

Kamil Żyła
WEiI Politechnika Lubelska, Poland

Sections discussed at the Conference

Computer and Information Science - Computer vision challenges

Computer architecture

Networking

Software

Web development

Intelligent systems

Other Computer and Information Sciences

Electrical and Electronic engineering

Mathematics

Earth and related Environmental sciences

Conference Sponsors and Partners

Publishing Society

Slovak Republic

www.The-Science.com

Science & Technology Park Zilina

Slovak Republic

Faculty of Transport and Traffic Sciences
Department of Information and Communication Traffic

Published by

EDIS
Publishing Institution of the University of Zilina

Univerzitna 1
01026 Zilina
Slovak Republic

Summary

Dear Reader,

it is our pleasure to introduce you this proceedings. This book contains all accepted papers from conference, which is described below in more details. We hope published papers contribute to the academic society and provide interesting information for researchers world wide.

Conference details:

- Conference full name: **Research Conference in Technical Disciplines**
- Conference short name: **RCITD**
- Conference edition: **3ʳᵈ**
- Conference dates: **October 19 - 23, 2015**
- Conference web page: **www.rcitd.com**
- Conference online archive: **www.rcitd.com/archive**

Conference paper approval process:

Each registered paper was evaluated in double tier approval process.

1. Scientific Committee evaluation (in average 2 reviews were prepared per paper).
2. Conference Editorial Board.

Only papers recommended by these committees were accepted for online presentation at the conference and for publication in this conference book.

Conference presentation:

All accepted papers were presented at the conference during conference dates. Asynchronous online chat was prepared for each paper, where all conference members could freely discuss the topic. During the conference, the Section Chairmen Committee steered the conference discussion. Short presentation were proposed for effective conference discussion.

Conference post processing:

- Each published paper is available in **CDRom** and **online** proceedings (online archive).
- All papers in online archive are available for downloading for potential readers (**open access policy**).
- For effective linking and citation each paper is provided with **DOI**, we recommend to use DOI for referencing.
- The online archive may be possibly evaluated by several Abstracting and Indexing services in future.
- Authors of around 20% of top rated papers were recommended for preparing extended version into academic journals.

Sincere thanks for:

- **Scientific Committee** and **Section Chairmen Committee** for their volunteer work during reviewing.
- Conference **partners** for promotional work and their contribution.
- **Editorial Board** for enormous workload and patience.

Stefan Badura
Conference Organizing Committee
The-Science.com community
Publishing Society

October, 2015

Table of Contents

Earth and related Environmental sciences

Zbigniew Piepiora, Paweł Piepiora

New balancing technique for green cloud computing and environmental Sustainability

Dr. Ayman E. Khedr
Faculty of Computers & Information
Helwan University
Ayman_khedr@helwan.edu.eg

Dr. Mona Nasr
Faculty of Computers & Information
Helwan University
m.nasr@helwan.edu.eg

Hesham Elmasry
Faculty of Computers & Information
Helwan University
Heshamelmasry1980@yahoo.com

Abstract—**The growing demand of Cloud Computing Infrastructure has increased the energy consumption of data centers, which has become a critical issue.**

Data centers hosting cloud computing applications which consume huge amounts of energy, contributing to high operational costs and leads to high carbon emissions which are not environmentally friendly. Therefore we need to propose a green cloud load balancing (GCLB) solution that intended to reduce energy consumption In cloud data center while maintaining the service level agreement (SLA) between the customer and the cloud service provider.

In order to design such solutions, deep analysis of Cloud is required with respect to their power efficiency. Thus, in this paper, we discuss various elements of green clouds which contribute to the total energy consumption and how it is addressed in the previous studies. We also discuss the implication of this solution on the energy efficiency (EE) and the quality of service (QOS).

Keywords: load balancing, green cloud computing, energy Efficiency, quality of service and data center

I. INTRODUCTION

The energy-efficient load balancing techniques on cloud computing aims to maintains and enhances both quality of service (QOS) and resource utilization while reducing the power consumption to achieve the environmental sustainability concept.

The generally definition of cloud computing (CC) comes from The National Institute of Standards and technology (NIST) as follows: "cloud computing is a model for enabling convenient, on-demand network access To a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." [1]

NIST also offers up several characteristics that it sees as essential for a service to be considered "Cloud". These characteristics include;

- *On-demand self-service:* the ability for an end user to sign up and receive services without the long delays that have characterized in traditional IT.

- *Broad network access:* the ability to access the service through the standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile Phones, laptops, and PDAs).

- *Resource pooling:* The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided Resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, Network bandwidth, and virtual machines.

- *Rapid elasticity:* Capabilities can be rapidly and elastically Provisioned, to quickly scale out and rapidly released to quickly scale in. To the end user, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

- *Measured Service:* cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the Provider and consumer of the utilized service.

A. Cloud computing service model

As shown in figure 1 the stack model for CC consist of three major services each of them have certain responsibilities which are shifted to the cloud service provider to allowing the consumer of cloud services to focus more on their own business requirement and less on the underlying technologies.

These major services can be summarized as follow:

- *Software as a Service (SaaS)*

The capability provided to the customer is to use the Provider's applications running on a cloud infrastructure. the applications are accessible from various client devices through a thin client interface such as a web browser (e.g. Web-based email). The customer does not manage or control The underlying cloud infrastructure, including network, servers, operating systems, storage, or even Individual Application capabilities [2].

- Platform as a Service (PaaS)

The capability provided to the customer is to deploy onto The cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure Including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

- Infrastructure as a Service (IaaS)

The capability provided to the customer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating Systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications.
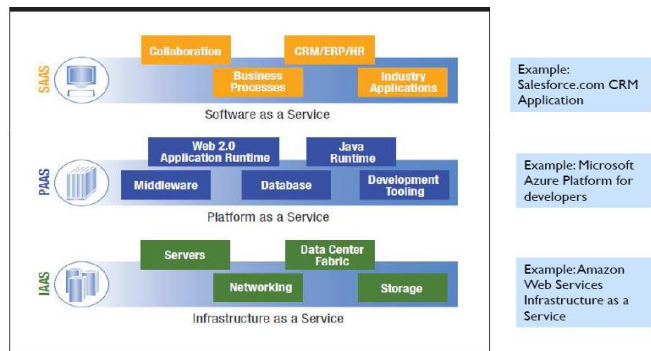
Figure 1.   Cloud computing stack model

(Source: Mather, Kumaraswamy and Latif, 2009)

B. Cloud computing roles

Roles define the responsibilities, access and profile of different users that are part of a cloud computing solution. Figure 2 presents these roles defined in the three service layers [4].
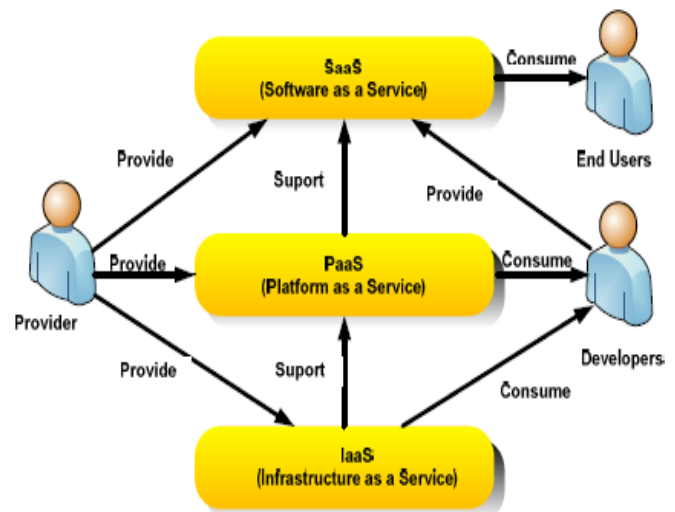
Figure 2.   Cloud computing roles

(Source: Marinos and Briscoe, 2009)

The provider is responsible for managing, monitoring and guaranteeing the availability of the entire structure of the cloud computing solution. It frees the developer and the final user from such responsibilities while providing services in the three layers of the architecture, the developers use the resources provided by IaaS and PaaS to provide software services for final users.

C. Cloud computing deployment model

According to the intended access methods and availability of cloud computing environments, there are different models of deployment; they include private cloud, public cloud, community cloud and hybrid cloud, which are briefly analyzed in table 1.

 -Private cloud: The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

-Public cloud: the cloud infrastructure is made available to the general public and is owned by cloud service provider.

- Community cloud: the cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns. It may be managed by the organizations or a third party and may exist on premise or off premise.

- Hybrid cloud: the cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology [5].

TABLE I.   CLOUD COMPUTING DEPLOYMENT MODELS.

	Infrastructure Managed By[1]	Infrastructure Owned By[2]	Infrastructure Located[3]	Accessible and Consumed By[4]
Public	Third Party Provider	Third Party Provider	Off-Premise	Untrusted
Private/ Community	Organization Or Third Party Provider	Organization Third Party Provider	On-Premise Off-Premise	Trusted
Hybrid	Both Organization & Third Party Provider	Both Organization & Third Party Provider	Both On-Premise & Off-Premise	Trusted & Untrusted

(Source: Cloud security alliance team, 2009)

D. Green cloud computing and environmental sustainability

Green cloud computing (GCC) is the practice of implementing policies and procedures That improve the efficiency of computing resources in such a way as to reduce the energy consumption and environmental impact of their utilization [7][8].this means how to take the maximum benefits from CC to the service providers and customers while minimize the negative impact on environment such as energy consumption and carbon emissions footprint , in order to achieve this goal we need an efficient solution to Achieve this difficult equation .

Environmental and energy conservation issues have taken center stage in the global business arena in recent years. The reality of rising energy costs and their impact on international affairs coupled with the increased concern over the global warming climate crisis and other environmental issues have shifted the social and economic consciousness of the business community.
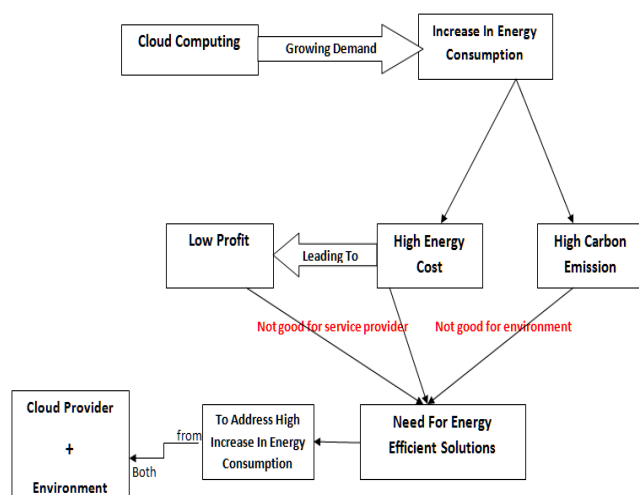

Figure 3.   Cloud and environmental sustainability (Source: Kansal and Chana, 2012)

As shown in figure3the growing demand for cloud computing Increases the energy consumption in data centers, which has become a major concern for both industry and society [9].this increase in energy consumption not only increases energy cost but also increases carbon emission. High energy-Cost results in reducing cloud providers' profit margin and high carbon emission which has negative impact on the environment [10]. Therefore we can use the Load balancing technique as energy saving solution to achieve the GCC Concept.

II.   BACKGROUND AND RELATED WORK

This section presents the various work and studies that has been done so far by the scientific community dealing with energy efficient and load balancing techniques in cloud data center.

- Energy efficient techniques

The first study discusses energy efficient computing and stress about the importance of saving energy [12]. They present the environment protection agency (EPA) report in their paper which highlights some of the following important points:

- 61 billion kWh (kilowatt hours) was consumed by servers and data center alone in 2006.

- IT equipment which is necessary to run the data center alone consumed a considerable amount of energy.

The second study try to extend the tickles Idle by introduce an Eco-friendly daemon which uses dynamic voltage and scaling to reduce the power consumption and strictly maintaining performance However, we need to predict the future workload to adjust the frequency of the CPU accordingly. If the windows size is too large, then we have to wait for a while to get the required windows size data to make an accurate prediction. Finding the right parameters to sample the workload to estimate the future workload is a challenge [13].

The third study proposed an approach for power conservation in server systems based on fast transitions between active and low power states. The goal is to minimize power consumption by a server while it is in an idle state [14]. Instead of addressing the problem of achieving energy proportional computing as proposed by barroso and holzle, the authors require only two power states (sleep and fully active) for each system component. The other requirements are fast transitions between the power states and very low power consumption in the sleep mode [15].

Load balancing techniques

The motivation of the survey of existing load balancing techniques in cloud computing is to encourage the amateur researcher to contribute in developing more efficient load balancing algorithms, this will benefit interested researchers to carry out further work .

The first study proposed a new content aware load balancing policy named as workload and client aware Policy (WCAP) it uses a unique and special property (USP) to specify the unique and special property of the requests as Well as

computing nodes. USP helps the scheduler to Decide the best suitable node for the Processing the request, this strategy is implemented in a decentralized manner with Low overhead [16].

The second study proposed a join idle queue load balancing algorithm for dynamically scalable web services, this algorithm provides large scale load balancing with distributed dispatchers by, first load balancing idle processors across dispatchers for the availability of idle processors at each dispatchers and then, assigning jobs to processors to reduce average queue length at each processor, by removing the load balancing work from the critical path of request processing , it effectively reduces the system load ,incurs no communication overhead at job arrivals and does not increase actual response time [17] .

The third study proposed a new server based load balancing policy for web servers which are distributed all over the world. It helps in reducing the service response times by using a protocol that limits the reduction of Requests to the closest remote servers without overloading them; a middleware is described to implement this protocol [18].

III. Green Cloud Load Balancing Model

Based on server utilization, there are often periods of idleness where the data center servers are not used and during this period these servers still consume about 65-70% of the power consumed when it is fully utilized [19] [20] [21] [22] without doing any useful work, therefore we introduce in this paper a GCLB model to eliminate the wasted power consumption during the idle period while maintaining the QOS.

The straight forward scheme is to set the server to sleep once the queue is empty and is waken up upon a new job arrival [23] this approach works but with some limitations and it is not efficient due to the following concerns:

- The fast transition from the running mode into the sleep mode might be too costly. If the idle duration or non-idle duration is short, the server is wasting time in transition without taking enough duration of sleep (or execution of jobs).

- The average response time threshold would be always violated if the wake-up transition time is above the average Response time threshold. All jobs after the wakeup will Experience at least the wake-up transition delay; this is another big drawback of this simple approach.

In our design of energy saving schemes, we propose the following constant parameters of time periods to overcome the above raised concern:

- Time to sleep threshold (TTS) it is the minimum length of the idle duration before the server is put into the sleep power mode.

- Time to wake threshold (TTW) it is the maximum length of the sleep duration which the server can stay in the sleep power mode continuously.

-Time to delay threshold (TTD) if the job arrives earlier before the expiration of (TTW) it will remain in queue till the Expiration of (TTW) period.

- Green Cloud Load Balancing Model

Our GCLB model can be described as the following steps:

I- once the queue is empty; the server intends to stay in the idle power mode for TTS threshold.

1- If a new job arrives before the expiration of the TTS threshold, the server will immediately serve the new job. Otherwise, the server will enter to the sleep power mode and then it will be enforced to stay in the sleep power mode for TTW threshold.

2- If a new job arrives before the expiration of the TTW threshold the new job will wait in queue for TTD threshold. Otherwise, the server will continue in the sleep mode until the expiration of TTW period.

3- once the server is in the running power mode, the server runs at a constant speed ratio (R) in the running power mode serving jobs in the queue until the queue become Empty .

4- if there are servers with low utilization their load will transferred to the medium load servers .

5- if there are servers with high utilization , the system will wake up one of the sleeping server and transfer the extra load to it .

6- Once the queue is empty, repeat Step I

Energy Saving
1-for all compute servers S do
2-calculate current CPU utilization for each compute server S
3-end for
4-if Utilization=0% // this compute node is idle
5-if Tidle ≥ TTS
6-put the compute server into sleep mode
7-else
8- if Utilization ≤30% // this compute server is under- loaded
9-transfer the load to the medium –loaded servers
10-end for
balance
11-if Utilization >65% utilization // this compute server is over- loaded
12-Wake up one of the sleeping compute node
13-Then
14-Transfer extra load from over-loaded server to new wake-up server
15-end if

Figure 4.   Green cloud load balancing algorthim

The green cloud load balancing algorithm has two basic sections:

the energy section is responsible for achieving energy saving by switching idle compute server to sleep power mode, and if the compute server is using less than30% of its resources, GCLB Algorithm transfer their load to one of The medium load server. It does this by first gathering the utilization percentage of each active compute server.

The balancing section is responsible for maintaining QOS and response time for job requests by wake up one of sleeping nodes that can help the overloaded servers to lighten its load, and this is happened if one of the servers has utilization over (65%).

the threshold of (65%) utilization was chosen since when (35%) of the resources are available, at least one more virtual machine can be accommodated using three out of five available configurations.

IV.   RESEARCH METHODOLOGY

This section presents the general view of the system architecture followed by a more detailed view of the component's architecture additionally; it shows the measurement's system that describes the power, Performance and CPU utilization's measurements. Then, the experimental methodology will be presented.

In order to achieve the objectives of this research and due to the nature of the study we opt to employ cloud Based E-learning model as shown in Figure 5, where the new system model is being constructed and evaluated, we find it useful to employ the modeling and experimental methodology, because such methodology enables us to describe the proposed system model and further validate it through experiment processes where it becomes possible to compare and analysis results between the current and proposed model.
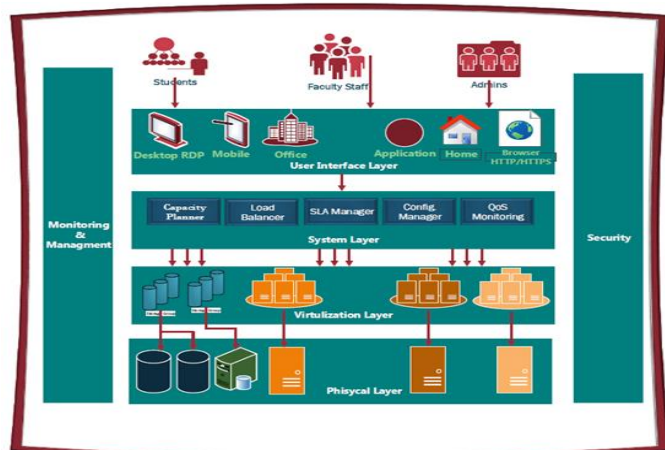
Figure 5.   Cloud based E-learning model

As shown in figure 6, an actual cloud system e-learning model has been implemented we find it is useful to use the measurement and experimenting methodology for collecting, analysis and validating the proposed GCLB model.

Figure 6.   Research methodology steps

- Data gathering methods

The overall approach to this thesis is to collect useful, quantifiable data to conclude whether or not increasing the server utilization efficiency and sleeping idle servers will be advantageous in reducing power consumption and maintain the response time.

The following areas are explored to gather sufficient data for a power saving analysis.

- CPU Utilization (Usage)

CPU utilization is a key performance metric and major indicator of power consumption. it can be used to track CPU performance regressions or improvements, and is a useful data point for performance problem investigations.

In this paper the proposed GCLB algorithm will depend on an idea of setting a CPU utilization threshold to distinguishing between the under-Loaded and over-loaded states of the host.

When the algorithm is invoked, it compares the current CPU utilization of the host with the defined threshold. If the threshold is exceeded the algorithm detects a host overloaded and vice versa.

We will use Processor :(%) Processor Time counter in windows performance monitor built-in tool to monitor the amount of time the CPU spends processing a non-idle thread.

- Power consumption

We will use the power metering and budgeting (PMB) tool to measure and monitor the power consumption in order to evaluate and validate the proposed GCLB model.

A power meter is a hardware component of the system that reports information about power consumption, in watts. This information is typically provided as part of a power supply or by using a base board management controller.

This infrastructure tool promotes energy efficiency on cloud data center servers by providing power consumption and management features, additionally PMB provides extra options to configure power metering and budgeting to help cloud provider and end users to balance power and performance to meet their needs.

The PMB infrastructure feature provides two type of information:

Power Metering Information

This information is used to determine how the computer system or sub-components use power. Power consumption is monitored, or metered, by a power meter in the system. Power metering also provides the current configuration of a power meter, such as metering capabilities and power consumption thresholds.

Power Budgeting Information

This information is used to determine the power limit, or budget, that is supported by the computer system. Depending on the hardware platform, this information might also allow you to configure the system's power budget.

Performance (Throughput)

Throughput is a measure of how many units of information a system can process in a given amount of time. We will use System\% user time and System\% privileged time counter in windows performance monitor tool to measure The percentage of time the processor is spending executing User processes.

- Data analysis Methods

Due to the nature of the quantitative research type and the numerical data which collected through the above data collection methods, we find it is useful to use the graphical chart and comparison table methods to analysis and validate the collected data.

We will compare and analysis the Power consumption, CPU utilization and throughout data which collected before and After we implement the proposed GCLB model and display The results through graphical chart and comparison table to discover and understand the relationship between the number of users, CPU utilization and power consumption and how the proposed model can achieve energy efficiency while maintain the performance and server availability.

- Experimental Setup

In this section, we will discuss the cloud E-learning model which will use in our experimental setup to evaluate and validate the proposed GCLB Model, we discuss the software and hardware used and the modifications made to them as shown in Figure 6.

- Hardware required

- Cloud Host server (Hyper V)

The study will use dell precision T5400 blade server, with processor Intel (R) Xeon (R) CPU X5450@ 3.00GHz 2.99GHz (4 processors), installed memory 16 GB and system type: 64-bit operating system, x64-based processor.

- Web Front End Servers (Guest Nodes)

The study will use three dell precision T5400 workstation, with processor Intel (R) Xeon (R) CPU X5450@ 3.00GHz 2.99GHz (2 processors), installed memory 6 GB and system Type: 64-bit operating system, x64-based processor.

- Software required

- Cloud Host server (Hyper V)

We will use windows server 2012 R2 data center with Microsoft Hyper V 2012 feature.

Web Front End Servers (Guest Nodes)

We will install on each web front end server, windows server 2008 R2, Microsoft Network load balancing (NLB) 2012, PHP 5.3 None thread safe installer, PHP Manager 1.2 for IIS 7.5, SQL server 2012 and Moodle 2.6.
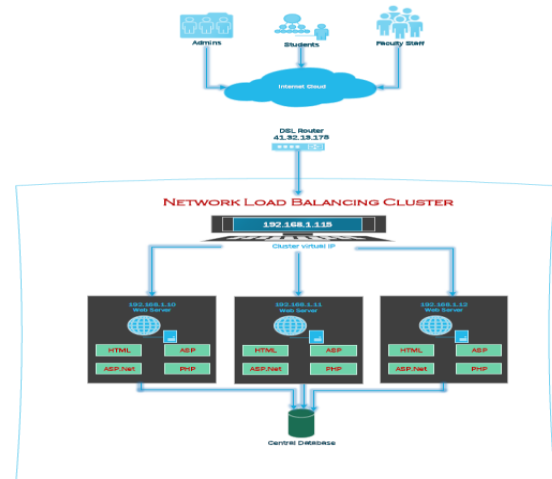
Figure 7.   Experimental setup Model

V.   EXPERIMENTAL RESULTS AND DISCUSSION

This part presents the results of the experiments that were conducted based on the methods described in pervious section, The data obtained throughout the experiments were analyzed and interpreted .summaries of results are generally presented in figures. Typical graphical charts relating to the experiments will be provided in this paper to analyze the relationship between the energy consumption and values of the utilization threshold. The values of the result show that the lower threshold has higher influence on the energy consumption than the upper threshold. This can be explained by the fact that an increase of the lower threshold by implementing the proposed GCLB eliminates the low utilization of the resources leading to higher energy saving.

For the benchmark experimental results we have used a Normal Load Balancing (NLB) policy. This policy does not apply any power saving optimizations and implies that Workload will distribute equally between all nodes regardless the CPU utilization percentage and the server idle periods. The second policy applies GCLB model which eliminating the servers idle period by sleeping the data center servers during idle periods based on the CPU utilization and GCLB algorithm.

Comparative results of the energy consumption and average throughout times for each of the two policies with different sets of request arrival rates

The NLB policy exhibit higher energy and power consumption, the reason is primarily due to the fact that servers once activated, stay in operational or active state throughout the duration of the run. The idle power of activated servers when not in use, contributes to the substantially high power and energy consumption.

Using our power saving based approach, energy and power consumption of the cluster is reduced and gives a savings of at least 40% when compared with NLB baseline policy during idle period, as shown in figure 15 and 18

On response times, Normal Load Balancing (NLB) Baseline policy perform the best because, servers are more often in operational state for the requests to be processed compared to GCLB schemes .

- Comparison of GCLB and Normal Load Balancer NLB

The results obtained by using our GCLB algorithm would obviously bring higher power savings over conventional load balancing algorithm, In the normal load balancing approach, the cluster controller assigns virtual machine requests to compute nodes sequentially. This effectively balances the load across the compute nodes, but leaves these nodes in an "on" state with usually low utilization. Figure 1 shows the power consumption of the Normal load balancing algorithm.

The major performance difference between our GCLB algorithm and the Normal Load balancing approach is that compute nodes that are idle using GCLB are switched to the low power sleep mode. The Normal Load balancing approach always keeps all compute nodes powered on, no matter the number of active virtual machines. The Normal Load balancing approach is effective in load balancing across the available compute nodes, but such a relatively simple load balancer consumes a large amount of unnecessary power.

We will use our experimental cloud Moodle to monitor the power consumption, response time, disk and memory for web front end servers under different workload before and after applying the Green Cloud Load Balancing (GCLB) model.

Finding analysis and discussion
- power consumption saving during peak period

After we measure power consumption for cloud host server During *peak* period before and after we implement the GCLB model as shown in figure 11 and 12 the power Consumption reduced from 155 watt to 149 watt with power Saving percentage (4 %)

- Power consumption saving during idle period

After we measure power consumption for cloud host server During *idle* period before and after we implement the GCLB Model as shown in figure 15 and 16 the power consumption Reduced from 167 watt to 100 watt with power saving Percentage (40 %)

The GCLB algorithm is more effective in reducing power Consumption during idle period due to the low CPU Utilization and huge power wasted which can be eliminated Through put the idle servers in sleep power mode state and Consolidate the workload on the active servers which Increase CPU usage and reduce power consumption.

- Throughput

After we measure the throughput for cloud server before and After we implement the GCLB model during peak period As shown in figure 10 and 11 and during idle period as Shown in figure 15 and 16 the proposed GCLB model Maintain the throughput and the cloud performance without SLA violation

GCLB VS NLB during peak period

Figure 8.   Cloud Server during peak period before GCLB

Figure 9.   Cloud server during peak period after GCLB

Figure 10. Throughput during peak period before GCLB

Figure 11. Throughput during peak period after GCLB

Figure 12. Power consumption during peak period before GCLB

Figure 13. Power consumption during peak period after GCLB
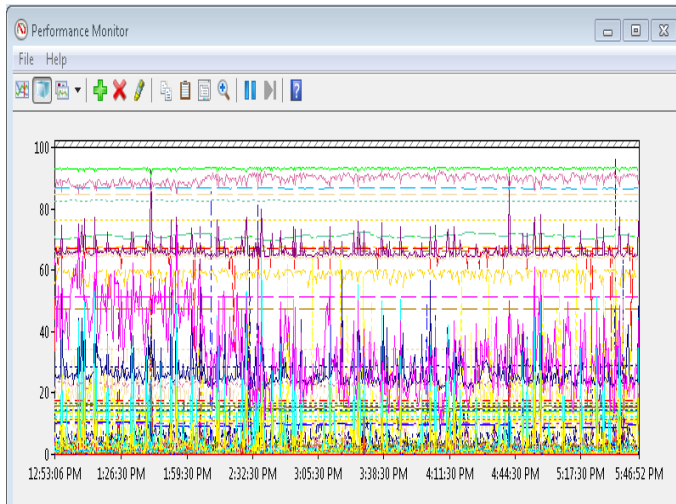
GCLB VS NLB during idle period

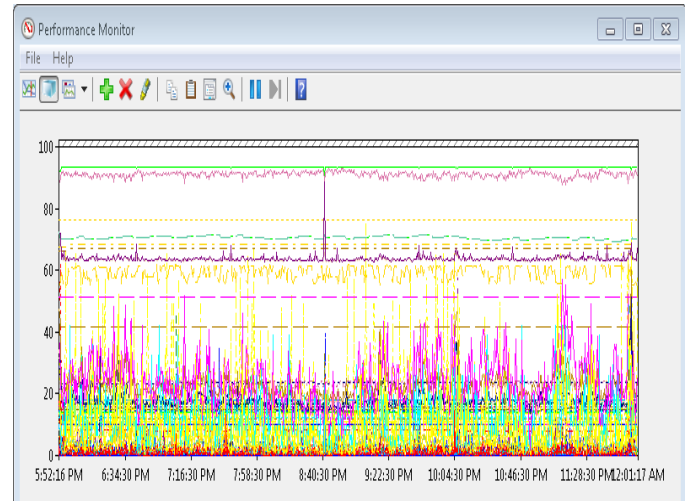Figure 14. Cloud Server during idle period before GCLB

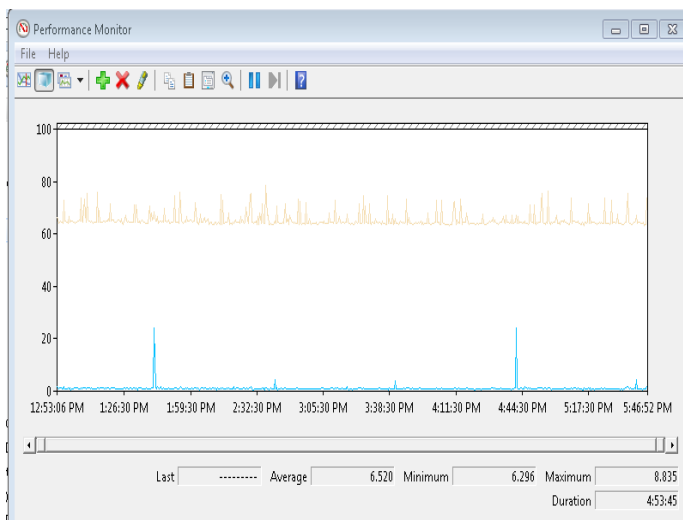Figure 15. Cloud Server during idle period after GCLB

Figure 16. Throughput during idle period before GCLB
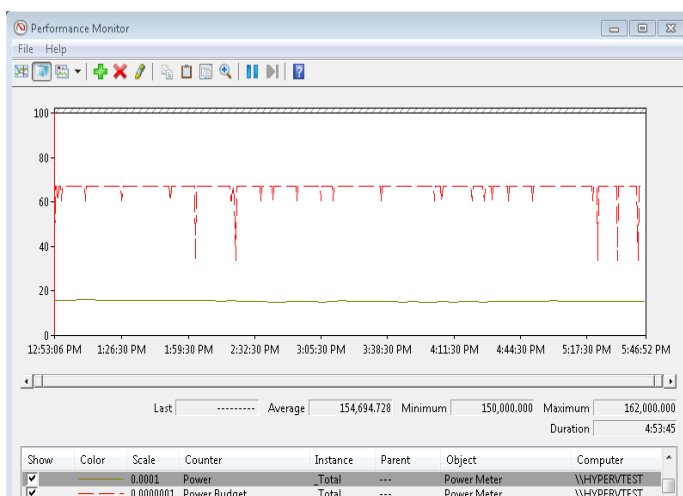

Figure 17. Throughput during idle period after GCLB

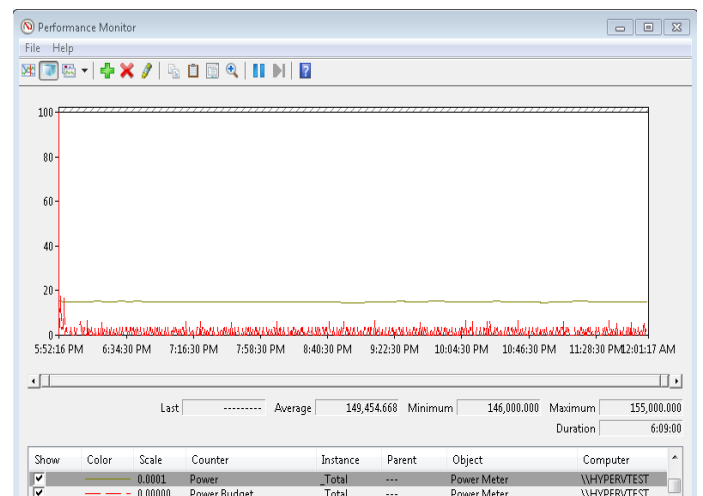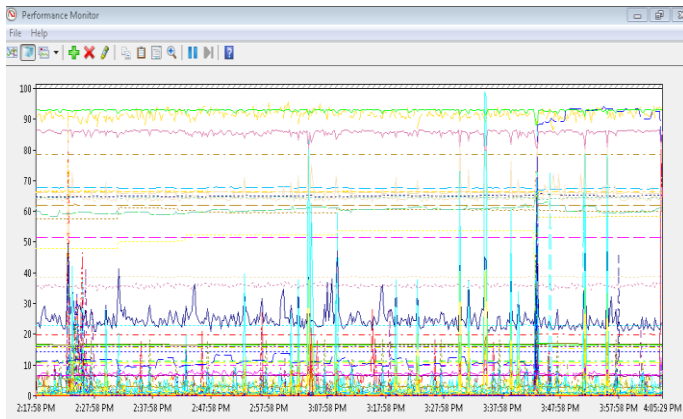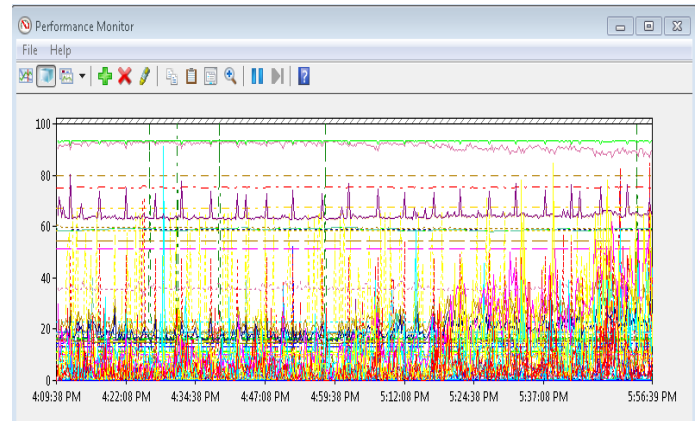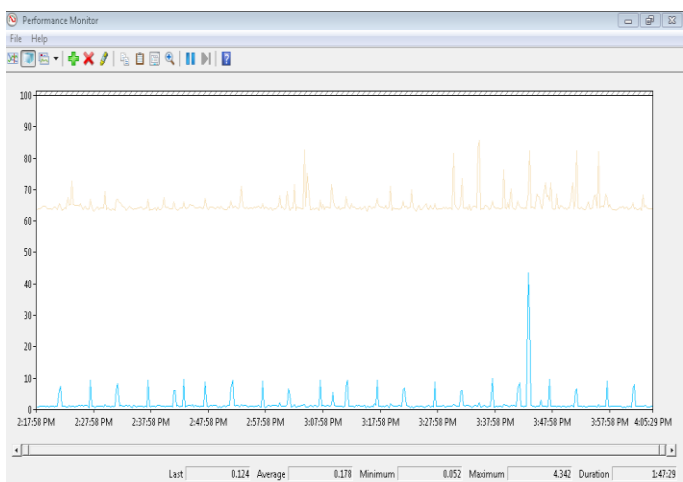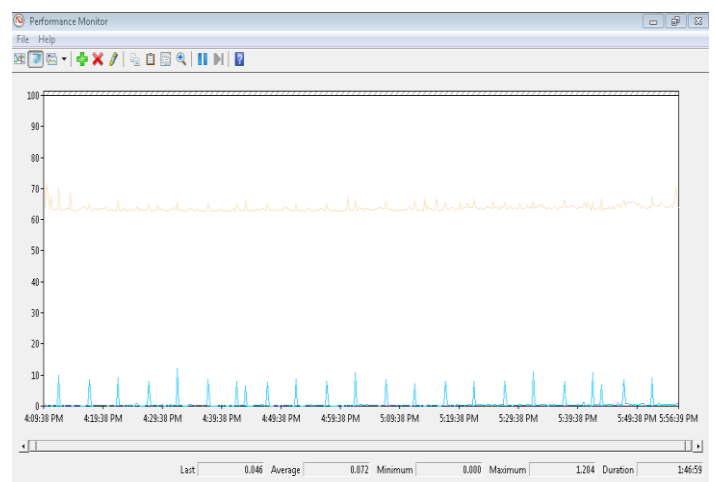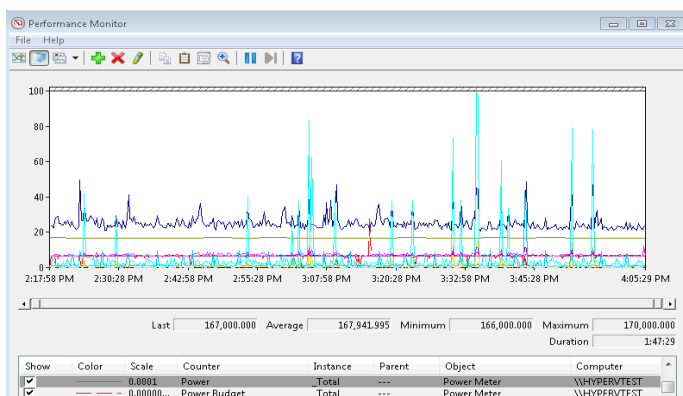Figure 18. Power consumption during idle period before GCLB
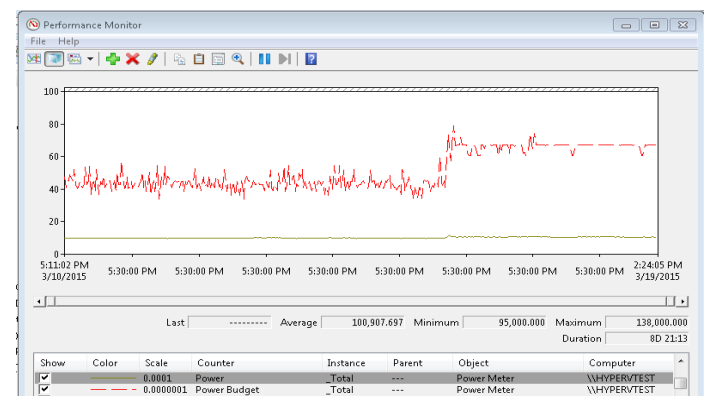
Figure 19. Power consumption during idle period after GCLB

REFERENCES

[1] P. Mell and T. Grance. "The NIST definition of cloud computing". Pp.1-106, (2011).

[2] M .AL-Roomi, S. Al -Ebrahim, S. rais and I. Ahmad, "Cloud Computing Pricing Models: A Survey" International Journal of Grid and Distributed Computing Vol.6, No.5 (2013), pp.93-106.

[3] T. Mather,S. Kumaraswamy, S. Latif "Cloud Security and privacy an Enterprise perspective on Risks and Compliance", 2009.

[4] A.Marinos and G. Briscoe, "Community cloud computing," in First International Conference Cloud Computing, CloudCom, volume 5931 of Lecture Notes in Computer Science, pp. 472–484. Springer (2009).

[5] C. Rao,M. Leelarani and Y. Kumar"cloud computing services and deployment models", International Journal Of Engineering And Computer Science, volume 2 issue 12, pp.3389-3392 ,December 2013 .

[6] Cloud security alliance team, "Security guidance for critical areas of focus in cloud computing ",available From:https://cloudsecurityalliance.org/guidance/csaguide.v1.0.pdf [accessed April.2009]

[7] S. Kabiraj, V. Topka, and R. C. Walke, "Going Green: A Holistic Approach to Transform Business", International Journal of Managing Information Technology (IJMIT), Vol. 2, No. 3, August 2010, pp. 22-31.

[8] J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker,"Green Cloud Computing: Balancing Energy in Processing, Storage, and Transport", Proceedings of the IEEE, Vol. 99, No. 1, January 2011, pp. 149-167.

[9] K. M. Nagothu, B. Kelley, J. Prevost, and M. Jamshidi, "Ultra low energy cloud computing using adaptive load prediction", Proceedings of IEEE World Automation Congress(WAC) , Kobe, September 2010, pp. 1-7.

[10] S. K. Garg, C. S. Yeob, A. Anandasivamc, and R. Buyya, "Environment-conscious scheduling of HPC applications on distributed Cloud-oriented data centers", Journal of Parallel and Distributed Computing, Elsevier, Vol. 70, No. 6, May 2010, pp. 1-18.

[11] N. Kansal and I. Chana "Cloud Load Balancing Techniques: a step towards green Computing",IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012 .

[12] R. Brown, E. Masanet, B. Nordman and B. Tschudi,, "Report to Congress on Server and Data Center Energy Efficiency", Public Law,July 2007,pp.109-431.

[13] S .Huang, W. Feng, " A workload-aware, eco-friendly daemon for cluster computing", Technical Report, Computer Science, Virginia Tech (2008).

[14] D. Meisner, B. T. Gold, and T. F. Wenisch, Power Nap: eliminating server idle power, ACM SIGPLAN Notices, vol. 44, no. 3, pp. 205–216, 2009.

[15] L. Barroso and U. Holzle, "The case for energy-proportional computing," Computer, vol. 40, pp. 33 – 37, Dec 2007.

[16] H. Mehta, P. Kanungo, and M. Chandwani, "Decentralized content aware load balancing algorithm for distributed computing environments", Proceedings of the International Conference Workshop on Emerging Trends in Technology (ICWET), February 2011, pp. 370-375.

[17] Y. Xiea Q. Kliotb G. Gellerb A..Larusb and A.Greenber, "Comparative Analysis of Existing Dynamic Load Balancing Techniques", International Journal of Computer Applications, IJCA 70(26): pp. 25-29, May 2013.

[18] Nakai A.M., Madeira E. and Buzato L.E. (2011) 5th Latin-American Symposium on Dependable Computing, PP.156-165.

[19] A .Greenberg,., J .Hamilton,., D .Maltz,and P .Patel, "The cost of a cloud: research problems in data center Networks", ACM SIGCOMM Computer Communication Review, Volume 39 Issue 1, January 2009, pp. 68-73.

[20] C. Lefurgy, X. Wang, and M. Ware "Server-level Power control", ICAC'07: International Conference on Autonomic Computing, June 2007.

[21] X. Fan, W. Weber, and L. Barroso, "Power Provisioning for a warehouse-sized computer", ISCA'07: International Symposium on Computer Architecture, 2007.

[22] L. Barroso and U. Holzle, "The Case for Energy-Proportional Computing". IEEE Computer, pp. 33–37, December 2007.

[23] A.Gandhi, V.Gupta, M.Harchol-Balter, and M.Kozuch. "Optimality analysis of energy-performance trade-off for Server farm management", Performance Evaluation, Volume 67 Issue 11, pp. 1155-1171, November, 2010.

Multi-level Verification with Synthesis

Using the Cubed-C High-level Synthesis Framework

Michael Dossis

dept. of Informatics Engineering
TEI of Western Macedonia
Kastoria, Greece
Michael.Dossis@gmail.com

Abstract—**Current developments towards the increased complexity of electronic systems and devices determine the use of formal and automated synthesis techniques at least for major custom parts of these devices. Using formal synthesis and verification techniques are a means to cut down significantly the development and verification times of complex electronic integrated circuits and systems. The Cubed-C synthesizer enables the formal verification of complex circuits by providing the automatic generation of cycle-accurate C simulators from the synthesized FSM model. In this way the designers are not required to spend time with lengthy and cumbersome RTL and gate-level simulations. It also provides an immediate way to compare the functionality of the input ADA/C code with the generated FSMs. Two example applications from the area of computer graphics and DSP filters are discussed in the experimental part of this work, which prove the usability of the presented methodology.**

Keywords- E-DA, HLS, Formal Verification, Programming languages

I. Introduction

Digital circuits that are used in embedded, high-performance and portable computing systems have highly complex components, design control & module hierarchy as well as interconnection schemes. These product features cannot be dealt anymore with conventional methods such as RTL coding, since these methods suffer from highly iterative design flows and prolonged development times. Thus, often products miss the market windows due to excessive project delays and a lot of engineering effort is lost. In order to deal with this complexity, commercial and academic organisations have invested in High-Level Synthesis (HLS) and optimisation techniques, so as to achieve design automation, quality of implementations and short specification-to-product times [1-5], [7]. However existing HLS tools usually produce lower than average quality of implementations and hardware models that are difficult to handle, with a lot of assumptions about the computing platform and transformation heuristics in order to deal with the NP-complete synthesis problems.

Logic programming [6] and compiler generators can benefit HLS tools, found in other areas, such as artificial intelligence and software compilers. Most discussed and formulated problems of HLS such high-level optimizations, scheduling, allocation and binding are explained in [1-5], [7]. High-level optimizations resemble software compiler optimizations, allocation is selection of functional units and storing resources for the data and operations objects found in high-level program

code, binding is the actual mapping of the above units to real hardware elements such as flip-flops, latches and combinatorial blocks such as functional operator hardware units, and scheduling is the arrangement of elementary operations to Finite State Machine (FSM) states or in other words real system's clock cycles. Although these problems were extensively studied in research labs, the optimization of real-world complex applications, and their mapping onto custom hardware fails to produce competitive implementations due to their inability to handle complex, nested control flow and large data objects, as well as sophisticated interfaces through complex hierarchy and module configuration.

The Cubed-C framework [4] combines formal techniques from compiler generators and artificial intelligence logic programming so as to deliver formal synthesis of high-level programs (in ADA or C) into fully synthesizable and simulatable RTL (VHDL or Verilog). Thus short spec. to product times can be achieved and lengthy RTL or gate-level simulations can be avoided.

Section II includes related bibliography. Section III discusses the Cubed-C formal verification flow. Extended experiments are discussed in the following section. Conclusions and future work are discussed in the last section.

II. Bibliography

A. Issues with existing High-level Synthesis tools

High-Level Synthesis research efforts started in the 80s and the first linear processing HLS tools appeared in the academic and industrial labs, in the early 90s. Important problems that researchers of HLS were called to handle were allocation, scheduling and binding, as they were outlined in the introduction of this paper. The most complex of these three problem solvers is the building of a reliable scheduler [3]. It is well known that when the system complexity increases linearly, the complexity of the scheduler algorithm increases exponentially and for some very complex applications, scheduling is NP-complete [3], [7]. This difficulty to handle complex code becomes critical, difficult and even prohibitive in practice when source code models with complex module and control flow hierarchy (e.g. nested while, ifthen and for loops) are to be processed by the HLS tool [4], [7].

Existing HLS tools are still not widely accepted by the engineering community because of their poor results, especially for large applications with complex module and control-flow hierarchy as stated before. However, such

applications are to be benefited the most from automated techniques such as HLS. For large-scale applications, the complexity of the synthesis transformations (front-end compilation, algorithmic transformations, optimizing scheduling, allocation and binding), increases exponentially, when the design size increases linearly [3], [4], [5], leading to suboptimal solutions when synthesis heuristics are employed to cut down the long optimization and processing times.

Many of the available HLS tools impose proprietary extensions or restrictions (e.g. exclusion of while loops) on the programming model of the specifications that they accept as input, and various heuristics on the HLS transformations that they utilize (e.g. guards, speculation, loop shifting, trailblazing) [2]. Most of them are suitable for only linear, and dataflow dominated (e.g. stream-based) applications, such as pipelined DSP, image processing and video/sound streaming, and which again are not possible to handle when they contain any of the excluded programming constructs (such as while loops and loop breaks). Nevertheless, when design's FSM increases to more than a douzen states, manual RTL coding approaches are prohibitive for delivering products in realistic times.

The most popular commercial existing HLS tools include the Catapult-C from Calypto (previously developed by Mentor Graphics), and Cynthesizer from Forte Design Systems. They both accept as input a small subset of the System-C and C++ languages. However, both of these tools have too complicated user interfaces, for the average system modeller and developer and they are the most expensive of their class since they are licensed for something less than 300K dollars per year. Therefore, these E-CAD products are inaccessible for many small and medium sized ASIC/FPGA design SMEs. Other commercial or industrial HLS tools are the Symfony C compiler from Synopsys, the Impulse-C from Impulse Accelerated Technologies, the CyberWorkBench from NEC, the C-to-silicon from Cadence, and the free web-based tool C-to-verilog from an Israel-based group.

Amongst the academic or research-based HLS tools are the SPARK tool [2] which accepts as input a small subset of the ANSI-C language (e.g. while loops are not accepted), and a conditional guard based optimization method [7] which set the basis for optimizing conditional source code at the beginning of the previous decade.

Alternative approaches include a multi-speculative technique to synthesize complex adders during datapath synthesis, which again helps only towards linear data flow oriented designs [8], a fixed-point accuracy analysis and optimization of polynomial data-flow graphs with respect to a reference model that is found in many digital signal processing applications [9], a technique to improve nested loop pipelining for High-Level Synthesis, called Polyhedral Bubble Insertion [10], an equivalence checking method of FSMs with datapaths based on value propagation over model paths, for validation of code motion transformations usually applied during the HLS scheduling phase [11], a formal method for accurate high-level casting of optimal adders and subtractors for HLS [12], and an exploration approach, called spectral-aware Pareto iterative refinement, that uses response surface models (RSMs) and spectral analysis for predicting the quality of the design points

without resorting to costly architectural synthesis procedures [13].

B. Verification in High-level Synthesis

In the Cubed-C approach, high-level C or ADA code is first executed to verify the correctness of the specification model in a rapid way. Then the debugged code is passed to the Cubed-C HLS synthesizer to generate correct-by-construction, Register Transfer Level (RTL), VHDL implementations of computation-intensive applications. RTL-level simulations are carried out to verify rapidly the designs and prove the argument of the correctness at the level of the automatically generated RTL VHDL implementations. The Cubed-C framework consists of the frontend and the backend compiler, which communicate with each other via the ITF database. The frontend compiler was built using compiler-compiler techniques and the backend compiler using logic programming techniques [4]. Moreover the ITF syntax and semantics are formally defined in [14]. Both of these methodologies are already patented with international patents [15]. The backend synthesis is optimized with the PARCS scheduler. PARCS is a formal optimizer which attempts to parallelize as many operations in the same clock cycle as possible, as long as control/data dependencies and resource constraints are obeyed.

Another more recent option of the Cubed-C tools is the automatic, formal and rapid generation of ANSI-C testbenches from the same hardware engine that is synthesized with the HLS tool. These testbenches are cycle-accurate simulators that the user can compile and execute with any of the available C development environments such as the GNU-C compiler and linker. By doing so the user can see through the FSM states execution the changes in the values of registers. Moreover, resetting the registers, providing input to all the inputs of the design and examining the values of the output registers of the design at any time in the cycle-accurate simulation are possible. Although this method is similar to formal verification is different than the latter, since the Cubed-C method integrates both synthesis and verification methods in the same development step. In this way, functional and other bugs can be caught quickly and corrected without a lot of engineering effort.

This synthesis and verification flow in the Cubed-C framework is shown in Figure 1. There, red frames include the verification processes and blue frames incorporate the synthesis tasks. If the same test vectors feed the high-level functional and the backend cycle-accurate simulations then the results of the two simulations (which was done in our experimental tests at the end of this paper) can be compared in order to confirm the guaranteed matching of the two behaviours.

Model checking is emerging the recent year as a formal means of verifying the system model with mathematical assertions. However for large applications it is not practical since the state space is exploding to prohibiting levels. A recent attempt to deal with this is found in [17]. A method based on a formal attribute grammar, used to translate human language into correct Computation Tree Logic from specific comments in the HDL code, is described in [18]. Transformations of the behavioural RTL FSMD model are proposed in [19] in order to deal with false negative results of the equivalence checking of

state sequences from conditional paths in equivalent checking. Static and dynamic analysis of the RTL source code is exercised in [20] to define and compute code coverage of assertions in the RTL code. Simulation-based verification is applied in [21] by generating stimuli for a number of scenarios which are triggered in various ways. These stimuli are very compact while satisfying the coverage requirements.
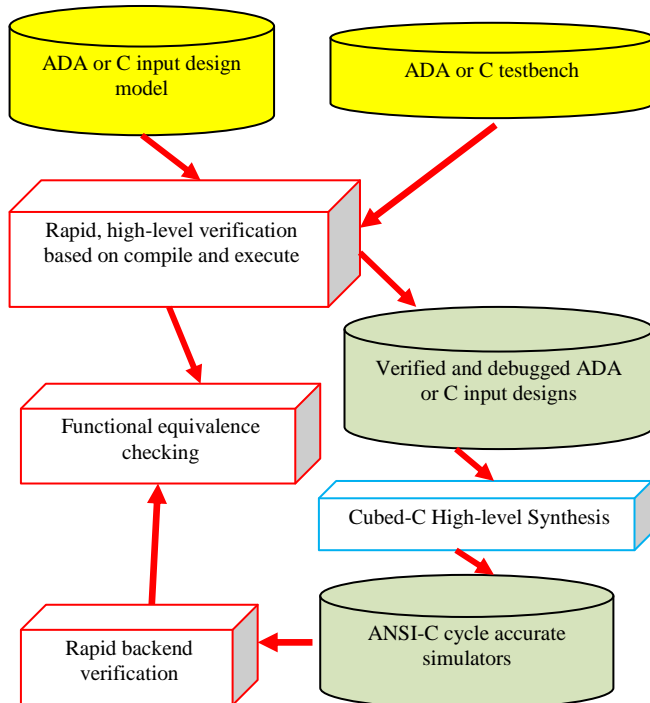
Figure 1.   Formal verification in the Cubed-C synthesis flow

III.   VERIFICATION IN THE CUBED-C SYSTEM

The formal, rapid Cubed-C verification method converges seamlessly into bug-free implementations. This is due to the formal and rapid nature of the Cubed-C HLS compiler. Due to the limits in the number of pages only two applications are discussed in the experiments section, however it is capable to prove the usability of the Cubed-C design/verification flow.

All of the executed experiments featured a convergence on the functionality of the high-level algorithmic code with the RTL simulations and the C testbench execution, in a fraction of time required with traditional approaches such as exhaustive and repetitive RTL/gate-netlist simulations.

Ongoing development targets generated cycle-accurate simulators and the inclusion of other targeted verification languages such as C++ and System-C. Recently, a new component integrated into the HLS Cubed-C backend compiler generates an ANSI-C based testbench extracted directly from the internal optimized FSM of each of the generated hardware modules. One cycle-accurate C simulator is extracted from every one of the synthesized coprocessor modules. Along this, the backend compiler generates a corresponding compile script so that the user can easily and rapidly compile, link and execute the cycle-accurate simulators using the GNU tools.

The Cubed-C integrated development and verification flow is shown in Figure 1. First the user builds a system model in a high-level programming language code such as ADA or C. The user also builds the required verification testbench in the same language using high-level algorithmic code. Then both the model-under-test and the testbench are compiled and coexecuted on the host platform. The testbench includes generated and expected test vectors for automatic verification.

After all of visible design bugs are removed the user ports the debugged design code to the Cubed-C synthesis tools and a number of hardware modules are automatically generated by the tools. Along with the RTL code, and for every one of the hardware modules a C cycle-accurate simulator is generated. The user compiles and executes the C simulators and verifies that the FSM engine is compliant to the source code in terms of functionality. This functional equivalence is guaranteed by the formal nature of the synthesis transformations anyway. For increased confidence, any RTL simulations can be executed as well, although this is not necessary. In this way both the RTL – level and the source code – level verification tasks are converging into the same conclusions about the correctness of the implemented system functions.

IV.   EXPERIMENTAL RESULTS

Many cycle-accurate application simulators were compiled and executed. Here the focus will be in a few experiments. First an application for drawing lines from computer graphics is generated, compiled and executed. Figure 2 shows a snapshot of the executed simulator run between states 0 and 3, as well as the option to reset the design and give values to the inputs before that.

Figure 2.   Cycle-accurate simulation of the computer graphics application near the begining of the test

Usually the verification of the cycle-accurate and the high-level testbenches were accompanied with regular RTL FSM simulations and in all cases the same behavior was confirmed. However, with the Cubed-C methodology this is not needed due to the formal nature of both synthesis and verification flows.

Figure 3 demonstrates a portion of the simulation screen of the cycle-accurate simulator of the computer graphics simulation near the last few states. Figure 4 shows the end option of viewing the coprocessor's results. From all these figures is obvious the usability of the generated coprocessors' testbenches. The simulator is customized to the needs of the particular coprocessor functions, and it includes options such as reset the design, give values to inputs, read the outputs and proceed to the next state in the coprocessor FSM operation.

Figure 3.   Snapshot of the cycle accurate simulation of the line drawing algorithm application in the middle of the execution

Figure 4.   The end of the simulator execution for the line drawing hardware

The second (amongst many) benchmark which is discussed here is a 5 tap FIR DSP filter. Arbitrary input samples and filter coefficients were used but this doesn't harm at all the argument of this work and it is suitable to implement real existing FIR filter taps. Figure 5 demonstrates the high-level ADA code testbench run which is a rapid one. The testbench execution snapshot includes the input values the filter coefficients and the filter's history trace.

Figure 5.   The high-level functional testbench execution of the FIR filter

The FIR filter was synthesized with the Cubed-C synthesizer and the cycle-accurate C testbench was extracted automatically and formally from the generated FSM. Figure 6 shows the cycle-accurate simulator near the beginning of the execution where the input values and the coefficients are given.

Figure 6.   Cycle-accurate simulator run during the begining of the execution

After resetting the registers and setting the values for the inputs the filter's cycle-accurate C model is executed in order to reach the completion of the FSM's operation and read the outputs. The completion of the testbench is shown in Figure 7, where it clearly shows the value matching with the high-level ADA testbench values of Figure 5. It is worthy to mention that in all cases the compilation and execution of both levels of testbenches (functional and cycle-accurate) were seamless, they used the GNU environment and that the cycle-accurate models verified the same operation as the ADA/C specification.

Moreover the execution of both verification levels and the Cubed-C synthesizer did take less than a minute in all cases. The use of the GNU tools proved the independence of the Cubed-C flow from any proprietary tools and methods.

Figure 7.   The completion of the cycle-accurate FSM execution and reading the outputs of the simulator

V.   Conclusion and Future Work

The Cubed-C verification method is rapid, robust and easily converging into bug-free implementations. This is supported by the formal and rapid nature of the Cubed-C HLS compiler. Due to the limits in the number of pages only two applications were discussed in the experiments above, but they are capable of demonstrating the usability of the followed design and verification flow.

All of the executed experiments featured a convergence on the functionality of the high-level algorithmic code with the RTL simulations and the C testbench execution, in a fraction of time required with traditional approaches such as exhaustive and repetitive RTL/gate-netlist simulations.

Further improvements for the user are envisaged in the generated cycle-accurate simulators and the inclusion of other targeted verification languages such as C++ and System-C.

References

[1]   B. L. Gal, E. Casseau, and S. Huet, "Dynamic Memory Access Management for High-Performance DSP Applications Using High-Level Synthesis". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 16, no. 11, pp.1454-1464, 2008.

[2]   S. Gupta, R. K. Gupta, N. D. Dutt, and A. Nikolau, "Coordinated Parallelizing Compiler Optimizations and High-Level Synthesis". ACM Transactions on Design Automation of Electronic Systems, vol. 9, no. 4, pp. 441–470, 2004.

[3]   R. A. Walker, and S. Chaudhuri, "Introduction to the scheduling problem". IEEE Design & Test of Computers, vol. 12, no. 2, pp. 60–69, 1995.

[4]   M. F. Dossis, "A Formal Design Framework to Generate Coprocessors with Implementation Options". International Journal of Research and Reviews in Computer Science (IJRRCS), August 2011, ISSN: 2079-2557, Science Academy Publisher, United Kingdom, www.sciacademypublisher.com, vol. 2, no. 4, pp. 929-936, 2011.

[5]   P. G. Paulin, and J. P. Knight, "Force-directed scheduling for the behavioral synthesis of ASICs". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 8, no. 6, pp. 661–679, December 1989.

[6]   U. Nilsson, and J. Maluszynski, Logic Programming and Prolog, John Wiley & Sons Ltd., 2nd Edition, 1995.

[7]   A. A. Kountouris, and C. Wolinski, "Efficient Scheduling of Conditional Behaviors for High-Level Synthesis". ACM Transactions on Design Automation of Electronic Systems, vol. 7, no. 3, pp. 380–412, 2002.

[8]   A. A. Del Barrio, Rom´an Hermida, S. O. Memik, Jos´e M. Mend´ıas, and Mar´ıa C. Molina, "Multispeculative Addition Applied to Datapath Synthesis". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 31, no. 12, pp. 1817-1830, December 2012.

[9]   O. Sarbishei and K. Radecka, "On the Fixed-Point Accuracy Analysis and Optimization of Polynomial Specifications". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 32, no. 6, pp. 831-844, June 2013.

[10]   A. Morvan, S. Derrien, and P. Quinton, "Polyhedral Bubble Insertion: A Method to Improve Nested Loop Pipelining for High-Level Synthesis". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 32, no. 3, pp. 339-352, March 2013.

[11]   K. Banerjee, C. Karfa, D. Sarkar, and C. Mandal, "Verification of Code Motion Techniques Using Value Propagation". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 33, no. 8, pp. 1180-1193, August 2014.

[12]   R. Sierra, C. Carreras, G, Caffarena, and C. A. López Barrio, "A Formal Method for Optimal High-Level Casting of Heterogeneous Fixed-Point Adders and Subtractors". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 34, no. 1, pp. 52-62, January 2015.

[13]   S. Xydis, G. Palermo, V. Zaccaria, and C. Silvano, "SPIRIT: Spectral-Aware Pareto Iterative Refinement Optimization for Supervised High-Level Synthesis". IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 34, no. 1, pp. 155-159, January 2015.

[14]   M. Dossis, "Intermediate Predicate Format for Design Automation Tools", Journal of Next Generation Information Technology (JNIT), vol. 1, no. 1, pp. 100-117, May 2010.

[15]   M. Dossis, Patent number 1005308, 5/10/2006 by the Greek Industrial Property Organisation, 2006.

[16]   D. Amanatidis and M. Dossis, "High level synthesis of geometric active contours", Proceeding of the 2nd Global Virtual Conference, Slovakia, April 7-11, 2014.

[17]   L. Wu, H. Huang, K. Su, S. Cai, and X. Zhang, "An I/O Efficient Model Checking Algorithm for Large-Scale Systems", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 23, No. 5, pp. 905-915, 2015.

[18]   C.B. Harris, and I.G. Harris, "Generating formal hardware verification properties from Natural Language documentation", Proceedings of the 2015 IEEE International Conference on Semantic Computing (ICSC), Anaheim, CA, pp. 49-56, 7-9 February, 2015.

[19]   R.A. Hernandez, M. Strum, and W.J. Chau, "Transformations on the FSMD of the RTL code with combinational logic statements for equivalence checking of HLS", Proceedings of the 16th Latin-American Test Symposium (LATS), Puerto Vallarta, Mexico, pp. 1-6, 25-27 March, 2015.

[20]   V. Athavale, Ma Sai, S. Hertz, and S. Vasudevan, "Code coverage of assertions using RTL source code analysis", Proceedings of the 51st ACM/EDAC/IEEE Design Automation Conference (DAC), Moscone Center, San Francisco, CA, USA, pp. 1-6, 1-5 June, 2014.

[21]   Yang Shuo, R. Wille, and R. Drechsler, "Improving Coverage of Simulation-Based Verification by Dedicated Stimuli Generation", Proceedings of the 2014 17th Euromicro Conference on Digital System Design (DSD), Verona, Italy, pp. 599-606, August 27-29, 2014.

Simulation Analysis of Characteristics and Application of Software-Defined Networks

Ivan Grgurević, Zvonko Kavran, Anthony Pušeljić (student)

Faculty of Transport and Traffic Sciences, University of Zagreb

Department of Information and Communication Traffic

Zagreb, Croatia

ivan.grgurevic@fpz.hr

Abstract - **Software-defined network (SDN) is an approach to computer networking that allows network administrators to manage network services through abstraction of higher-level functionality. This research includes comparison of multiple scenarios of the software-defined network, which are based on different types of coverage and local area networks (LAN), i.e. a traditional LAN. Differences are evident in the scenario of network performance and can be perceived as advantages and disadvantages of SDN in relation to the traditional network. The parameters used in the analysis are data rate, packet delay (i.e. latency), packet loss, throughput, the cost of network performance and others. The application and the simulation demonstration of a software-defined network is shown in the graphical network simulator GNS and emulator Mininet. This research has analysed the advantages and disadvantages of a software-defined network over a conventional network, taking into account various parameters and stakeholders.**

Keywords - Software-defined network/networking, simulation analysis, controller, Application programming interface (API)

I. INTRODUCTION

Nowadays, we are witnessing a very high degree of application of virtualization technologies with the growing customer demand for a fast establishment and delivery of services and placement services within the Cloud Computing concept. In addition, users require flexible and automated network environment that is adaptable to current applicative requirements. Such new challenges require responses by the application of a different approach in relation to the classical network infrastructure management. Cloud computing allows users to store data and install software on the servers that are connected through the Internet. With the help of a web browser and special customers, these services are flexible and the users pay only for what they use.

Software-defined network (SDN) is a network architecture in which the networks control is separated from the packet forwarding and it contains the possibility of direct programming. Such migration of control, which is sometimes strongly related to an individual network device, in the external computing devices allows basic infrastructure separation of applications and network services, which are therefore able to treat the network as a logical or virtual entity. SDN enables dynamic adjustment of the network environment to the current application requirements or the user's needs, and simplifies management and increases the scalability of the network, which is particularly manifested through a simple implementation of additional network services and components. An additional benefit of SDN is the possibility of using the network components from different manufacturers, basically without having to know how to operate the devices since the complete network environment is managed from a single point, or through the SDN controller. The SDN network architecture consists of a controller SDN, OpenFlow network devices and a communication channel that connects them.

Today, the largest application of SDN is present in data centers which are also known as software-defined data centers (SDDC). Such data centers contain all the elements of the infrastructure needed for networking, storage, processing (Central processing unit - CPU), the realization of security and virtualization, and are being delivered as a service. Development, provisioning, configuring, and other operations of the whole infrastructure are separated from the hardware and executed by the software.

The aim of this research was to conduct an analysis of the characteristics and the application of software-defined networks. The analysis is based on a comparison of conventional networks and software-defined networks with the display of significant differences. The research includes a simulation of different network topologies using the graphical network simulator GNS3 on Linux. For the purposes of the simulation, it was necessary to specify the differences between network architecture of traditional networks and software-defined networks, and to conduct the process of designing software-defined network via a graphical network simulator GNS3 and emulator Mininet. The research is in fact an analysis of different scenarios and parameters (data rate, packet delay, packet loss, throughput, the cost of network performance, etc.).

II. BACKGROUND AND RELATED WORK

Numerous available articles and research are dealing with the analysis of the characteristics and architecture of software-defined networks / networking, and the analysis is mainly based on the impact due to changes in certain performance of the network and the application of SDN controller [1], [2], [3]. Within the development of the Internet of Things (IoT) concept, many authors reveal the application of software-defined networks / networking and access in the IoT environment, and thus achieve the differentiation level of service due to the different needs of IoT in different (heterogeneous) scenarios, especially related to the wireless networks [4], [5], [6]. The development of software for the

simulation of the operation of information and communications networks has achieved efficient testing of various networks and network elements, ways of networking and the presentation of various possible scenarios, which is also present within software-defined networks. Typically used software for the implementation of software-defined networks simulation is the OpenNet [7], Mininet [8], ns3 [9] and EstiNet [10]. The article [11] presents a comparative analysis of the existing simulators for SDN according to different characteristics and functions.

According to [12], the authors were interested in research of the SDN technology and its possibilities, and were thereby using Mininet simulator and POX SDN controllers. The results were compared with the results obtained by the application of network devices and the use of "traditional" network. The throughput in a software-defined network is increased in comparison to a "traditional" network and the number of lost packets in a software-defined network is smaller.

Within the SDN analysis conducted by the Open Network Foundation (ONF) it has been concluded that separate control and data planes result in better programmability, automation and better control of the network, which results in scalable and flexible networks which allow, for example, business companies to easily adapt to variable business needs [13]. Analyzing the issues of SDN, the Cisco Systems company has come to the conclusion that SDN greatly helps to simplify operations by automating and centralizing network business management [14]. One part of the research also analyzes the traffic parameters as part of the transport engineering in SDN networks, using various simulation methods and simulation experiments [15], [16].

III. OVERVIEW OF CHARACTERISTICS AND ARCHITECTURE OF SOFTWARE-DEFINED NETWORK

The SDN concept is based on the need to separate and redefine a network construction, and its implementation uses the following three principles:

1) Control and forwarding planes: Control planes are separated from the forwarding planes. Forwarding planes are still located in the switch, while control planes are moved to the SDN controller in the form of software.

2) Control intelligence: Control intelligence is centralized at SDN controller.

3) Network programmability by applications: The network can be programmed beginning from the applications. Applications interface can be exposed to the controller to manipulate the network.

The main objective of SDN is to achieve better management of networks with large extent and complexity and to ensure that all logical decisions of control level are made from the central point. This central access will reduce the need for the N-number of intelligent nodes in an N-nodes topology. The basic role of every network software is to program the path that will allow the traffic to flow. Now, when the dependence of software on the hardware is reduced, there is no need for intelligent software to operate on all nodes.

SDN is based on the concept of logical starting of software in a centralized location and programming of switches using the southbound Application Program Interfaces (API). Figure 1 shows the logical layers of SDN. At the lowest level there are network elements such as switches, computers, servers and other network devices. It is important to note that the switches are located on top of the lowest layer. The middle layer is a layer of controller that communicates with the switches.
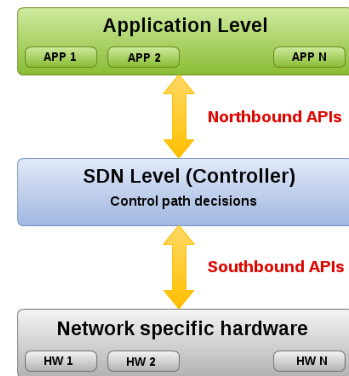
Figure 1.   Logic layers of SDN

The highest level is the application level in which the user can define the applications that will allow the definition of the network flow. As result, a network approaches the applications as one logic switch thus providing control of the entire network from one logic point and simplifies the network design and all of the operations within the network. SDN also simplifies the operating of the network devices because they no longer have to understand, but only to process a lot of protocol standards led only by the instructions of the SDN controller [14].

A. SDN Controller

The central controller (SDN controller) is a software entity that needs to have a global view on the entire network. The network operating system, launched logically for the choice of path, needs to be launched on the central SDN controller. The controller has an overview of the entire network and it can determine the optimized flow and program hardware ports. The basic characteristics of the controller are:

- Detection of end user devices such as laptops, desktops, printers, mobile terminal devices, etc.

- Detection of network devices that form the network infrastructure such as switches, routers and wireless access points.

- Management of network devices topology by maintaining information about the details of the link between the network devices and directly connected terminal devices.

- Control of database maintenance managed by the controller and performing of necessary coordination with the devices to ensure the synchronization of flow entry of devices with that database.

B. Southbound API

Within the architecture of software-defined network, the southbound API are being used for communication of SDN controllers with network switches and routers. Southbound APIs mitigate the efficient network control and allow the SDN controller to dynamically make changes according to the real-time requirements and needs.

C. Northbound API

Within the SDN network, northbound APIs are being used for communication of the SDN controller with the services and applications launched within the network. Northbound APIs can be used in order to mitigate the innovations and provide an efficient orchestration and automatization of network which can align due to its programmability with the needs of various applications. Northbound APIs are most critical of all within the SDN environment, because the value of SDN is related to innovative applications which can be potentially supported and provided and they have to support a wide range of applications.

IV. PLANNING OF SOFTWARE-DEFINED NETWORK

Many organizations inforce the initiative of the implementation of SDN solution, but there is a question of the best performance onto the more automated network architecture and what is to be considered and applied within. In many cases, the software-defined solution does not need to look any different from the conventional network. It is important to define the impact of the SDN model on the existing services and to use samples of applications that connect and checkout the continuity of the service before and after the implementation. That will prevent the disruption of service and eliminate all implementation-related problems. However, regardless of the number of preparations, some of the circumstances are still unpredictable. Therefore, it is important to have an alternative plan that allows the administrators to return the previous network configuration. The implementation of SDN without the proper knowledge represents a certain risk, but the ignoring of SDN represents a significant risk for IT organizations and IT experts. In the case of IT organizations, the risk is that they will not be able to solve the problems for which SDN has been designed, which results in the lack of competitiveness. The risk for IT professionals is that they can fall behind in learning and education related to this approach and thus will not have the competitive value for the current or for the future employer. SDN security needs to be built into the architecture, as well as delivered as a service to protect the availability, integrity, and privacy of all connected resources (and information).

According to previous analyses and research it can be concluded that in the upcoming period SDN will have a significant impact on corporate networks and roles of the network experts. Because of that, it is important that the IT organizations and the IT experts develop a plan for the SDN implementation. The implementation may vary depending on the size and the complexity of the network, as well as the experience of the IT team. New skills and additional training is needed. With proper planning, most organizations can quickly and easily take advantage of SDN solutions.

V. SIMULATION DEMONSTRATION OF SOFTWARE-DEFINED NETWORK OPERATION

A. Development of the network topology

Graphical Network Simulator 3 (GNS3) [17] was used for the purpose of creation of the conventional network topologies and the network configuration switches. It will be installed on the Linux OS, which allows the combination of virtual and real devices and allows the simulation of complex networks. It uses Dynamics emulation software to simulate Cisco's Internetwork Operating System (Cisco IOS). The software used in the configuration of the switches is used in the actual physical devices. The network topology used to compare the conventional configuration and the SDN network consists of five switches and two Linux PCs connected as shown in Figure 2.
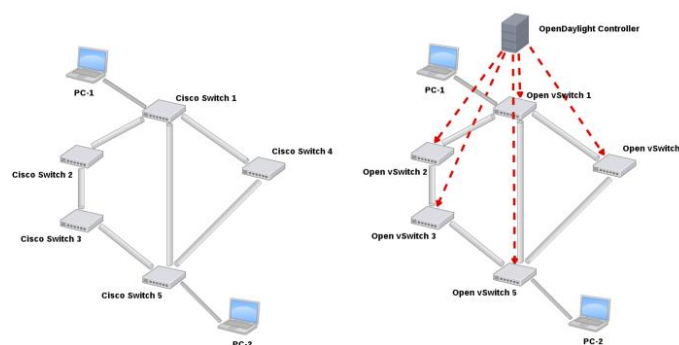
Figure 2.   Network topology of conventional network and SDN network [18]

By using GNS3 simulator, a network is created as shown in Figure 3. Before any traffic can flow from PC-1 to PC-2, it is necessary to configure all the switches, to make sure that the traffic from PC-1 to PC-2 flows via the shortest route, which is made possible by using FIB (Forwarding Information Base) on each switch.
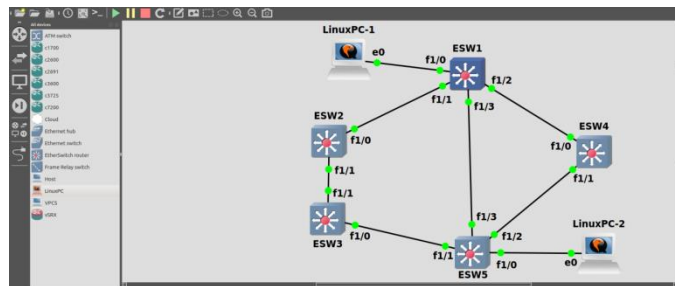
Figure 3.   Conventional network topology in GNS3 simulator

Once the ports, IP addresses and configured switches are assigned, the network topology must be learned by all switches.

Even though the network topology contains a small number of network devices and is not complex, it can be concluded that the configuration of larger networks has higher demands for the number of connections and the time. If the network consists of a thousand network devices and hosts, which is very frequent in today's networks, every switch and every flow has to be particularly configured for the appropriate traffic. This process takes a number of procedures and increases the total

time. Within the SDN solution all the procedures, from switch configuration and learning of network topology, are performed by SDN controller from one centralized point and within a very short period of time, which makes it an advantage over the conventional networks. The only condition is that switch has to be connected onto the SDN controller, and all the other work is performed by the controller itself.

The development of the SDN network topology is performed by Mininet [8] emulator in order to show how the controller operates. This emulator allows creation of virtual networks and initiates a real kernel, switch and the application code on virtual machine, which is in this case VM Virtual Box. Mininet is installed on the Linux OS and uses an appropriate script in *Python* programming language to initiate previously created network topology. The SDN controller is needed with the use of Mininet. For the purpose of testing OpenDaylight controller was used [19]. The topology consists of five Open vSwitches and two Linux PCs all connected as shown in Figure 2. After successful initiation of Mininet, it is necessary to start your own creation of topology specially intended for this case and written in Python programming language under the name of TestTopology. The command for initiation of the test network is:

sudo mn –mac –controller=remote,ip=192.168.165.1,port=663 –custom TestTopology.py –topo=mytopo, and the meaning of the individual parts of the command are:

- sudo mn: initiates command with *root* privilege

- --mac: sets MAC addresses of *hosts* similar to IP addresses, which makes it easier to read the generated traffic shown in Wireshark

- --controller=remote: informs the Mininet that SDN is not on the local computer

- Ip=192.168.165.1: IP address of the SDN controller, as well as the IP address of the *host* computer where the controller is started

- Port=6633: Standard TCP port for connecting the switch onto the controller

- --custom TestTopology.py –topo=mytopo: initiates its own topology written in *Python*.

By entering an accurate code, Mininet will create the network by adding controllers, hosts, switches and links that will configure the hosts and initiate switches. Figure 4 shows the visibility of switches and their connection in the OpenDaylight controller.
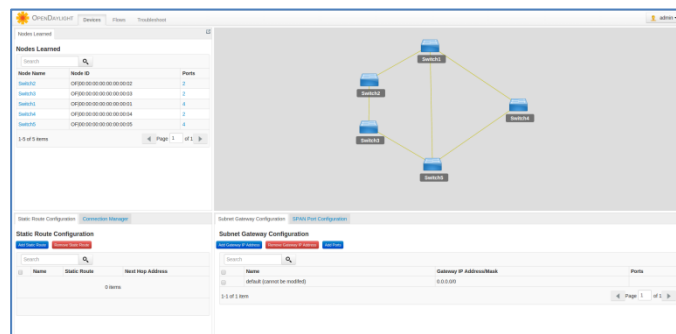
Figure 4.   The learned nodes in *OpenDaylight* controller

Even though this is about the creation of a virtual network, the used controller is also used in real physical networks. It is evident that this process facilitates the processes of a conventional network, where every single switch has to be configured manually, which is within the SDN controller quickly performed by separating the control planes from the data planes, which are still present within the switch. After the controller knows about the switch, the next step is to gain insight into the entire view of the network (i.e. learn about the details of switch devices and about the connections between them). This is conducted in two steps: the first step is to learn about the individual switches, and the second is to learn about the connections between the switches. The first step is performed by feature request and feature reply mechanisms. The controller sends feature-request message at the moment the so-called TCP handshake is conducted. The newly connected switch replies with the feature-reply message. The feature-reply message informs the controller about the capabilities of the switch, details of the port and the available operations. In the next step, the identification of the switch connections is made by *Link Layer Discovery Protocol (LLDP)* frames that are sent onto the connected ports of switches.

B. Performance measurement

Measuring of the performances includes two different scenarios, where the measurement of the permeability and packet delay within the client server communication based on the TCP protocol will be the first one, and the second one will measure the packet loss based on the UDP protocol. Scenario 2 is different in ending of the link triggered by closing the port between Switch 1 and Switch 5. Network topologies of conventional and SDN network are made in GNS3 simulator, to ensure the same conditions, on 4 Linux PCs by using real software with Cisco switches and Open vSwitch software based on SDN switches. Figure 5 shows the presentation of the used topology for Scenario 1.
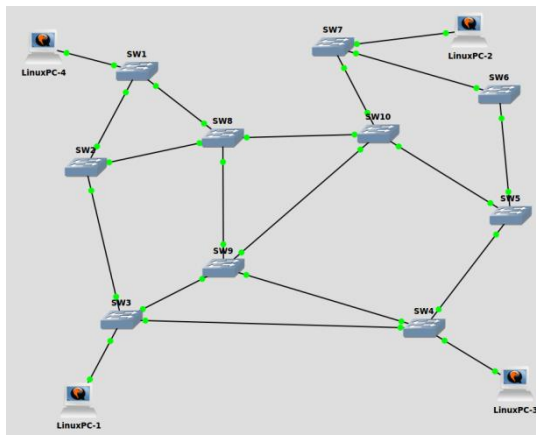
Figure 5.   Topology for Scenario 1

Conventional network is made of the following devices:

- 10 Cisco c3725 Ethernet Switch Router, and

- 4 Linux PC (PC-1 and PC-3 as client, PC-2 and PC-4 as server).

SDN network is made of the following devices and controllers:

- 10 Open vSwitch 1.11,

- 4 Linux PC (PC-1 and PC-3 as client, PC-2 and PC-4 as server), and

- OpenDaylight controller.

The traffic is generated within 5 minutes by the Distributed Internet Traffic Generator (D-ITG) [20] triggered on all PCs.

Before the measurement takes place, the necessary procedures are needed in order to enable the measurement, and these are:

1) To prevent deviations in measurement, all of the PCs are supposed to have their clocks synchronized and linked to the public Network Time Protocol (NTP) servers. For that reason, it is necessary to create gateways towards the Internet. That will be conducted by Cloud in GNS3 simulator, which is in fact a tunnel between the host computer and Linux PC. The IP address on the TAP host computer is also a default gateway for virtual Linux PC. In order for Linux to retrieve certain web addresses it is important to define the DNS server as the Google DNS server. The NTP server is a reference for the synchronization of the clock and that is, in this case, CARNET NTP server located in Zagreb (University Computing Centre of the University of Zagreb).

2) Creating of the Linux Bridge: Open vSwitch used in the SDN network is Virtual Box Appliance in the GNS3 simulator. Used ports must be added to Linux bridge so that the switches could communicate with OpenDaylight controller. After having performed successful needed configurations, the generator on PC-1 and PC-3 is initiated as sender, and the PC-2 and PC-4 as receiver.

The topology used for the presentation of Scenario 2 is shown in Figure 6. The settings are the same as inScenario 1, the only difference being the existence of one sender (PC-1) and one receiver (PC-2) and the use of UDP protocol.
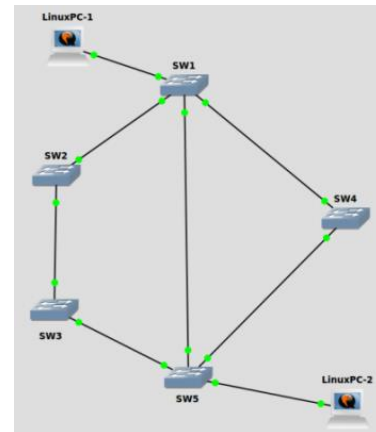
Figure 6.   Topology for Scenario 2

The measurement results are shown in Table 1 and Table 2. The results in the first measurement show that SDN offers equal performances if set conditions are identical. There is a slight difference in numbers because, since the topology is known in advance, the first packets in the beginning of the SDN solution will be sent faster than they would be in a conventional network. The difference is that the switches must first learn the topology and this creates the delay in relation to SDN. Although the conditions were identical and the used software real, these results in reality can be differentiated and vary because of influencing factors such as distance, links and, of course, the used hardware.

TABLE I.   MEASUREMENT RESULTS FOR SCENARIO 1

Scenario 1		
Parameters	*Traditional network*	*SDN network*
Total packets	209,253	222,052
Avg. delay (s)	0.004798	0.004424
Bytes received	214,275,072	227,381,248
Avg. bitrate (Kbit/s)	5,714.098545	6,063.7873
Avg. packet rate (pkt/s)	697.521795	740.20841

TABLE II.   MEASUREMENT RESULTS FOR SCENARIO 2

Scenario 2		
Parameters	*Traditional network*	*SDN network*
Total packets	265,425	285,875
Avg. delay (s)	0.002623	0.002791
Bytes received	271,795,200	323,456,000
Avg. bitrate (Kbit/s)	7,248.044165	8,625.693047
Avg. packet rate (pkt/s)	884.771016	952.941046
Packets dropped	59,920 (18.42%)	544 (0.19%)

It is very difficult to compare a traditional network with SDN on the basis of the measured performances because SDN

is designed with the objective of flexible and easy network management. The SDN solutions are different depending on the manufacturers and the network can be configured according to the need. The performance is adjustable and it depends on the purpose of the organization and why it needs to use network services.

The second measurement shows higher difference in the results. While for SDN the loss was only 0.19%, for the conventional network it was 18.42%. The difference in reality can oscillate, but the SDN solution will definitely yield better results. Knowing the concept of the topology, the controller knows where to direct the packet if the link is disrupted or a certain port is closed, and its performance is very fast. In conventional switch devices, the topology has to be primarily learned because the switch operates only with the nearest unit and has no knowledge of the current state in the network. STP has four conditions, and these are: *blocking*, *listening*, *learning* and *forwarding*. Once the port is blocked it remains in that state for the next 20 seconds. Then it spends the next 15 seconds in the state of learning. If these two states are summed together with the *Hello* time of 2 seconds, the final time is 52 seconds. The difference in the measurement results is therefore high because the switches have to learn the topology, which is not necessary in case of SDN and this is an additional advantage.

VI. CONCLUSION

The software-defined network includes the architecture which can be described as dynamic, economic and adjustable which makes it ideal for the dynamic nature of today's applications. Separation of the control planes and data planes allows directly programmable network control and separation of the low-layer infrastructure for the purpose of applications and network services. SDN offers centralized view onto the network, providing the controller with SDN so that they can operate as control planes, which makes them a strategic control point within the SDN network. It communicates with switches/routers by using the southbound API, and with applications by using the northbound API. The centralized, programmable SDN environments are easily adjustable to the variable needs of the company. The key advantages of SDN are agility and flexibility due to its separated architecture. SDN allows the organizations to quickly develop new applications, services and infrastructures in order to satisfy the variable business goals, flexible selection and operation of the network. Implementing the SDN solution requires good planning. Organizations should have clear idea about the advantages that are planned to be achieved by implementing SDN. In many cases, software-defined solution does not have to look different from conventional network, and SDN solutions are different depending on the manufacturers.

During network configuration there are substantial differences that could be noted between SDN and the conventional network. In conventional network every switch must be configured separately, which requires more procedures and time. If the network consists of a thousand network devices and hosts, which is very common today, every switch has to be individually configured according to the current traffic flow

and its changes, which additionally increases the number of procedures and the needed time.

Within the SDN solution, all of the listed procedures, from switch configuration and learning of new topology, are performed by SDN controller from one centralized point in the short time of the first connection. This concept is the main advantage of the SDN network in relation to conventional networks. The switch must be connected onto the SDN controller and the rest is performed by the controller itself. The advantage of learning of the entire topology and the view onto the entire network is shown on the basis of the results of Scenario 2 where the difference of the packet loss is high in regard to a conventional network.

REFERENCES

[1] B. Yeong Yoon, S.M. Kim, J.H. Lee: "Transport SDN Architecture for Distributed Cloud Services", The 12th International Conference on Optical Internet Proceedings (COIN), IEEE, Jeju, South Korea, 2014, pp, 1-2.

[2] S. Zhang, C. Kai, L. Song: "SDN based uniform network architecture for future wireless networks", Conference proceedings of International Conference on Computing Communication and Networking Technologies (ICCCNT 2014), Hefei, China, 2014, pp. 398-402.

[3] G. Sun, G. Liu, Y. Wang: "SDN architecture for cognitive radio networks", Conference proceedings of 1st International Workshop on Cognitive Cellular Systems (CCS), IEEE, Duisburg, Germany, 2014, pp 56-60.

[4] Z. Qin, G. Denker, C. Giannelli, P. Bellavista, N. Venkatasubramanian: "A Software Defined Networking Architecture for the Internet-of-Things", Conference proceedings of Network Operations and Management Symposium (NOMS), IEEE, Krakow, Poland, 2014, pp. 1-9.

[5] V.R. Tadinada: "Software Defined Networking: Redefining the Future of Internet in IoT and Cloud Era", International Conference on Future Internet of Things and Cloud (FiCloud), IEEE, Barcelona, Spain, 2014, pp. 296-301.

[6] H. Huang, J. Zhu, L. Zhang: "An SDN_based management framework for IoT devices", 25th IET Irish Signals & Systems Conference 2014 and China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014), IET Limerick, Ireland, 2014, pp. 175-179.

[7] OpenNet, available at: http://github.com/dlinknctu/OpenNet (25.06.2015)

[8] Mininet emulator, available at: http://www.mininet.org (17.06.2015)

[9] Ns-3, available at: https://www.nsnam.org/news/release-3-1/ (17.06.2015)

[10] EstiNet, available at: http://www.estinet.com/ (20.06.2015)

[11] M.C. Chan, J.X. Huang, T. Kuo, L-H. Yen, C-C Tseng: "OpenNet: A Simulator for Software-Defined Wireless Local Area Network", IEEE Wireless Communications and Networking Conference, Istanbul, Turkey, 2014, pp. 3332-3336.

[12] M.J. Todorović, N. D. Krajnović: "Simulation Analysis of SDN Network Capabilities", 21st Telecommunications forum TELFOR 2013, Serbia, Belgrade, 2013, pp. 38-41.

[13] ONF White Paper: "Software-Defined Networking: The New Norm for Networks", Open Networking Foundation, Palo Alto, CA, USA, 2012

[14] Cisco Systems: "Software-Defined Networking: Why We Like It and How We Are Building On It", White Paper, Cisco Systems, Inc., 2013

[15] S. Agarwal, M. Kodialam, T.V. Lakshman: "Traffic Engineering in Software Defined Networks", International Conference on Computer Communications (INFOCOM), Conference proceedings of INFOCOM, IEEE, Turin, Italy, 2013, pp. 2211-2219.

[16] M.R. Nascimento, C.E. Rothenberg, M.R. Salvador, C.N.A. Correa, S.C. De Lucena, M.F. Magalhaes: "Virtual Routers as a Service: the RouteFlow approach leveraging Software-Defined Networks",

Proceedings of the 6th International Conference on Future Internet Technologies, New York, NY, USA, 2011, pp. 34-37.

[17] Graphical Network Simulator 3, available at: http://www.gns3.com (20.06.2015)

[18] A. Pušeljić: "Analysis of Characteristics and Application of Software Defined Networks", Master thesis, Faculty of Transport and Traffic Sciences, University of Zagreb, Zagreb, 2015, unpublished.

[19] OpenDaylight Controller, available at: http://www.opendaylight.org (02.07.2015)

[20] Distributed Internet Traffic Generator, available at: http://traffic.comics.unina.it/software/itg (25.06.2015)

Massive Machine-Type Communications: An Overview and Perspectives Towards 5G

Ivan Jovović*, Ivan Forenbacher*, Marko Periša*

* University of Zagreb, Faculty of Transport and Traffic Sciences,
Department of Information and Communication Traffic,
Zagreb, Republic of Croatia, EU
ivan.jovovic@fpz.hr, ivan.forenbacher@fpz.hr, marko.perisa@fpz.hr

Abstract– **This paper provides an overview of the main characteristics and development of the new generation of mobile communication systems known as Next Generation Mobile Network (NGMN). As predicted, current exploitation of the fourth generation mobile communication systems (4G) will reach 50 billion connected devices by 2020. Expectations are that its successor, currently developing fifth generation (5G), is going to be operable by 2020. This new generation of mobile communication systems tends to become a technology platform that will enable the development of new applications, business models, industries, such as massive machine-type communications. This will be possible primarily through the creation of acceptable ecosystem that could provide a massive machine-type communication using a single platform based on the Internet of Things (IoT) concept. NGMN enables the integration of all so far known and used machine-type communications, creating an environment of smart cities and a fully networked society under the new concept of Internet of Everything (IoE). However, such network also poses specific performance requirements reflected through higher transmission speeds, higher data volumes, reduced energy consumption, higher quality of service, and growth in the number of services and users of currently deployed mobile generation.**

Keywords- Next Generation Mobile Network, 5G, Machine Type Communication, Internet of Things, Internet of Everything

I.   INTRODUCTION

Nowadays, mobile communications have a significant impact on the society and are an important factor in economic development and computerization of society. After a number of generations and the current commercialization of 4G mobile communication systems (MCS) it is expected that the upcoming 5G is going to be operable by 2020 [1], [2], [3] [4]. Up to this period, 5G is emerging as one of the main fields of research and development [4], [5]. Predictions [6], [7], suggest that 5G will contribute to creating a fully mobile and connected society, which will result with a stronger socio-economic progress of the community.

Over the past decades, the development of information and communication (IC) and computer technology has resulted in universal computerization of society, including the application of various telematics solutions. Telematics solutions have found their purpose primarily in the field of traffic and transport, through the implementation of the Intelligent Transportation Systems (ITS) [8]. However, telematics solutions are experiencing continuous increase of application fields and various users, including the large number of smartphone users, owners of the digital gas meters with the possibility of remote control, owners of smart vehicles, owners of smart homes, telemedicine service users, etc.

More recently, no less important are daily use of technologies such as communication between machines (M2M, Machine to Machine), the devices (D2D, Device to Device) or the vehicles (V2V, Vehicle to Vehicle) and the use of associated services and applications [9]. The goal is to consolidate all so far known and used machine-type communications (MTC, Machine Type Communication) through a single infrastructure. MTC technology is based on the idea that machines have a growing value proportional to the number of the networked units [10]. This would result in the concept of IoE and the possibility of creating a smart cities environment and a fully networked society by simple increasing the number of networked machines [11], [12], [13]. Currently, total number of all existing networked machines is hundreds of millions, and the annual growth rate is around 25%. As a result, it is expected that by 2020 the total number of networked machines (equipment, vehicles, goods, etc.) will reach 50 billion [7], [9]. Therefore, network operators will be able to expand their business activities, service portfolios and increase revenues. This technology, combined with the existing ones, has a high potential for the development of future applications. Authors in [6], [7] defined six fields of application which will result with an increase of machine-type communication in a few years from now: (1) automatization and the structural control of buildings, (2) transport and logistics, (3) health, (4) public safety and supervision, (5) monitoring of environmental and utility services, and (6) monitoring of power plants and electric energy distribution.

The main objective of this paper is to provide an overview of the main focus fields of development and implementation of mMTC communication that will be possible with the advent of 5G networks. Authors find this field very important in the upcoming time period of 2015-2020. Until then, complete MTC communication (MMC, Massive Machine Communication / mM2M, massive Machine to Machine communication / mMTC, massive Machine Type Communication) will pave the way for a significant increase of new networked machines, equipment and / or vehicles providing the new innovative services and applications. Consequently, as predicted, in a ten year period from now, this will result with the creation of new ecosystem based on the mentioned IoE concept [14], [15].

II. MAIN CHARACTERISTICS OF THE FIFTH-GENERATION OF MOBILE COMMUNICATION SYSTEMS

With every new MCS generation it was possible to experience two times higher speeds of data transfer than it was in the previous generation. In NGMN system, i.e. 5G MCS, one of the main requirements are the increase of data speeds and capacity with the significant reduction of latency level. Figure 1 shows the comparison of latency level between 4G and 5G networks [16].
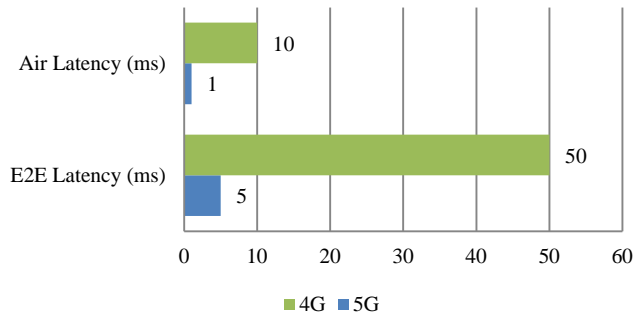
Figure 1 - Latency levels between 4G and 5G systems [16]

The integration of new services and applications is just as important as increasing the speed of data transfer and / or reducing the level of latency. 5G will be the MCS that uses its architecture and functionalities to enable the full implementation of services based on the IoT concept. This will result with a completely new communications between different types of machines (M2M, D2D, V2V, etc.) regardless of whether they're mobile, far apart or interconnected with different IC access technologies. With the advent of new services and the increase of the number of terminal devices, 5G MCS will increase the range of access requests for a mobile connection to the Internet network (MBB, Mobile Broadband) [17].

Required transmission speeds vary based on the volume of the transferable data. Figure 2 provides an overview of the range of data transmission speeds between 4G and 5G MCS. files.
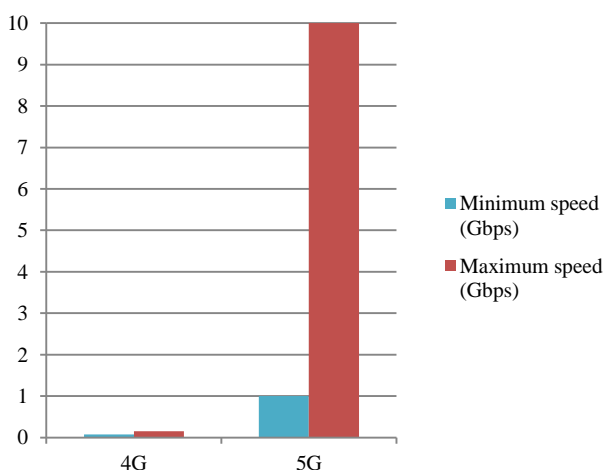
Figure 2 - Comparison of data transfer speeds between 4G and 5G systems [16]

The range starts from a very low level for sensor systems to very high level for the transfer of Ultra High Definition (UHD) video These requirements are reflected in the expected level of latency that will have to be low, as for security applications (emergency services, e-call about a car accident, alarm systems e-maintenance, etc.). However, there will also be services and applications that will be latency agnostic [18]. In addition, the size of the packets will vary from small to large, depending whether it's a smartphone applications or file transfer application, respectively. MCS 5G will enable ubiquitous provisioning of access to a wide array of services and software solutions, thus affecting daily routines and allowing continuous progress of overall society and a significant reduction in energy consumption [19]. Figure 3 shows a comparison of the main performance parameters of the 4G and 5G networks.
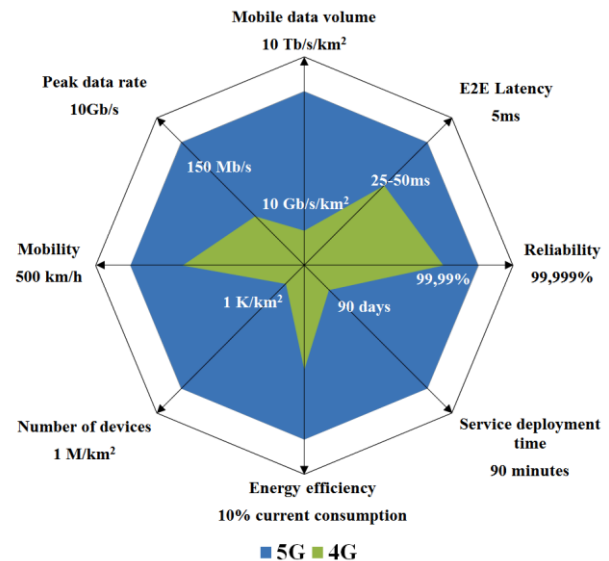
Figure 3 - The range of minimum expected requirements between 4G and 5G systems [4]

Energy-efficiency is an important issue to solve with the advent of 5G network. The drastic increase of the infrastructure by 2020 may easily result with exceeding acceptable level of energy consumed. This is because the biggest part of energy consumed comes from hardware and only smaller part from antennas [20]. To avoid any network failures due to energy depletion 5G will have the ability to shut down certain parts of its infrastructure on demand without losing the required coverage and capacity [21].

Compared with today's networks, the greater efficiency of radio spectrum and application of mMIMO (massive Multiple Input Multiple Output) will increase the capacity resulting with an average mobile data speeds up to 1 Gbps [16]. By reducing the latency to 1 ms, it will be possible to send a larger amount of multimedia content. Because of these new features, it is expected to connect potentially one trillion of terminal devices that will have smaller dimensions and greater battery autonomy. In order to assure a certain level of quality of service, reliability of service will amount to 99.999% [3], [4], [22], [23], [24]. Additionally, 5G will enable higher level of privacy for business and private users, and data protection.

5G systems will bring significant improvements and benefits to for users and for network operators all the main stakeholders (Figure 4). The network operators' capital expenditure will be lower due to the cost-effectiveness of IC equipment [25], [26]. This is because the future core network of the 5G system will have the ability to dynamically configure the characteristics of communication hardware according to different standards. At the same time, there will be an integration of different technologies such as nanotechnology and Cloud communications, [1], [27].
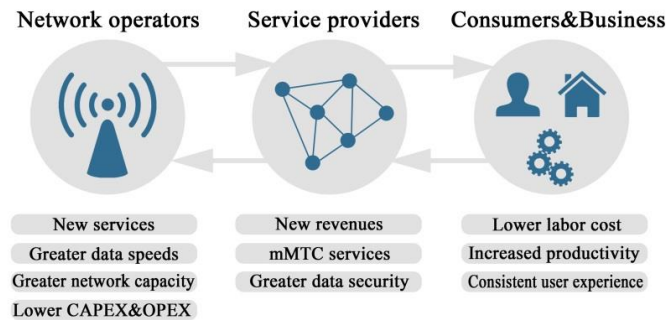
Figure 4 - Summary of main benefits for 5G stakeholders

Furthermore, it is expected that the IoT concept and 5G network will penetrate into various industries where they can contribute to interconnection and integration of different industrial facilities, medical equipment, vehicles, etc. [2]. In this way, this upcoming trend will meet required demands (which are impossible on 4G technologies to such an extent) for diverse services, including production services, medical services, transportation and other industries, thus resulting with a completely new concept and ecosystem based on the IoE paradigm.

This new ecosystem will be completely end-to-end 5G-based which will provide completely mobile and interconnected society. This, in turn, will enable to create values and benefits towards all relevant stakeholders through existing and new use-case scenarios [2].

III. PERSPECTIVES TOWARDS MASSIVE MACHINE-TYPE COMMUNICATIONS AND 5G

5G MCS will enable a complete implementation of the IoT concept, along with the advance of all machine-type communications through a single infrastructure (a prerequisite for MMC communications). An increasing number of end-devices will make maintenance and planning of today's telecommunications networks expensive and complicated. During the same period, an increase in the number of terminal devices will increase the volume of generated data traffic up to ten times [16].

A. Internet of Things concept

IoT is the concept of the information and communication network, where objects ("things") from diverse environments are mutual connected into a single large-scale network based on the Internet Protocol (IP). It's the basis for the development of smart environments such as smart homes, roads, factories, cities, etc. As a result, all these connected objects are part of a single converged ecosystem. While the term IoT was first introduced in 1999 it became very popular during the past several years. It was introduced formally in 2005 with the release of the IUT-T (International Telecommunication Union - Telecommunication Standardization Sector) report [28]. IoT consists of smart machines interacting with other machines, objects, things, environment and infrastructure [29]. Therefore, M2M communication it's often associated with MTC IoT provides remote access to various interconnected devices such as things (machines, computers, devices, sensors, products, etc.), systems (business applications, support systems, analytical systems, data warehouses, control systems, etc.), and people (customers, employees, partners and customers) [30].

IoT usually integrates the sensor data with analytics and business applications to improve productivity, service, and increase production and market share [6]. Some of the main challenges that accompany its implementation include security, privacy and trust, managing heterogeneity, limited network capacity, managing large amounts of information and processing large amounts of data in order to provide useful information / services. While IoT will reach its full potential during the next five years it is considered as a transitional technology towards Internet of Everything.

B. Internet of Everything Concept

Internet of Everything (IoE) includes interconnection of people, objects, things, data and processes. Its usefulness will result from the impact obtained by interconnecting people, processes, and data over IP network (Figure 5).
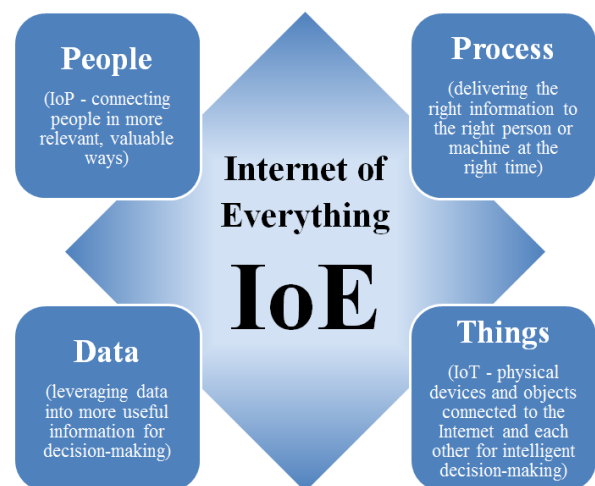
Figure 5 - Entities forming Internet of Everything concept [12]

Figure 6 shows the evolution of Internet-of-X concept starting with Internet of People (IoP, connecting people), followed by IoT concept, which will finally result with interconnection of everything as a part of IoE concept.

There is extraordinary potential of the IoE concept. According to the research conducted by Cisco Systems Inc. Corporation, 99.4% of physical objects that will one day be a part of the IoE concept aren't currently interconnected [31]. Not only that these objects (things) do not make the connection, but they aren't even a part of the IoT world. This is primarily because there are no such services that would result with the interconnection of these devices.
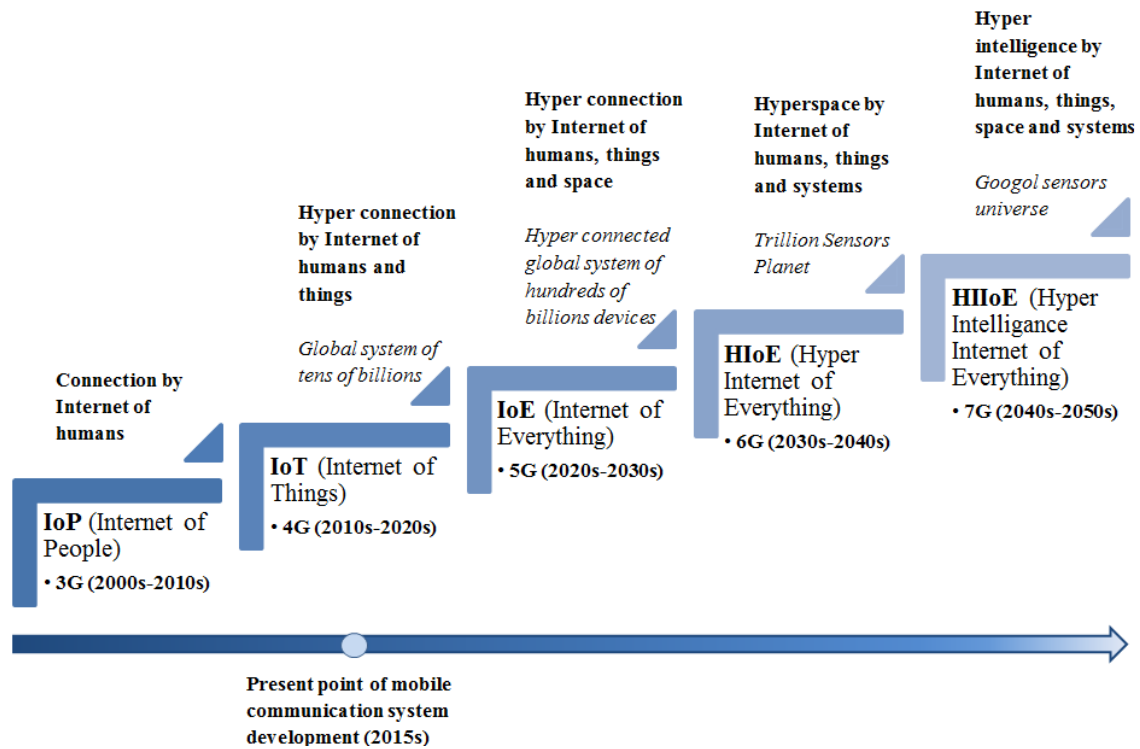
Figure 6 - The evolution of mobile communication systems and supporting services in the period from 2000 to 2040 [5]

Development of the 5G MCS will have an impact on economics, especially on the private sector. This is because of the simultaneous development of 5G objects, services and applications. Predictions claim that the interconnection of the currently unconnected objects may result with a total profit of $19 trillion; only private sector will make a profit of $14.4 trillion [31]. Connection of the objects will most commonly be performed through 5G MCS with the use of full communication functionality of machines such as the MMC, People to Machine communications (P2M), V2V, V2I, direct D2D (dD2D), etc.

C. Massive Machine-Type Communication

MMC communication will enable the connection of tens of billions of IP-based devices through 5G MCS. For example, as shown on Figure 7, 5G will be a systematic part of the Smart Cities in which 5G services and applications will have an impact on smart networked households (Domotics), smart/intelligent vehicles, tele surgery, fun, and time-critical applications that require an immediate reaction [32].

MMC concept includes a set of radio ICT and techniques, thus enabling the expected growth rate in the number of terminal devices and related services and applications. These access technologies are divided into three types [17], [3]:

- Direct access, terminal devices connect and communicate via direct access node
- Aggregation point access
- MTC communication between devices

The development of ICT and related services is gradually changing complete society in the last 20 years as well as the habits of each individual. However these are only the foundations for social changes which will follow after the implementation of future IC networks. For example, more than a third of the world population is nowadays constantly connected to the Internet; smartphones are becoming more and more popular and the development of mobile sector is expecting 4 billion users by 2017. NGMN will be the first complete system to fully implement converged IC network in which wired and/or wireless communications technologies will be using a single infrastructure [2]. Described 5G IC infrastructure will be the base for the development of fully interconnected society.
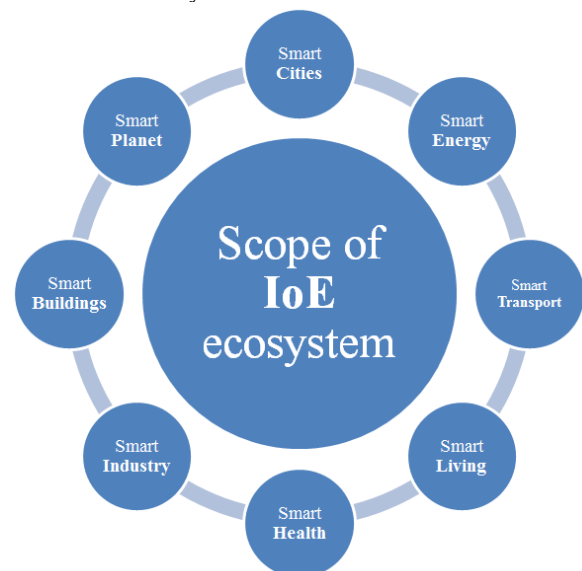
Figure 7 - Examples of application of Internet of Everything concept based on 5G

The adoption of these services and applications is part of the IoT / IoE ecosystem. In particular, the 5G era will be characterized with a fully interconnection of all things. Mobile network services, which were until recently limited only to smartphones, are gradually introduced to devices such as smart watches and sunglasses. In near future, each person will have several such devices.

D. Direct Device to Device Communication

As previously described, the MMC communication and the supporting IoT / IoE services provided via 5G platform will tend toward the creation of a fully networked and connected society. In a fully connected ecosystem enabled by 5G, the important role will play the ability to connect objects via D2D, i.e. dD2D technology [33]. D2D communication is implemented within the 4G MCS but it's not widely operable, mainly because of its still limiting performance (level of latency, capacity, data rate, the level of confidence, etc.). Consequently, D2D will not become a part of everyday operable technologies primarily because due to these limitations when compared with its direct competitor - V2V solutions based on VANET (Vehicular Ad-Hoc Networks) technology.

Although the VANET solutions have their disadvantages, they are developed for the specific environment and their usability in V2V communication is currently more prominent. However, the concept of NGMN and earlier NGN (Next Generation Network) system are designed for all-IP communication via single infrastructure. This will result with the convergence between different access and transmission technologies. Therefore, it was decided that one part of MTC communication will be enabled by dD2D technology over 5G [34]. Unlike the original D2D concept, dD2D will allow the planning and implementation of V2V communication within the IoE ecosystem. As mentioned earlier, with the development and implementation of 5G, several V2V requirements will be met, such as low latency of 1ms, default 99.999% reliability of communications, high availability and dependability and low failure rates [22], [3].

IV. Conclusion

The main objective of this paper was to present an insight into future developments, possible implementation and key features of the massive machine-type communication by introducing 5G. It is expected that in the next five years, mobile networks are going to experience a significant change compared to the current state. Transmission rate will be higher, the number of users and associated machines will increase by 10-100 times, and the traffic volume will increase up to 1000 times. It is necessary to consider the entire field as it opens new opportunities for network operators, service providers, and users within the value chain.

IC development and computer technologies have evolved simultaneously with the telematics systems and their field of application. Although the application of telematics solutions has its wider exploitation within the application segment of ITS, in the past decade their application is found in other fields as well. This is mainly due to the development of MTC Communications and later to the development of modern MCS

and wider use of the accompanying smartphones. Virtual connection of all types of devices such as, home appliances, vehicles, industrial machinery, vending machines, computers, etc. is becoming easier. The development of the IC networks and their environments, such as the MCS and WLAN with reduced size, high performance, and low-cost communication hardware has enabled easier way of connecting almost everything from the everyday living environment. Today, MTC communication is possible even for devices that until a few years ago were not designed with the purpose of network communication, such as, air conditioners, gas meters, vehicles, televisions, and the list is still growing. MTC technologies such as M2M, D2D and V2V are becoming a reality and an everyday need through the concept of IoT which allows communication directly with and without human intervention (automatic process control, monitoring and collection of data between devices, providing services through devices, etc.). A significant number of things/objects will find their application in IoT segment through various fields of activity.

In order to globally realize a trend and a plan for development of ICT and related services, it is necessary to fully develop 5G MCS as a single platform for the successful provision of cutting edge IC and telematics services. Estimated completion of the development and start of the implementation is beyond 2020. To realize all the predictions of the technological development, 5G networks in relation to the current MCS will have to assure significant improvements, particularly the increase of capacity, multiple increases of data speeds, and a high level of reliability, availability and dependability with low levels of latency in the network. As a result, this will allow a massive MTC communication that will contribute to IoT concept which will become a part of everyday life and enable transition to IoE concept.

IoE concept will bring a significant change for the society. It will change the way people live and contribute to their higher quality of life. With such development of ICT, implementation of new 5G MCS, and provision of various new services, it's possible to make MMC homes, roads, vehicles, cities, countries, etc. an integral part of an instrumented world of widely distributed artificial intelligence in which almost everything becomes a communication unit (objects, things). Those machines in the common ecosystem interact with each other and solve complex tasks, often independently without any human intervention. Maintenance of buildings, bridges, roads, etc. will be raised to a higher level with the possibility of remote e-maintenance assisted with sensors and software as part of MMC and IoE world. Options for the application of modern telematics solutions are almost limitless and the development and the implementation of 5G MCS will create a platform for their exploitation.

The future studies should systematically analyse consumer preferences towards 5G services. These details may help operators and service providers expand their knowledge about the future market, thus reducing the uncertainty risk associated with the introduction of new services and products. This is an important issue in heterogeneous environment such as 5G ecosystem where technologically innovations are the key growth drivers.

ACKNOWLEDGMENT

REFERENCES

[1] IMT-2020 (5G), "5G Concept," IMT-2020 (5G) Promotion Group, China, 2015.

[2] NGMN Alliance, "NGMN 5G White Paper," Next Generation Mobile Networks Ltd, Frankfurt am Main, 2015.

[3] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine,* vol. 52, no. 5, pp. 26-35, 2014.

[4] EC / 5G PPP, "5G Vision / The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services," European Commission, Heidelberg, 2015.

[5] 5G Forum, "5G New Wave / Towards Future Societies in the 2020s," 5G FORUM, Seoul, 2015.

[6] European Commission, "Europe 2020 / A strategy for smart, sustainable and inclusive growth," Europe Union, European Commission, Brussels, 2010.

[7] METIS, "Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS final project report)," The METIS 2020 Project, Sweden, 2015.

[8] ETSI, "Intelligent Transport Systems," European Telecommunications Standards Institute, Sophia-Antipolis Cedex, 2014.

[9] Ericsson, "More than 50 billion connected devices," Ericsson AB, Stockholm, 2013.

[10] G. Lawton, "Machine-to-Machine Technology Gears Up for Growth," *IEEE Computer,* vol. 37, no. 9, pp. 12-15, 2004.

[11] D. Soldani and A. Manzalini, "Horizon 2020 and Beyond: On the 5G Operating System for a True Digital Society," *IEEE Vehicular Technology Magazine,* vol. 10, no. 1, pp. 32-42, 2015.

[12] J. Bradley, L. Buckalew, J. Loucks and J. Macaulay, "Internet of Everything in the Public Sector / Generating Value in an Era of Change," Cisco Systems Inc., San Jose, 2014.

[13] D. Boswarthick, O. Hersenet and O. Elloumi, M2M Communications: A Systems Approach, Oxford: Wiley-Blackwell, cop., 2012.

[14] Q. C. Li, N. Huaning, A. T. Papathanassiou and W. Geng, "5G Network Capacity: Key Elements and Technologies," *IEEE Vehicular Technology Magazine,* vol. 9, no. 1, pp. 71-78, 2014.

[15] L. Gavrilovska, V. Rakovic and V. Atanasovski, "Visions Towards 5G: Technical Requirements and Potential Enablers," *Wireless Personal Communications,* pp. 1-27, 2015.

[16] Samsung, "5G Vision," Samsung Electronic., Ltd., Gyeonggi-do, 2015.

[17] J. F. Monserrat, G. Mange, V. Braun, H. Tullberg, G. Zimmermann and O. Bulakci, "METIS research advances towards the 5G mobile and wireless system definition," *EURASIP Journal on Wireless Communications and Networking,* vol. 53, no. 1, p. 20, 2015.

[18] S. Hossain, "5G Wireless Communication Systems," *American Journal of Engineering Research,* vol. 2, no. 10, pp. 344-353, 2013.

[19] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," *IEEE Instrumentation & Measurement Magazine,* vol. 18, no. 3, pp. 11-21, 2015.

[20] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck i M. Debbah, »Energy-Efficient Power Control: A Look at 5G Wireless Technologies,« *IEEE Transactions on Signal Processing,* pp. 1-28, 2015.

[21] R. L. G. Cavalcante, S. Stanczak, M. Schubert, A. Eisenblaetter i U. Tuerke, »Toward Energy-Efficient 5G Wireless Communications Technologies: Tools for decoupling the scaling of networks from the growth of operating power,« *IEEE Signal Processing Magazine,* svez. 31, br. 6, pp. 23-34, 2014.

[22] IMT-2020 (5G), "5G Vision and Requirements," IMT-2020 (5G) Promotion Group, China, 2014.

[23] K. Moskvitch, "Tactile Internet: 5G and the Cloud on steroids," *Engineering & Technology Magazine,* vol. 3/2015, pp. 48-53, 2015.

[24] C. Shanzhi and Z. Jian, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Communications Magazine,* vol. 52, no. 5, pp. 36-43, 2014.

[25] J. Sanchez, I. G. Ben Yahia, N. Crespi, T. Rasheed and D. Siracusa, "Softwarized 5G networks resiliency with self-healing," in *1st International Conference on 5G for Ubiquitous Connectivity (5GU), 2014*, Akaslompolo, 2014.

[26] Ericsson, "5G Radio Access," Ericsson AB, Stockholm, 2015.

[27] S. Shorgin, K. Samouylov, I. Gudkova, O. Galinina and S. Andreev, "On the benefits of 5G wireless technology for future mobile cloud computing," in *First International Science and Technology Conference (Modern Networking Technologies) (MoNeTeC), 2014*, Moscow, 2014.

[28] IERC, »Internet of Things, Position Paper on Standardization for IoT technologies,« IERC / European Commission, Bruxelles, 2015.

[29] IERC, »Internet of Things, IoT Semantic Interoperability: Research Challenges, Best Practices, Recommendations and Next Steps,« IERC / European Commission, Bruxelles, 2015.

[30] ETSI, "TR 102 725 V1.1.1 / Machine-to-Machine Communications (M2M), Definitions," European Telecommunications Standards Institute, Sophia-Antipolis Cedex, 2013.

[31] J. Bradley, Reberger, C., A. Dixit and V. Gupta, "Internet of Everything: A $4.6 Trillion Public-Sector Opportunity / More Relevant, Valuable Connections Will Boost Productivity, Revenue, and Citizen Experience, While Cutting Costs," Cisco Systems, Inc., San Jose, 2013.

[32] O. Vermesan and P. Friess, "IERC, Cluster's Priorities & Rolling Work Programme Overview," European Research Cluster on the Internet of Things, Bruxelles, 2013.

[33] G. Fodor and S. Sorrentino, "D2D Communications – What Part Will It Play in 5G?," Ericsson, 10 07 2014. [Online]. Available: http://www.ericsson.com/research-blog/5g/device-device-communications/.

[34] Q. Jian, S. Xuemin, J. Mark, S. Qinghua, H. Yejun and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Communications Magazine,* vol. 53, no. 1, pp. 209-215, 2015.

Balancing Expectations to the Health Software Production Process Standard

Alpo Värri

Department of Signal Processing
Tampere University of Technology
Tampere, Finland
Alpo.Varri@tut.fi

Abstract – **This paper presents the stakeholder expectations to the new version of the ISO/IEC health software life cycle standard 62304. This software production standard is central to the medical device industry but the new version is expected to cover even more scope including also other health software than just regulated medical device software. This paper discusses how to balance the expectations of the law makers, regulatory bodies, software producers and users etc. Compared to the present version the new version should take more into account the special characteristics of developing low risk mobile health application software and cybersecurity while maintaining the endorsement of the regulatory bodies.**

Keywords – Health Software Development, Medical Device Software, Software Life Cycle Standard, ISO/IEC 62304, Software safety, Medical Device Directive

I. INTRODUCTION

Since the Therac 25 incident [1] in which the software controlling a radiation therapy machine was found responsible for the death and injuries of several patients, the industry and regulatory bodies have been aware of the importance of safety requirements on medical device software. That incident taught that it is not only testing that is important to assure the safety of medical software but also the process applied to the development of the software needs to be well-controlled.

The European Union regulates the production of medical devices, including medical software with the Medical Device Directive, MDD [2]. The requirements of this and other "new approach directives" can be fulfilled by adhering to so-called harmonized standards which are considered as the state-of-the-art of the methods in the area that they cover.

The "**ISO/IEC 62304:2006 Medical device software -- Software life cycle processes**" is the harmonized standard for medical software production in the European Union. It is also the generally used standard for this purpose in other continents. The standard covers software safety classification into three classes (A, B and C), software development process, software maintenance process, software risk management process, software configuration management process, software problem resolution process and quality management system issues. As it is already a rather old standard, the methods of software production have developed and requests to extend its scope to non-regulated health software have been expressed, it is time to revise the standard. The revision is somewhat complicated due to the many different expectations for the revised version from different stakeholders. This paper attempts to capture these expectations and find a suitable balance to the in order to reach an acceptable compromise.

II. STAKEHOLDERS AND THEIR EXPECTATIONS

A. Politicians, Law Makers and The General Public

The rightful requirement of the general public is that medical devices, including software are safe and that the private health information is stored securely and that it is protected from unauthorized access. The general public has the right to assume that the medical applications they buy and use function correctly and give them true information. This is possible only when the operational logic in the software is based on sound medical knowledge and the software does not make too many mistakes. The politicians recognize these rights and attempt to enforce them through legislation.

Medical device failures which result in patient injuries and media attention inspire the politicians to demand for tighter regulation and control of the medical device manufacturers. This may be the case even when the guilty manufacturer did not follow the existing regulation, the act which would have prevented the incident. The politicians are motivated by the assumed concern of the general public on the issue. Loud voices from the politicians influence the law makers to tighten the regulations.

Although the "*safety is the first priority*" is a very appealing slogan, the other side of the coin is often forgotten. The more demanding it is to fulfill the requirements of the regulations is, the more product ideas are not implemented as their costs will become too high for the potential customers to pay for.

B. Regulatory Bodies

The regulatory bodies implement the policies defined by national laws. Some of this legislation originates from international agreements or in Europe from the MDD [2]. The principle of the MDD is that the manufacturers follow the state-of-the art methods in ensuring the safety of the medical devices, including software-only medical devices. The state-of-the-art should be reflected in the standards that the regulators approve as the methods of fulfilling the requirements of the law. This creates pressure for the standardization bodies to maintain and update the relevant standards as the state-of-the-art develops in their field.

The regulatory bodies require that the development organizations manage the risks associated to the use of their products. In the medical devices domain the applied standard

for this purpose has been the ISO 14971 [3]. There exists also a technical report (IEC/TR 80002-1) which gives guidance in how to apply the ISO 14971 to medical device software [4].

The regulators' task is to make inspections to companies which produce medical devices. In these inspections, the regulators check if the companies comply with the standards they claim to comply with. For this reason the regulators wish that the standards are easily auditable.

In order to make it less costly to make useful medical devices available in worldwide markets the regulators co-operate within the International Medical Device Regulators Forum, IMDRF [5]. One of the work items of IMDRF is "Software as a Medical Device" (SaMD) of which a document has been produced [6]. Although the main purpose of the document is to harmonize vocabulary and provide criteria to classify SaMD into safety impact categories, it also specifies requirements to software design and information security with respect to safety.

Cybersecurity has become more and more relevant theme also to health software since the tendency is to connect more and more devices and software together in order to facilitate information communication and automation in health care. Some high visibility security breaks, for example to the wireless communication to pacemakers [7] has motivated the Food and Drug Administration (FDA) to issue guidance how to manage cybersecurity in medical devices [8]. This puts pressure to the revision of the ISO/IEC 62304 to address security issues more thoroughly in order to represent the state-of-the-art.

C. Health Care Organizations

Health care organizations expect that the medical devices including SaMD are safe, easy to use, efficient, support the work processes of the organization and are interoperable to avoid duplication in data entry. Still, the products should be affordable and the manufacturer should be responsive to feedback when errors are found. This calls for requirements for the software maintenance processes of the manufacturers into the standard ISO/IEC 62304.

Some advanced health care organizations are applying the ISO/IEC 80001 [9] family of standards to improve and manage the safety of the medical devices connected to their networks. In order to comply with this standard, the health care organization requires information about the software which is used in the systems connecting to its IT network. This forms a requirement to the ISO/IEC 62304 to make such information available to the customers in the documentation of the software.

Some health care organizations also develop software, for their own use. They should be able to comply with the requirements of the standard. Thus the standard should be neutral towards the organization which develops the software not assuming that the developer organization is a company.

D. Large Health Software Producing Companies

Large companies, particularly those producing software classified as a medical device have development processes adapted to the present version of the ISO/IEC 62304. They recognize the need for a software development standard approved by the regulatory authorities to represent the state-of-the-art in the field because the lack of such a standard would probably make it more resource demanding for them to prove the adequacy of their processes to represent the state-of-the-art. They are therefore interested in maintaining this approval status for the ISO/IEC 62304 as the first priority. Their second priority is that the changes introduced to the standard are as minimal as possible to minimize the need for re-engineering their development processes because the re-engineering has cost implications.

E. Small Health Software Producing Companies

Small companies producing health software have various degrees of readiness to comply with the ISO/IEC 62304 standard. Particularly the start-ups require guidance in setting up their processes to comply with the standard. The problem in the start-ups is sometimes that they are so busy producing their first viable medical software product that they only pay attention to the regulatory requirements in the later stages of the design process. With the ISO/IEC 62304 this is, unfortunately to them, too late. It is not possible to "decorate" the product with the regulatory compliance afterwards if the processes required by the ISO/IEC 62304 were not in place in the beginning as Fig. 1 shows.

The small companies wish that it is possible to comply even when there is only a small number of developers working in the company. This means that the amount of procedures and documentation should be manageable by a small organization. Sometimes the standard identifies a lot of roles and functions from the developing organization. It should be possible to combine some of these roles to a single person, not requiring independence of the roles when not absolutely necessary.

Unfortunately small companies have rarely the possibility to influence the requirements of the standard because they cannot afford their expert(s) to attend the meetings of the standards development revision teams. Therefore the needs of the small companies are not so well addressed in the revision work.

Many small companies are in the mobile health application business. The development of mobile applications uses often software modules from other developers. There are also frequent updates to these modules as well as to the medical applications themselves. Risk management is a challenge in this kind of an environment. This form of software development is somewhat problematic to the existing version of the ISO/IEC 62304 which assumes a certain rigidity of the design on the basis of which the implementation is built and tested. The better suitability of the ISO/IEC 62304 to the mobile application world consisting mostly of low risk products is therefore in the wish list of the small mobile health application companies.

F. Standardization Organizations

The standardization organizations expect that standards exists in relevant areas and that they are up-to-date. Therefore the status and relevance of the ISO standards should be checked every five years and revisions should be undertaken if necessary. For the ISO/IEC 62304 the parent IEC 62A

committee set the task to expand the scope of the standard from medical device software to health software covering not only regulated medical devices but also the unregulated health related software embedded in devices and the stand-alone health software running on general purpose computing hardware.

The recently announced goal of the ISO Technical Committee 215 Health Informatics is to work towards a family of standards which are compatible to each other. This means that the requirements of one standard are not in conflict with another standard. This calls for the ISO/IEC 62304 revision team to interact with the different working groups of the technical committee to find interfacing areas. One such area is the link between ISO/IEC 62304 and ISO/IEC 80001 series of standards. The work has actually already been started. One of the first steps is to agree on the vocabularies to be used in both standards. Another obvious area is in the data security area. In the IEC side, some compatibility to the IEC 60601 series of standards [10] should be maintained.

The business models of ISO and IEC are that they earn the most of their income by selling standards. This means that they are not willing to make the majority of their standards available for free. This limits the circulation of the ISO/IEC 62304 to those willing to pay making it somewhat more unlikely that, e.g. the medical mobile application start-ups have access to it as early as they should. Some freely available educational material of the new version of the standard should therefore made available - not, however, slowing down the sales of the actual standard.

III. FINDING THE BALANCE

It is important that a medical device software engineering standard exists which is approved by the regulatory bodies. Therefore it is important that the most important regulatory bodies are somehow involved in the ISO/IEC 62304 revision process. At minimum, the regulatory bodies should give feedback about the draft standard when it is circulated for comments.

Sometimes it is difficult to draw the borderline between regulation and standards. The standards revision groups should beware stepping on the regulators' side in drawing regulatory policies. For example, the standard drafters are tempted to define what kind of medical evidence is required to prove that a software-based medical analysis method works, but this is actually a regulatory task. Therefore the software engineering standard should discuss this requirement only in quite general terms.

ISO/IEC 62304:2006 refers normatively to ISO 14971 [3] as the method for risk management. While this standard remains as a valid source for medical device production risk management, there may be other risk management standards which provide comparable safety to the resulting software. One

example of those is the ISO 31000 series of standards [11]. These could potentially be used at least in the lowest risk categories by a software company which has previously produced software for other safety critical uses than medicine.

The amount of mobile medical applications is growing rapidly. It is becoming increasingly difficult to the regulatory bodies to keep track of them all and apply the existing rules to them all. If there exited a way of providing compliance with a set of reduced requirements of the present ISO/IEC 62304 and a large proportion of the mobile health application developers would seek that compliance as a quality label for their product, this might improve the overall quality of these applications. There should not, however, be a quick way out in producing software for the life critical applications. The elaborate design process should still be required in this area as previously.

The ISO/IEC 62304 revision team should make a survey of the standards relevant to its field now that a major revision process has been started. Not only the health informatics and medical device sectors should be investigated but also the general field of software engineering should be studied. There are, however, practical limitations in taking into account all wishes for the revised version of the ISO/IEC 62304. The revision is carried out by experts who are often busy in their main work and participation to standards development is only a sidetrack to many. It is therefore difficult to survey all possible standards and other documents which are relevant to the standard under revision.

In principle, the commenting process of the standards should enable the incorporation of additional expertise to the standard document but in practice only a small proportion of the affected organizations invest resources in commenting on the draft standards. The national standardization bodies should have good contacts to the relevant stakeholders of the standard to activate them at least when there is the last chance to propose modification to the contents to the standard.

The revised ISO/IEC 62304 standard cannot go to details in all areas that it is expected to cover. For example, usability is an area with its own standards and ISO/IEC 62304 should refer to those. Particularly the usability standard ISO/IEC 62366 [12] is intended for the medical device area and should be preferred. Similarly, the data security domain should be covered by referring to one or more standards of that field. The recently revised ISO standard 27799 [13] is a good starting point. It should be noted, however, that not all the requirements of the ISO 27799 are applicable to all health software. For example, it is not feasible to expect that an electronic thermometer (containing embedded medical device software) for general practitioners' appointment should hold an elaborate back-up system for its data as suggested by the ISO 27799 standard. Another technical specification worth referring to is the ISO 14441 [14] regarding security requirements in electronic patient record system software.
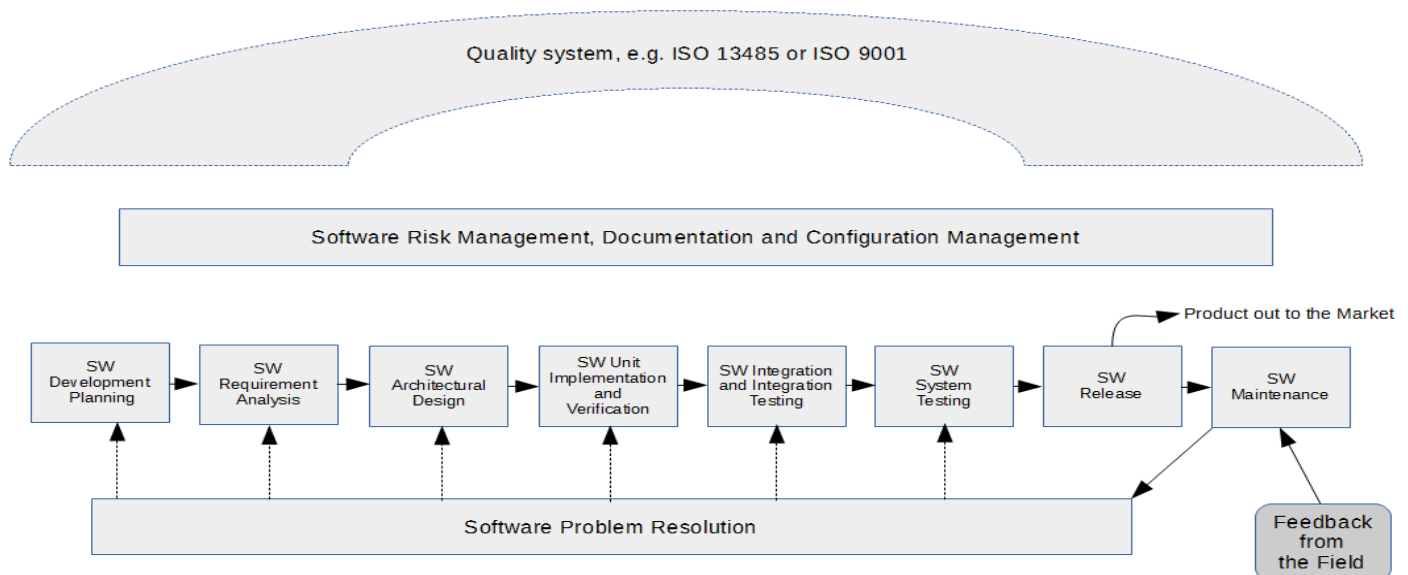
Figure 1.   The scope of the standard ISO/IEC 62304 extending throughout the life cycle of the development cycle of the software including its maintenance after release. The risk management, documentation and configuration management activities continue in all stages of the process. Feedback from the field, particularly an error report activates the software maintenance activities as a problem resolution process. Although the quality system according to a certain standard is not explicitly required, the standard requires that the producer of the software demonstrates the ability to produce high quality software.

In practice, the stronger emphasis on security requires the manufacturer to extend its risk management of cybersecurity risks. For example, instead of assuming that the communications channel is secure, the manufacturer needs to think: what if the channel were attacked and the message content altered; how could this be detected.

IV.   CONCLUSIONS

There are a lot of expectations to the revised version of the ISO/IEC 62304 medical software engineering standard. The regulatory endorsement of the standard must be maintained. The scope extension to all health software requires some new thinking in the principles of the standard recognizing that all potential users of the standard are no longer regulated medical device manufacturers but small mobile health software developers are also within scope. The interfaces to other standards need to be checked and cybersecurity requirements should also receive more attention in the new version.

ACKNOWLEDGMENT

REFERENCES

[1]   Therac-25. http://en.wikipedia.org/wiki/Therac-25

[2]   DIRECTIVE 2007/47/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 September 2007 amending Council Directive 90/385/EEC on the approximation of the laws of the Member States relating to active implantable medical devices, Council Directive 93/42/EEC concerning medical devices and Directive 98/8/EC concerning the placing of biological products on the market. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007L0047&rid=2

[3]   ISO 14971:2007 Medical devices - Application of risk management to medical devices.

[4]   IEC/TR 80002-1:2009 Medical device software -- Part 1: Guidance on the application of ISO 14971 to medical device software.

[5]   International Medical Device Regulators Forum, http://www.imdrf.org/

[6]   "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations. September 18th, 2014   http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf.

[7]   Barnaby Jack Could Hack Your Pacemaker and Make Your Heart Explode, June 23rd, .2013. http://www.vice.com/read/i-worked-out-how-to-remotely-weaponise-a-pacemaker.

[8]   U.S. Department of Health and Human Services, Food and Drug Administration, "Content of premarket submissions for management of cybersecurity in medical devices, Guidance for industry and Food and Drug Administration staff". October 2nd, 2014, http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM356190.pdf

[9]   IEC 80001-1:2010 Application of risk management for IT-networks incorporating medical devices -- Part 1: Roles, responsibilities and activities.

[10] IEC 60601-1 Series, Medical electrical equipment. www.iec.ch.

[11] ISO 31000:2009 Risk management - Principles and guidelines.

[12] IEC/ISO standard 62366, Medical devices - Application of usability engineering to medical devices. Edition 1, ISBN 2-8318-9313-5, 2007.

[13] ISO/IS 27799 Health informatics, Information security management in health using ISO/IEC 27002.

[14] ISO/TS 14441:2013 Health informatics - Security and privacy requirements of EHR systems for use in conformity assessment.

Fair value of the media sector companies quoted on the Warsaw Stock Exchange in Poland within 2006-2015 and their financial analysis

Rafał Parvi

Wroclaw School of Banking
Wroclaw, Poland
rafalp4@o2.pl

Abstract - **This paper examines share price of the companies listed on the WIG-MEDIA and their fair value between 2006-2015. Data from 2006 to 2015 were collected from the Stooq.pl (Polish portal of shares). Two hypotheses are tested: (1) value of the shares based on the market price; (2) value of the shares as the fair value of shares. Moreover, the financial condition of several companies with respect to the sector was examined with the selected financial ratios. In addition, the author expressed his view on the opportunities of the property companies market's WIG-MEDIA in 2015-2017.**

Keywords: foods, company, market value of shares, profit

I. INTRODUCTION

The modern media recipient, in spite of appearances, reads not only branch magazines, but also observes the reality outside the mainstream, that in on Internet portals such as Kwejk, Demotywatory, or reading comments on Onet.pl, and Gazeta.pl websites. These Internet services are the source of the young generation's voice. It is there, where we can observe behaviours of these people, who will take purchasing decisions in 9 years time. The young generation thinks in different terms, and will decide about their future in a way different than our one.

The challenge for the media market will be also demographic changes. The Polish society – by analogy, like these of the West – is getting old very fast. Media will have to reckon with not only audience outflow, but also with other patterns of consumption because young groups do not copy communication schemes of the older ones.

The media market is subject to dynamic changes. Primarily, the consumer and the direction of development of new technologies are changing. The media market is following them. The modern media recipient, in spite of appearances, reads not only branch magazines, but also observes the reality outside the mainstream, that in on Internet portals such as Kwejk, Demotywatory, or reading comments on Onet.pl, and Gazeta.pl websites. These Internet services are the source of the young generation's voice. It is there, where we can observe behaviours of these people, who will take purchasing decisions in 10 years time. The young generation thinks in different terms, and will decide about their future in a way different than our one.

The area, where changes are the most visible, is the technology. The consumer of the 21st century uses new technologies in full, but subjectively – he or she accepts these solutions, which give him or her something more. Alex Stil, the head of digital EMEA & global head of programmatic, Mindshare Worlwide, speaking of an important role of digital media in our lives, proved that the future belongs to big data [2].

Media will also change. Apart from the "mythic" passing from off-line to on-line, about which everybody talks about, the change of systems of access dividing the content into the basic and the premium one also awaits us. The process, which has been in progress for a very long time in spite of appearances, will be a great challenge for publishers in the coming years. The offer is being developed, but not only towards the on-line direction – thanks to digitisation of broadcasting, new channels quickly earning new watchers are being created [3, 5].

II. FAIR VALUE OF LISTED COMPANIES

Share price of the companies listed on the Stock Exchange should reflect also their fair value. The fair value can be defined in several ways. In view of the foregoing, the fair value is a value used repeatedly in accounting, and thus in Article 28 (6) of the Accounting Act of 29.09.1994 as "the amount for which a given asset component could be exchanged, and the liability could be paid on market transaction terms between interested and well-informed, unrelated parties.

In view of whether the price of shares that are quoted on the stock exchange corresponds to their fair value, should be found in the value alone, since, after all, the values may be diverse, like the value of a similar company, producing similar goods and operating in the same industry, will also be diverse for various reasons.

The subject of trade covers minority shares, and the main market participants are retail investors or minority institutional investors, thus the price of shares should reflect the fair value characterizing the liquid minority interest.

The value presented in this way seems righteous, which is confirmed by the premium paid by the investors who announce calls for subscription for shares and plan in this way the purchase of the controlling interest. Then the premium reflects the difference between the level of liquid minority interest and the level of controlling interest. Sometimes the level takes into account benefits resulting from synergy. An investor

purchasing the controlling interest in this way receives premiums that appear after taking over control of a company, in the form of funds, business management and making a number of strategic decisions.

The fair value of the share price should be determined in accordance with the idea of capital market, namely the market participants should have equal access to data, information and all messages concerning a given company [11, 13, 14]. However, the investors are divided into three groups:

a) People with access to the most closely guarded information that affects the price and the business value, namely those can the company's management board or shareholders,

b) Institutional investors with blocks of shares with simultaneous access to the company's management board,

c) Individual investors who have access to public information.

At this point, there should be no differences in particular groups, at least officially, however, it happens that a group that is closest to the company has information which can obviously change its value overnight or distort its value artificially, e.g. other data or informal data, and even fictitious data.

Worldwide we can also appreciate companies that treat individual investors seriously, namely provide them data concerning a company and treat them as equal and as strategic investors, with a large impact on a company in the present and in the future.

Transactions that are concluded on the Warsaw Stock Exchange relate to transactions between interested parties, namely a purchase or sale order should be submitted. Sometimes there are cases of wrongly submitted orders, but these are marginal orders that do not have a large impact on the transactions conducted on the Warsaw Stock Exchange.

Share liquidity is understood as a percentage quantity of shares in a company that are in the possession of minority shareholders. This is, at the same time, free float and the face value of such shares and the value of average daily trade in shares in a given company.

Thus, we may presume that if during a day a large trade in shares takes place and a greater part of shares remains in the hands of minor shareholders, the share price reflects their fair value. We cannot agree with the fact that at small or minimum trade the value of shares is reduced to its daily minimum and it would be its fair value. Then it is only a change in the share price to a lower one and it does not indicate its fair value, as the trade alone suggests that this is only a pure coincidence. Such an impact can be exerted by large shareholders, as they can, by using one block, decrease the share value, preventing smaller players from raising the price for one reason: the capital of smaller shareholders does not enable them to raise the share price.

There is one premise more to determine the fair value of share price. The investors are fond of investing in shares, namely they buy them as in the past they managed to earn on

them and they feel that presently the share price is ideal and reflects their fair value and will enable them to obtain fair dividend in the future.

Such a purchase or sale of shares can largely overestimate or underestimate the share value of a quoted company. Here the IT industry may serve as an example, namely shares in technological companies at the beginning of the new millennium, when shares in these companies were being purchased without any analysis in technical terms, but looking at their name and value, which was increasing overnight. In view of the foregoing, this led to excessively high business value above its fair value.

The share price should thus reflect the fair value of a company listed on the Warsaw Stock Exchange. For the value of these companies be fair, the market must make available to all investors information regarding companies listed on the Warsaw Stock Exchange [13]. The shareholders should be treated equally; therefore we cannot distinguish majority shareholders as those who should have information unavailable for minority shareholders. First of all, shares should be liquid securities, therefore they should be in free float and have real-time transferability, namely at any moment and at any time during the office hours of the Warsaw Stock Exchange on a business day.

The WIG-MEDIA index, presented in Figure 1, shows that from 2009 to the first quarter of 2011, the WIG-MEDIA sector's companies in Poland showed an upward trend in their values. However, from the second quarter of 2011, a significant downward trend can be noticed, until 2012 showed an upward trend in their value too, in which the augmentation and consolidation at the level of 4206.15 points on the index take place. The values reported on 18.09.2015 reflect the lateral trend and confirm it. However, the market values do not reflect their fair value.

Source: Stooq.pl

Figure 1.   WIG-MEDIA in the period from 01.2006 to 09.2015 [12].

III.   ANALYSIS AND VALUATION OF THE WIG-MEDIA SECTOR COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE (WSE) IN POLAND

In the WIG-MEDIA sector, three companies, the values of which approach to their maximum prices on 18.09.2015, and they are Wirtualna and SMT. However, one company does not show its maximum or even fair value, though it can show the

net profit and good financial condition, and this is COMPERIA and CYFRPLSAT. Some companies were overvalued by even 95%.

These companies are AGORA, HUBSTYLE and PMPG. The flagship companies, such as 4FUNMEDIA, ATMGRUPA, K2INTERNET and KINOPOL, stay ahead with the best results, as shown in Table 1-2

TABLE I. THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 18.09.2015.

Name	Average rating	rating
4FUNMEDIA	4.0/5.0	AAA
AGORA	3.0/5.0	AAA
ATMGRUPA	3.5/5.0	AA+
COMPERIA	4.0/5.0	AAA
CYFRPLSAT	4.0/5.0	B+
HUBSTYLE	5.0/5.0	AAA
K2INTERNT	4.0/5.0	BBB
KINOPOL	4.0/5.0	AA+
MUZA	2.5/5.0	B+
PMPG	2.5/5.0	CCC-
SMT	4.0/5.0	AAA
WIRTUALNA	No data	No data

Source: own development based on the data of the Warsaw Stock Exchange.

In the WIG_MEDIA sector, it can be seen that PMPG has a overestimated value in relation to the maximum value by more than 99%, but the WIRTUALNA and the SMT are presented as companies of a high value in the sector.

TABLE II. THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 18.09.2015.

Name	Current price PLN	Maximum price PLN from the beginning of the stock exchange quotation
4FUNMEDIA	4.34	21.6
AGORA	12.01	114.50
ATMGRUPA	3.19	9.95
COMPERIA	19.10	30.42
CYFRPLSAT	22.87	27.66
HUBSTYLE	0.89	12.15
K2INTERNT	10.80	24.54
KINOPOL	14.02	28.05
MUZA	3.44	19.30
PMPG	1.53	250.80
SMT	18.64	21.95
WIRTUALNA	37.23	37.60

Source: own development based on the data of the Warsaw Stock Exchange.

It is obviously a very good recommendation for future investors. These companies prosper properly and are quoted on the Warsaw Stock Exchange [1, 6, 7, 8, 9, 10]. In tables are nominal prices in PLN.

In table 3-4, the most important ratios presenting the financial condition of the WIG-MEDIA sector's companies were presented. The profit per share was generated in 10 examined companies. It shows that WIG-MEDIA invest a lot, what determines their value and the generated profit per share [2].

The price to the operating earnings shows the losses of the company, and this state of affairs was reported in the examined stock exchange quoted companies (4FUNMEDIA, AGORA, HUBSTYLE, K2INTERNET). The ATMGRUPA, COMPERIA, CYFRPLSAT, KINOPOL, MUZA, SMT, achieved a very high ratio, and the PMPG – an average one.

TABLE III. TECHNICAL EVALUATION OF THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 18.09.2015.

Name	P/OE (price/ operating earnings)	P/BV (price/ book value)	P/PF (price/ profit)
4FUNMEDIA	-1.67	0.62	1.02
AGORA	-235.65	0.53	0.53
ATMGRUPA	17.39	1.24	1.64
COMPERIA	21.44	1.68	1.74
CYFRPLSAT	7.96	1.53	1.50
HUBSTYLE	-6.78	0.96	1.23
K2INTERNT	-250.83	1.06	0.32
KINOPOL	13.98	4.31	2.52
MUZA	8.15	0.45	0.37
PMPG	3.19	0.68	0.30
SMT	11.20	2.74	0.79
WIRTUALNA	No data	No data	No data

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

Generally, the value of the index P/OE is a useful tool for determining an absolute share value referred to the operating profits. Using the operating profit instead of a net profit (the P/E index) allows for rejecting single events. Moreover, a net profit is easier to handle. Fewer companies incur losses at the level of an operating profit rather than a net value, which allows for a broader use of P/OE [4, 6, 7, 8, 14].

The P/BV index informs how the company's own capital is valued by the market at a given moment. A general interpretation of the index consists in the fact that the P/BV indices below 1 mean a low price of a company, whereas a value over 3 that a company is overrated [4, 9, 10].

The value of the P/PF index is expressed in the way that when the value of the index is lower, then the price for the purchased company's shares is theoretically lower too, which means that the enterprise is more attractive. It is used in order to demonstrate cyclic profits and losses of the analyzed companies (the income is much more stable than the company's profit) [4].

In contrast, analysing P/BV and P/PF, it should be noted that both the price to the book value and the price to profit demonstrate that five companies exemplary operate on the market and have a value of more than 1.0, and these are ATMGRUPA, COMPERIA, CYFRPLSAT, KINOPOL and SMT.

TABLE IV.   TECHNICAL EVALUATION OF THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 30.06.2015.

Name	Net profit (net loss) in thousands PLN	Depreciation in thousands PLN
4FUNMEDIA	597	520
AGORA	3841	30060
ATMGRUPA	3622	4446
COMPERIA	285	895
CYFRPLSAT	304500	0
HUBSTYLE	-7005	382
K2INTERNT	487	0
KINOPOL	5077	3258
MUZA	737	688
PMPG	No data	No data
SMT	2856	609
WIRTUALNA	9917	7433

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

TABLE V.   TECHNICAL EVALUATION OF THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 30.06.2015.

Name	EBITDA in thousands PLN	Assets in thousands PLN
4FUNMEDIA	1151	33483
AGORA	33669	1544586
ATMGRUPA	8381	308433
COMPERIA	1077	30564
CYFRPLSAT	853500	27141800
HUBSTYLE	-6409	18682
K2INTERNT	300	48759
KINOPOL	9005	108340
MUZA	1649	39802
PMPG	No data	No data
SMT	3726	155819
WIRTUALNA	23781	665110

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

Other companies do not significantly differ from the average values. Only AGORA, MUZA and PMPG is significantly below the thresholds and shows the negative values.

Table 5 presents the studies concerning, among others, the net profit, depreciation, EBITDA and assets of the WIG-MEDIA sector's companies.

According to the obtained values, it is clear that only HUBSTYLE showed a loss, which was confirmed by the previous ratios included in Table 3.

TABLE VI.   TECHNICAL EVALUATION OF THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 30.06.2015.

Name	Profit per share	Book Value per Share
4FUNMEDIA	0.147	7.000
AGORA	0.075	22.457
ATMGRUPA	0.043	2.567
COMPERIA	0.132	11.365
CYFRPLSAT	0.476	14.945
HUBSTYLE	-0.436	1.040
K2INTERNT	0.196	10.149
KINOPOL	0.256	3.262
MUZA	0.258	7.770
PMPG	No data	No data
SMT	0.253	6.645
WIRTUALNA	0.351	12.322

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

According to the book value per share and profit per share, it is possible to deduce that some companies are overvalued, and they are AGORA and CYFRPLSAT, and in the case of the 4FUNMEDIA, ATMGRUPA, COMPERIA, K2INTERNET, KINOPOL, MUZA company, they are undervalued. However, it is important not to follow this opinion because the values are only the book values [4], and the calculation of them is purely mathematical and financial. In the case of using the economic attitude and interpretation, it would occur that the companies do not have the fair value (table 6).

The profitability of the equity, as well as the profitability of assets is shown by ATMGRUPA, COMPERIA, CYFRPLSAT, KINOPOL, MUZA, PMPG, and WIRTUALNA, while 4FUNMEDIA, AGORA, HUBSTYLE, K2INTERNET and SMT does not have it. Therefore, according to the presented study, it is possible to observe that the flagship chemical concerns have the profitability and they are not threatened by any disturbance of the financial liquidity (table 7).

TABLE VII.   TECHNICAL EVALUATION OF THE PROPERTY WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 30.06.2015.

Name	ROE	ROA
4FUNMEDIA	-23.21	-17.65
AGORA	-2.58	-2.17
ATMGRUPA	6.01	5.09
COMPERIA	29.96	19.52
CYFRPLSAT	1.88	1.38

HUBSTYLE	-1.96	-1.70
K2INTERNT	-0.12	-0.08
KINOPOL	7.69	6.15
MUZA	2.00	1.14
PMPG	5.26	2.20
SMT	-0.03	-0.02
WIRTUALNA	20.34	14.54

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

Return on equity (ROE) measures the rate of return for ownership interest (shareholders' equity) of common stock owners. It measures the efficiency of a firm at generating profits from each unit of shareholder equity, also known as net assets or assets minus liabilities. ROE shows how well a company uses investments to generate earnings growth.

The return on assets (ROA) shows the percentage of how profitable a company's assets are in generating revenue.

Other companies have shown a substantial profit which was generated in 2015, and they were 12 companies, however, in the PMPG company the value could not be calculated and read due to the lack of data. Currently, the value of companies significantly deviates from the maximum value achieved a few years ago. The only exceptions are WIRTUALNA and SMT, which achieved almost the maximum value in its history. Other companies have the value less than 99% of the maximum one (Table 8).

However, the fair value which should be reflected by the share prices of the examined companies significantly differs from the calculated value, which was presented in Table 8. In some cases, it is even 99% of the current value. The fair value is considerably higher than the current value of the examined companies.

TABLE VIII.   THE WIG-MEDIA SECTOR'S COMPANIES QUOTED ON THE WARSAW STOCK EXCHANGE IN POLAND AS OF 18.09.2015.

Name	Fair value	Deviation from the fair value in PLN
4FUNMEDIA	15.21	10.87
AGORA	80.32	68.31
ATMGRUPA	8.64	5.45
COMPERIA	22.67	3.57
CYFRPLSAT	25.96	3.09
HUBSTYLE	5.57	4.68
K2INTERNT	20.63	9.83
KINOPOL	22.61	8.59
MUZA	15.24	11.80
PMPG	50.74	49.21
SMT	21.65	3.01
WIRTUALNA	37.60	0.24

Source: own development based on the financial data of the companies quoted on the Warsaw Stock Exchange in Poland.

Deviation from the fair value in PLN = DevFV

DevFV = Fair value - current value.

IV.   CONCLUSION

The share price of selected companies of the food sector's companies quoted on the Warsaw Stock Exchange in Poland is significantly underestimated by the current financial situation in the world.

The flagship companies of the WIG-MEDIA sector achieve enormous profits, which was proved in the examination of ratios in last years and a net profit in 2015. CYFRPLSAT, KINOPOL, SMT and WIRTUALNA are an example of it.

The share price of the property WIG-MEDIA sector's companies quoted on the Warsaw Stock Exchange in Poland is significantly underestimated by the current financial situation in the world [11, 13, 15].

The fair value of the WIG-MEDIA sector's companies quoted on the Warsaw Stock Exchange in Poland should be reached within two years, that is up to 2017 because it is the right estimation of further fast development of the Polish WIG-MEDIA sector.

REFERENCES

[1] Czekaj J., "Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia", PWE, Warszawa, 2014, pp. 201-205

[2] GOBAN-KLAS T., „Media i komunikowanie masowe. Teorie i analizy prasy, radia, telewizji i Internetu", Warszawa, PWN 2009, pp. 230-250.

[3] GUEDES B. O., BART C., NICO C., Media alternatywne, Wydawnictwo Uniwersytetu Jagiellońskiego, 2012, pp. 124-156.

[4] Jajuga K., Jajuga T. „Inwestycje", Wydawnictwo Naukowe PWN, Warszawa 1996, pp. 34-57.

[5] KALL J., Reklama, Polskie Wydawnictwo Ekonomiczne, Warszawa 2010, pp. 134-145.

[6] Parvi R., "Analysis of companies of the energy sector based on an example of the companies quoted on the Warsaw Stock Exchange in Poland and their fair value"CER – International Scientific Conference, LONDON 2014, pp. 33-36.

[7] Parvi R., "The analysis of companies of the Polish fuel sector based on an example of PKN Orlen and Lotos companies versus walue of fuel prices and their impact on fuel sector management", ESD – International Scientific Conference Economic and Social Development, Zagreb Croatia 2014, pp. 52-60.

[8] Parvi R., "Investment funds and their financial effectiveness", Goce Delchev University Macedonia & THOMSON Ltd. Slovakia, EDIS - Publishing Institution of the University of Zilina, 2014, pp. 264-267.

[9] Parvi R., "Valuation of share and their fair value of the banking sector's companies quoted on the Warsaw Stock Exchange in Poland within 2008-2014", Goce Delchev University Macedonia & THOMSON Ltd. Slovakia, EDIS - Publishing Institution of the University of Zilina, 2014, pp. 76-81.

[10] Parvi R., "Share price of companies listed on WIG20 and their fair value". EDIS - Publishing Institution of the University of Zilina EIIC – 3 rd Electronic International Interdisciplinary Conference 2014, pp. 169-177.

[11] Pierce R., "Stocks, Options & Spreads, Infinity Publishing" 2004, pp. 110-115.

[12] Stooq.pl - http://stooq.pl/q/?s=wig_media&c=10y&t=l&a=lg&b=0, 18.09.2015.

[13] Veale S., "Stocks, Bonds, Options, Futures", Prentice Hall Press 2001, pp. 88-125.

[14] Withers H., "International Finance", Warszawa, Kraina Książek, 2012, pp. 25.

[15] Zaremba, A., „Giełda. Podstawy inwestowania", Warszawa, Helion, 2014, pp. 67-70.

Insight into Make and Model Recognition of cars - SVM and FHOG design

Stefan Badura, Roman Badura, Michal Mokrys

Independent Researchers, Zilina

Slovakia

baduras@gmail.com, badura4@gmail.com, mokrysm@itall.sk

Abstract – **Car type recognition, or** *Make and Models Recognition* **(MMR) of cars is very often solved task in current computer vision area. This task finds its deployment in many areas of human beings lives and therefore many researchers resorted to it. This paper proposes a deep insight into MMR problem. Effective combination of widely used methods can lead to a design, which is able to cope with unwanted changes in illumination, background noise with acceptable results. Which approach or configuration is the most suitable for MMR? Answering this question introduces solving certainly challenging task. This article presents experiments with approach, which examines Falzenschwalb's** *Histograms of oriented gradients* **(FHOG) and configurable** *Support Vector Machine* **(SVM) merged into single solution, which can bring more light into MMR problem.**

Key-words – *MMR, FHOG, SVM, feature, analysis*

I. INTRODUCTION

Vehicle detection and classification systems play essential role in variety of applications widely needed in humans being lives. Such applications are used in traffic management and monitoring, automated parking entrances in parking houses, driver assistance and control access systems. Other deployments are found both in public and commercial sphere. Targeted advertisement or surveillance for crime preventing and safety are nowadays demanding.

From existing solutions of processing vehicles images well known is the task of Automatic License Plate Recognition. Analyzing deeper the use cases of understanding images of cars, four main research directions can be highlighted, mostly solving localization or classification problems:

- License plate localization and recognition.

- Car identification in the image. The result information proposes simple output about existence of a car in proposed input image. Simple application is recognition cars from pedestrians or bikes and motorbikes.

- Car recognition based on car body (sedan, hatchback, bus, truck). Application finds its utilization in traffic monitoring, restrictions of access, etc.

- Car recognition based on manufacturer assigned name (Seat Leon, VW Golf 4th generation, VW Golf 5th generation\ldots). This task is the most advanced with high sensitivity on details, environment changes, rapid changes in manufacturer production and maintenance. This problem is also referred as Make and Models Recognition (MMR). Application can be found in car tracking and surveillance, advertisement and marketing, etc.

Referring to the previous list, this work proposes approach for the make and model recognition of a car (MMR), which described the last item. This task is interesting from several points of view: the tremendous level of details, instantly growing the model market, etc. It is clear the task is challenging with many drawbacks. Finding solution capable to work in real-time, in environment with uncontrolled conditions, for large number of classes and even with higher reliability is the main motivation sets for research presented in this paper. Deep understanding of the topic, methods and data is essential.

Facing presented motivation, we introduce an approach, which has a potential to be successful for all of mentioned drawbacks and restrictions. From analysis of current state of art, two approaches with knowledgeable potential were identified. Falzenschwalb's Histogram of Oriented gradients (FHOG) and Support Vector Machine (SVM) are examined. We suppose gradients and in block normalized histograms are very stable under different illumination changes, a partially invariant in scale and space plane. As for SVM, this method proposes effective solution of N dimensional space optimization problem, which can be greatly exploited in tasks with large number of classes. Combination of given methods is taken as developing design, elaborated to MMR and discussed in this paper. Such approach proposes two positives:

- Examining promising design which was not presented in current state of art.

- Disassemble the MMR of cars task to understand the problem effectively.

Since many works and attempts are available nowadays, description of current state of art is proposed in next section. In 3rd section, a brief explanation of used method is proposed. The 4th section presents insight into our MMR model design followed by followed by description of datasets. Experiments are analyzed in 6th section. The last section concludes this paper and shows on positives and negatives of presented design.

II. STATE OF ART

Considering on analysis of current state of art, we can group methods for MMR of cars into two groups, based on the feature extraction process:

1. Methods which use pixels as features.

2. Methods which formulate a model (transformation) for feature extraction from pixels.

Most works relay to the second group, because pixels in row form are unstable, noise sensible, etc. In next paragraphs we analyze several interesting works. Properties which are considered to be the most important are highlighted.

The first described work is from authors V. S. Petrovic a T. F. Cotes [1], they utilize a dataset of 1132 frontal car images, where 105 from them are used in training phase. In total they classify 77 different classes of cars. The database is recorded under semi controlled conditions. The classification process starts with localization of reference segment in image, which is the license plate. Therefore a special attention is laid on license plate detection. They propose own design, where the license plate is supposed to have a rectangular shape. Corners propose primary information for this process – all candidates of corners are extracted based on gradient filters and grouped by hierarchical algorithms into rectangular areas with required properties. Described features are extracted using different approaches – different gradient filters (Sobel edges, edge orientation, locally and squared normalized gradients, etc.). Authors use Principal Component Analysis (PCA) for feature vector dimensionality reduction. For feature classification, different simple distance measure methods were used: scalar multiplication and Euclidean distance. The final positive classification rate moved over 93%.

Authors in [2] use another large dataset of cars- 1414 images grouped into 20 classes. But more interesting is the fact, the data was collected under uncontrolled conditions in highway. In this work license place detection for localization of ROI is used. For feature extraction 2DLDA (2 dimensional Linear Discriminant Analysis) transformation coefficients are computed and they are classified using the k-NN method. The proposed approach provided results of positive car recognition up to 85%, with the capability to be executed in real time.

Different approach was used in [3]. Authors propose design which is able to cope with partially corrupted data (car images contain noise caused by barrier). The data can be considered as data collected in uncontrolled environment. Together used dataset contain 20 classes of frontal car images, split into training (173 images) and testing subsets (480 images). The approach can be described as multi level car type recognition system using oriented contour points set based model as features and multi classifier. For each vehicle class, from 1 to n prototypes are generated, and voting classification system is build. The final result of positive recognition moved up to 93%.

Next these authors enlarge the database using 50 classes comprised of 291 high quality frontal view images captured in car parks [4]. The testing dataset is comprised of 830 outdoor images with variance in lighting, angle, distance and resolution. The final classification rate remains similar to the previous one (93.1%).

From current state of art another interesting work was presented in [5]. From used dataset we highlight the idea: more common classes have more images and rare cars have fewer images applied in their experiments. Together 262 frontal cars images grouped into 74 different 'make and model' classes is processed. For feature extraction a new design based on Harris corner strengths is proposed– locally normalized Harris strengths (LNHS), where an image is recursively divided into quadrants in which all edges are hierarchically summed and normalized. In classification phase k-nearest neighbors and naïve Bayes are examined. The achieved results for positive classification moved around 96%.

From other works we choose [6], where the dataset contains 560 images divided into training and testing subsets with proportion 7:3. The significant difference is in ROI localization. Symmetry principle is used for car logo detection as alternative to methods which use license plate for this purpose. Features are extracted from car logo as a composite vector containing computed appropriate features from three approaches– SIFT method, gradient operators, wavelets transform. As final classifier SVM with linear kelner function is used. Such design proposes positive classification rate up to 90%.

Authors Remigiusz Baran a Andrzej Glowacy, Andrzej Matiolanski [7] propose another approach, where they use own database collected in controlled conditions (scale invariant images, illumination varies slightly). This dataset consist from 54 classes in 2499 images. The car frontal mask is segmented and used in ROI localization process. As feature extraction method the SURF descriptor (Speeded-up robust features) is chosen. Vectors from this method are ordered into centroids vocabulary called sparse vector of occurrence counts (SVoOC). They are trained and classified using SVM (Support Vector Machine) classifier. Results received from experiment in this work move up to 97% for their own database. Approach presented in this work was able to work in real time.

In [8] is presented another approach which is interesting especially for achieved classification results. Authors use own database, which consist of 400 images in 40 classes. The ROI is extracted based on license plate, using character detection algorithm connected with outer extraction of outer rectangle and symmetry. As feature descriptor HOG (Histogram of Oriented Gradients) is used. The k-nearest neighbor (k-NN) and the support vector machine (SVM) were tested as classifiers. The achieved results were more than satisfied receiving of 100% of positive recognition rate.

To conclude this overview we include a method which is based on probability theory [9]. Authors in this work executed experiments on own database, which consists of 38 different car types with 15 images for each type (the database is referred as COMVis_cardataset). The data is collected under uncontrolled conditions and it is split in proportion 1:1 for training and testing purposes. Described approach tries to handle with unwanted background, which seems to be unique approach from the current state of art. Author uses blob-like patching in scale pyramid for region detection, which is described by Local Energy Based Shape Histogram (LESH). They propose also approach where car types are recognized based on rear ward view. The received results moved around 62%.

TABLE I.   OVERVIEW OF METHODS FROM THE ANALYSIS OF CURRENT STATE OF ART FOR THE TASK CAR TYPE RECOGNITION. THE MOST INTERESTING INFORMATION IS THE TYPE OF OF USED FEATURES, THEN THE CLASSIFER AND INFORMATION ABOUT ACHIEVED RESULTS, TESTED DATASET AND POSITIVE CLASSICATION RATE.

Source	Feature extraction	Feature (pattern) classification	Used database	Achieved results	Inputs taken controlled | uncontrolled conditions	Capable working in real time?
[1]	Different approaches – mostly gradient based	Nearest Neighbor	77 classes, 1132 images	93%	Semi controlled	N/A
[2]	2DLDA	k-NN	20 classes 1414 images Frontal view	94.7%	Uncontrolled	Yes
[3], [4]	Gradient features (Sobel edges)	Multiclassifier – voting system	20, 50 classes, > 480 images, frontal view	~93 %	Uncontrolled	No
[5]	Canny edge detector / Squared mapped gradients / Harris edge strengths / LNHS	$k-$nearest neighbors (K-NN) / Bayes classifier	72 classes, 262 images frontal view	96% LNHS + Bayes / 94.9 – LNHS + K-NN	Controlled	N/A
[6]	SIFT, gradient, wavelet	SVM	560	90%	N/A	N/A
[7]	SURF	SVM	54 classes (1360 training images + 2499 testing images 2499) Frontal view	91,7% - 97%	Controlled	Yes
[8]	HOG	k-NN, SVM	40 classes 400 images Frontal view	70-100%	Semi controlled	Yes
[14]	LESH	Bayes	38 classes, 570 images for both frontal and rear view	48 %	Uncontrolled	N/A

From other approaches used in MMR, authors often refer to SIFT and SURF methods for feature extraction [10], [11], [12]. In classification phase authors use also tools from computational intelligence– neural networks [13] and genetic algorithms. A brief overview for further reading can be found in [5]. The following table (table 1) proposes a short summary of previously described approaches, highlighting set of properties, which are considered to be the most significant in the task of car type recognition model design.

To conclude this section we highlight some notes, which are considered in MMR of car system design:

- Most works use the frontal part (frontal view) of the car for recognition.

- License plate is very important informational area of the car. Many works extracted ROI (Region of Interest) based on the color and gradient characteristic of the license plate.

- Estimating the true validity of existing models and their use in real environment is more than complicated and the fact is supported by variety of used datasets, collected under various conditions, classifying from few or several tens of classes (car types).

III.   METHODS

In this section we introduce a basic theory, essential for further design. From analysis of current state of art it can be concluded; the task of MMR is in most works constructed from three parts. The first part is effective ROI localization based on license plate or car mask. The second part refers to feature extraction process and the third part is the classification phase. Considering proposed processing stage we will concentrate on the 2nd and 3rd stage:

- Histogram of Oriented Gradients (HOG) as feature extractor.

- Support vector Machine Method (SVM) as classifier.
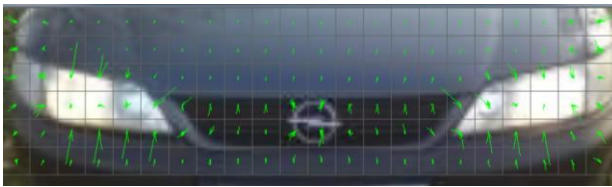
Figure 1.   Extracted features using FHOG method. Green lines represent gradients and their strengths for each cell in sample image
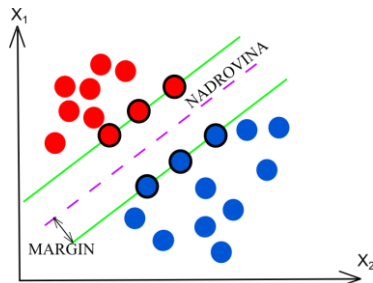
Figure 2.   linearly separable data, demonstration exsample in 2D space.
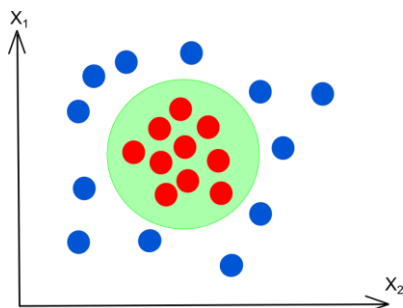
Figure 3.   non linearly separable data, demonstration exsample in 2D space.
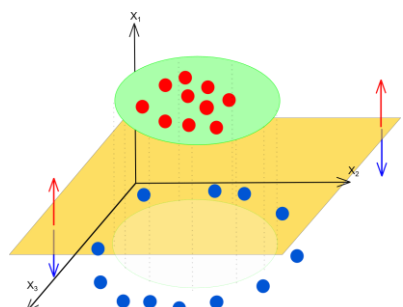
Figure 4.   Linearly non separable data in 2D Space, projected into 3D space - demonstration example

3.1 HOG and FHOG

The primary goal of this method is to characterize appearance and shape of objects, and to detect it in unknown image. HoG was presented in 2005 by Nawneet Dalal a Bill Triggs [15] for the first time.

This method divides the input image into small regions – cells (4x4, 8x8, 16x16). For each cell, histogram of oriented gradients is computed. By combination and normalization of these histograms into blocks of 2x2 cells a feature vector is extracted, which is stable against photometric and geometric

transforms. Authors in [16] propose advanced version of this method, which is designed to be more efficient in time – FHOG (F stands for Felzenswalb, who is one of authors). The main difference lays on the final size of extracted feature vector. In the basic version, the size of feature vector for single block consists from 36 values for contrast insensitive version, or 72 values for contrast sensitive version. But in FHOG the final feature vector is represented only by 31 values (18 contrast sensitive + 9 contrast insensitive + 4 normalization value within cells). A simple example of extracted FHOG feature is showed on the figure 1.

3.2 SVM

SVM was introduced in work of authors Cortes and Vapnik [17] and since then this method has became a very popular approach in tasks of numerical classification and regression.

SVM can be considered as machine learning approach, basically used for binary classification. For example, the simplest case of classifying features containing 2 classes in 2D space is using a line (see figure 2). But what is more common in practice, the classified features are not linearly separable (see figure 3). The SVM copes with this problem by using kernels, which can effectively projects data into more dimensional space (see figure 3) in which they are separable [18]. The criteria for selection of kernel functions are not defined; they should be selected experimentally with consideration of used dataset. The three most known kernel functions are:

- Polynomial

- Linear

- Radial Basis Function (RBF).

As it was already described, the SVM is binary classifier, and therefore the essential question to be answered is, how to solve multiclass problem? To be able classify many classes a multi classifier must be designed. Generally two approaches are often used, which are also analyzed in our experiments later:

- **One-against-one**. Multi-classifier consists from multiple binary classifiers, trained for each pair of considered classes. This approach is very exhaustive in training phase; because $\frac{k(k-1)}{2}$ binary models are created, where k is number of considered classes.

- **One-against-all**. This approach leads to training k models (binary classifiers), where each of them classify data from single class against data from other remaining classes.

For both presented models two different training strategies can be applied (see figure 5):

- **Single parametric strategy**. This approach trains single set of model parameters (in our case – parameters for SVM method) for the whole multi-classifier. A single instance of model parameters is searched for the whole problem. Then each binary classifier in multi-classifier has the same set (instance) of model's parameters trained for the solved problem.
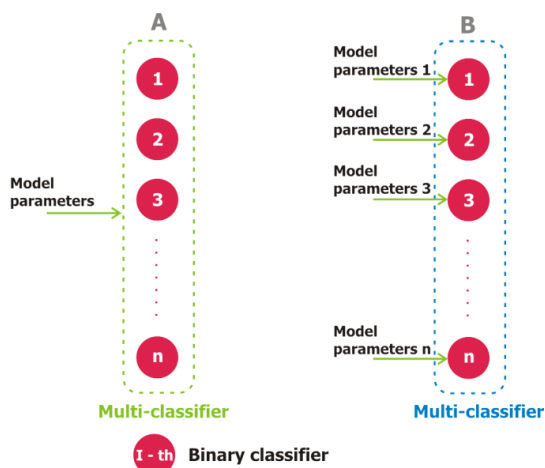
Figure 5. two approached in the process of training multi-class SVM model. Variant A - single parameters set for the whole model. Varian B multiple parameters set for model (each binary classifier has own set of trained parmaters).

- **Multi parametric strategy**. Opposite the previous approach we train parameters for each binary classifier, with the idea to propose optimized binary classifier in relation to the multi classifier.

Next theory refers to different approaches for results evaluation provided from multi-classifier. In experiments we refer to few of them, e.g. the most often used method, in one-against-one multiclassifier, is the strategy: Max Wins (the classes with most votes from all classifiers wins).

In this section a brief explanation to essential theory was proposed. To obtain more detailed insight into the methods, we refer to [15], [19] for HoG features, [17], [20], [21] and [22] for SVM theory and classification approaches comparison, next in [18], [23], [24], [25] the reader can look for SVM mathematical background, Lagrangians, numerical optimization, etc. In [26] information about multi-classifier approaches can be found for further reading.

IV. MODEL DESIGN

In this section a model design, used in experiments for solved MMR task is introduced. From analysis of current state of art, three main stages were identified, which we have adopted (see figure 6):

- Region of Interest (ROI) localization based on License plate.

- Robust and stable feature extraction.

- Classification of computed features.

As it was already mentioned in previous sections, the main goal of work presented in this paper is to analyze MMR problem deeply using FHOG and SVM methods, but proposed solution tends to be unique from several aspects. The following research questions are answered in designed experiments:

1) SVM as classifier is analyzed from several points of view for MMR: using different kernel functions, different multi-classifier approaches including not yet verified design one-against-one approach with multi-
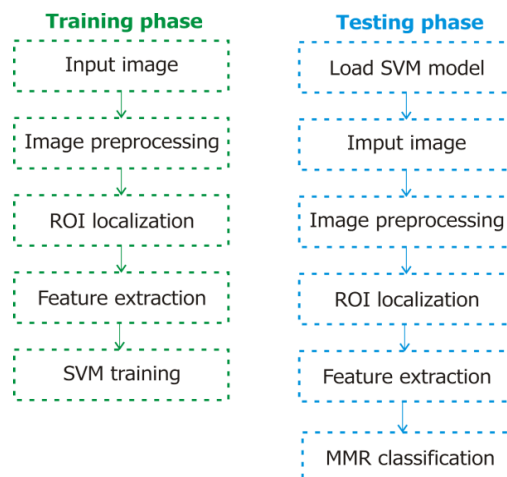
Figure 6. Workflow for the task of MMR. Training phase is essential for SVM model, where binary classifiers are trained. The testing phase serves for experiments verification. The testing and training phase were executed on different data.

parametric training phase. **We suppose multi-parametric approach can lead to more reliable results**, e.g. authors in [27] presented these different approaches to SVM in the task of handwritten digits classification.

2) **Which view of car is more informative**? Comparison two different approaches using different frontal and rear view of a car.

3) Since the computation in real time is more than essential aspect of image processing, presented experiments are looking for approaches, reducing the feature dimensionality. For feature extraction process the FHOG method is used, since HOG provides interesting results in works of other authors [8], replacing **HOG with FHOG should lead to more efficient computation**, while the recognition rates remains satisfactory.

4.1 Implementation

All methods and experiments were implemented in C++ using open libraries: OpenCV, Boost, dlib, LibSVM.

V. DATA DESCRIPTION

From works analyzed in section 2, many different datasets were identified. Results varied based on used datasets and methods. Most authors referred to own data collected for their purposes. The character of data varied also greatly. Some datasets were collected under controlled conditions, other under semi controlled (e.g. in indoor car) and few of them were recorded under uncontrolled conditions (for more information see table 1). In our experiments we decided to use two datasets. The detailed description of used databases can be found in table 2. Sample images from used datasets are showed in figure 7.

This section summarizes experiments executed on presented datasets in the task of MMR of cars. The main goal of presented results is to point at strengths and weaknesses of designed approach. Precise analyses and series of experiments

Figure 7. Sample images from used datasets. Upper row - db2. Bottom row - db1

TABLE II. DESCRIPTION OF DATASETS USED IN EXPERIMENTS

Database name	Num. of training images (classes)	Num. of testing images	Num. of Classes	Description
Db1_frontal [28]	304 (8)	266 (7)	38	Database of frontal car images, recorded under uncontrolled conditions
Db1_rear [28]	304 (8)	266 (7)	38	
Db2 [7]	1360 (80)	2499	17	Database of frontal car images, recorded under semi controlled conditions

provide significant insight into problematic as it was set in introduction.

All experiments were executed on personal computer of standard equipment. The next text presents a set of executed experiments with achieved results. Research questions introduced in section 3 will be analyzed and answered.

First of all we summarize; that all models were trained and then tested. In both case datasets presented in previous section were used. All results are measured on testing database. The training phase is not included in described results. But it is important to note the MMR task was executed on images extracted ROI. The ROI extraction was executed with high efficiency deploying the methods based on license plate localization.

Plots in figures 8-10 summarizes achieved results of positive recognition rates for described model designs. They differ each other in use of kernel function for SVM classifier – **linear, polynomial, RBF**. Each plot describes the percentage of positive recognition for 3 datasets. Each dataset was examined using three model configurations.

The 1st assumption:
a) The most relevant results, from all presented configurations, were achieved using one-against-all in combination with linear kernel function. The

dominancy of this model can be seen for all tested datasets.

b) The average accuracy of model with configuration one-against-all was higher than other examined cases.

c) Notable results can be observed when a comparison between single-parametric and multi-parametric configuration is performed. The logical assumption about better efficiency of multi-parametric compared to single-parametric system was not confirmed.

The 2nd assumption:
a) From presented experiments can be summarized: the rear view of the car is more informative. This fact is understandable, because the rear view contains more information about car brand and generally the manufacturers make more face-lifted backs of car than front mask.

The 3rd assumption:
b) Using FHOG seems to be an acceptable alternative to already used feature extraction approaches. The computational time during testing phase was deeply beneath 1 second, which is positive, since the classification was more than competitive to other approaches (see table 3).
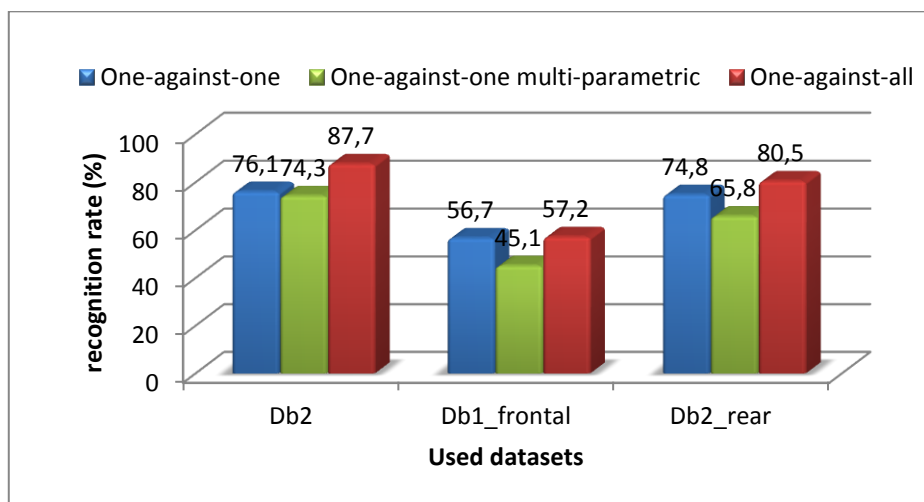
Figure 8.   Car model recognition using SVM (linear kernel) + FHOG with different SVM configuration for 3 databases.
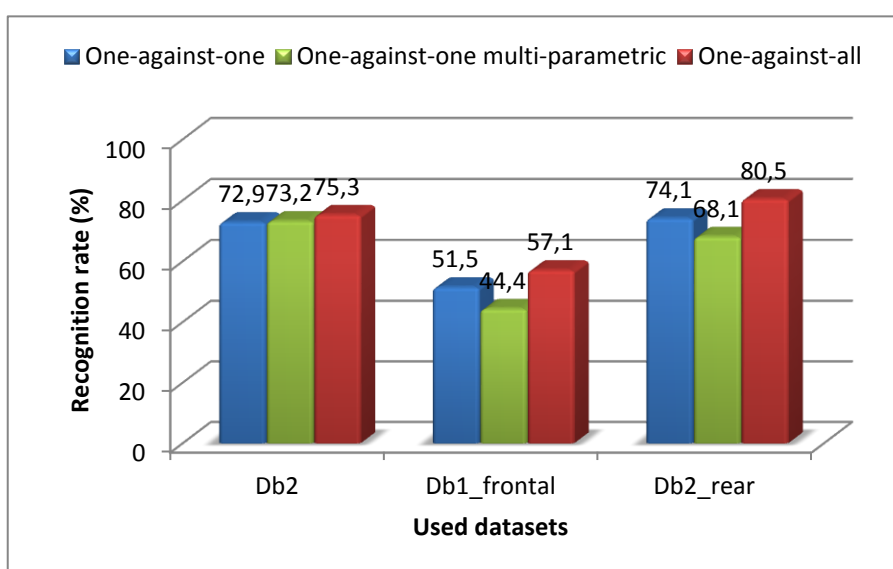
Figure 9.   Car model recognition using SVM (polynomial kernel) + FHoG with different SVM configuration for 3 databases.
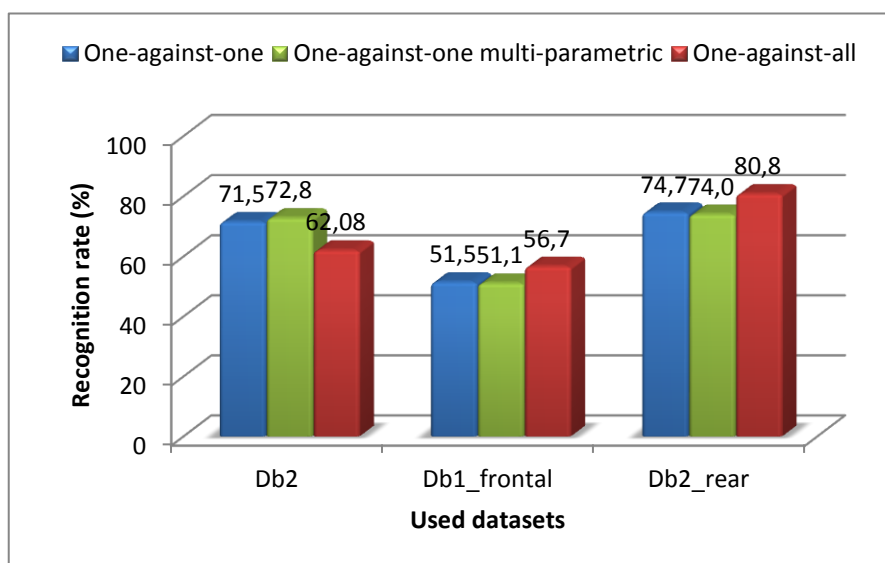
Figure 10. Car model recognition using SVM (RBF kernel) + FHOG with different SVM configuration for 3 databases.

TABLE III. COMPARISON OF RESULTS ACHIEVED IN PRESENTED EXPERIMENTS TO OTHER APPROACHES.

Dataset	Feature extraction	Classifier	Positive classification
Db2	SURF	SVM	91,7 %
Db2	FHOG	SVM	87.6 %
Db1_frontal	LESH	Integrálny rozdiel	48 %
Db1_frontal	FHOG	SVM	57,6 %
Db1_rear	LESH	Integrálny rozdiel	62 %
Db1_rear	FHOG	SVM	80.8 %

TABLE IV. EXAMPLE OF CONFUSION MATRIX, WHICH POINT OUT CLASSIFICATION RESULTS BETWEEN CONSIDERED CLASSES.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17
T1	75	3	0	0	4	0	1	0	0	0	2	3	3	0	0	1	0
T2	4	135	0	0	11	0	0	0	0	0	2	0	2	2	0	0	0
T3	2	3	193	0	3	1	1	6	0	0	1	4	0	0	9	14	0
T4	0	0	0	147	0	0	1	3	0	0	1	0	1	0	0	0	14
T5	0	0	0	2	116	2	0	4	0	0	1	2	0	0	0	0	0
T6	0	0	0	2	2	159	0	4	0	0	0	0	0	0	0	0	0
T7	0	0	0	0	0	1	73	1	1	0	0	1	0	0	0	0	2
T8	0	0	1	0	0	0	0	63	1	0	0	0	0	0	0	0	0
T9	0	1	0	0	0	1	1	23	41	0	0	0	0	0	0	0	1
T10	0	6	0	1	3	0	3	1	0	128	0	0	0	0	0	0	0
T11	1	0	0	0	0	0	0	0	0	1	147	0	0	0	0	0	0
T12	1	1	2	0	1	0	2	9	3	0	0	168	0	0	0	2	2
T13	2	2	0	0	0	0	1	0	0	1	0	1	151	3	0	0	0
T14	2	1	1	0	0	0	4	0	0	0	5	1	8	155	12	0	1
T15	0	2	0	0	19	0	0	0	0	0	1	8	0	0	182	1	2
T16	1	0	6	0	0	0	3	5	0	0	2	4	0	0	7	115	0
T17	0	0	0	2	0	0	1	3	0	0	1	0	0	0	0	0	142

To propose an even more efficient computation (time execution), a dimensionality reduction methods can be exploited. Since such methods require additional models to be trained and they are often not trivial, such possibility needs to be analyzed. Next figure proposes a brief summary of experiments results with Principal Component Analysis (PCA) as dimensionality reduction method. The relation between classification efficiency and PCA space dimensionality is explored.

VI. CONCLUSION

A deep analysis was presented for the task of MMR of cars in this paper. The MMR presents a very challenging task, which needs to be understood from several points of view.

This paper discussed many aspects of MMR problem. The recognition model was presented, which uses FHOG method for feature extraction and SVM classifier in the classification phase. Deep analysis of proposed solution, supported with
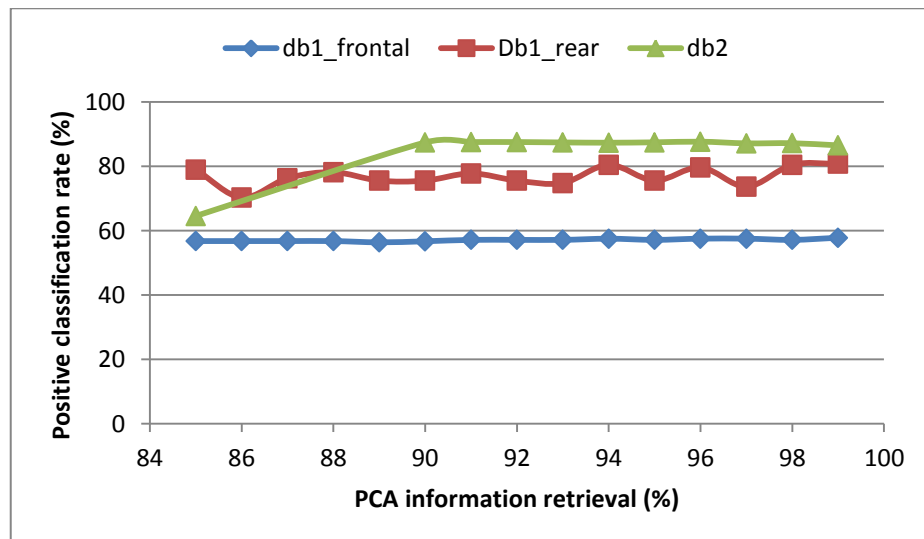
Figure 11. This figure displays positive classification rate dependence from PCA information retrieval level. Considering only 90% of information means using 45% of values from original feature vector, but the classification rate remains unchanged.

series of experiments were discussed and described logically. Achieved results are competitive in comparison to other approaches analyzed in current state of art.

Future experiments with proposed design should be aimed at e.g.: designing a tree structure of simple classifiers. Such approach can eliminate effectively classes which are difficult to classify. The table 4 shows an example of confusion matrix for one of presented experiments. It can be noticed; some classes bring a noise into classification. Such cases can be considered separately, to enhance the final classification rate.

Splitting the MMR problem into two stages e.g. manufacturer identification and even then car brand of manufacturer recognition can be other design of tree boosted recognizer. Other experiments can relay to different feature extraction processes although the most works refer to edges, which seem to be informative. The rear view of car holds a lot of information; this fact can be utilized to design approach considering this view.

REFERENCES

[1] V. S. Petrovic a T. F. Cotes, „Analysis of Features for Rigid Structure Vehicle," rev. British Machine Vision Conference, 2004.

[2] H. Huang, Q. Zhao, Y. Jia a S. Tang, „A 2DLDA Based Algorithm for Real Time Vehicle Type Recognition," rev. 11th International IEEE Conference on Intelligent Transportation Systems, 2008. ITSC 2008., Peking, 2008.

[3] P. Negri, X. Clady, M. Milgram, R. Poulenard a L. Prevost, „Multiclass Vehicle Type Recognition System using an Oriented Points Set based Model," 2008.

[4] X. Clady, P. Negri, M. Milgram a R. Poulenard, „Multiclass vehicle type recognition syste," rev. Artificial Neural Networks in Pattern Recognition, 2008.

[5] G. Pearce a N. Pearce, „Automatic Make and Model Recognition from Frontal Images of Cars," rev. IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2001.

[6] M. M. Nabeel a M. F. rEl-Dein, „Intelligent Vehicle Recognition based on Wireless Sensor," International Journal of Computer Science, zv. 20, %1. vyd.4, pp. 164-174, 2013.

[7] R. Baran, A. Glowacz a A. Matiolanski, „The efficient real- and non-real-time make and model," Multimedia Tools and Applications, 2013.

[8] S. Lee, J. Gwak a M. Jeon, „Vehicle Model Recognition in Video," International Journal of Signal Processing, Image Processing and Pattern Recognition, zv. 6, %1. vyd.2, 2013.

[9] I. Zafar, E. A. Edrisinghe a B. S. Acar, Localised Contourlet Features in Vehicle Make and Model Recognition, Loughborough University, 2012.

[10] L. Dlagnekov, „Video-based car surveillance: license plate, make, and model recognition," 2005. [Online]. [Cit. May 2014].

[11] S. Badura a S. Foltan, „Advanced scale-space, invariant, low detailed feature recognition from images – car brand recognition," rev.

IMCSIT'10 – International Multiconference on Computer Science and Information Technology, 2010.

[12] S. Foltan a S. Badura, „Robust car brand recognition from camera image," rev. Mendel 2010, 16th International Conference on Soft Computing, 23-25, 2010.

[13] H. J. Lee, „Neural network approach to identify model of vehicles," rev. Proc. of the 3rd int., 2006.

[14] M. S. Sarfraz a M. H. Khan, „A PROBABILISTIC FRAMEWORK FOR PATCH BASED VEHICLE TYPE RECOGNITION".

[15] D. Navneet a T. Bill, „Histograms of oriented gradients for human detection," Computer Vision and Pattern Recognition, zv. 1, pp. 886-893, 2005.

[16] P. Felzenswalb, R. Girshick, D. McAllester a D. Ramanan, „Object Detection with Discriminatively Trained Part Based Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, zv. 32, %1. vyd.9, 2010.

[17] C. Cortes a V. Vapnik, „Suppport Vector Networks," Machine Learning, zv. 20, %1. vyd.3, pp. 273-297, 1995.

[18] T. Wen a A. Edelman, „Support Vector Machine Lagrange Multipliers and Simplex Volume Decompositions," 2000. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.32.4011&rep =rep1&type=pdf. [Cit. 28 3 2014].

[19] P. Felzenswalb, R. Girshick a D. McAllester, „Cascade Object Detection with Deformable Part Models," rev. IEEE Conference on Computer Vision and Pattern Recognition, 2010.

[20] C. Lehmann, T. Koenig, V. Jelic, L. Prichep, R. John, L. Wahlund, Y. Dodge a T. Dierks, „Application and comparison of classification algorithms for recognition o f Alzheimer's disease in electrical brain activity," Journal of Neuroscience Methods, zv. 161, p. 342–350, 2007.

[21] A. Y. Ng a M. I. Jordan, On Disciminative vs Generative classifiers: A comparision logistic regresion and naive Bayes, Berkeley: University of California, Berkeley.

[22] X. Wu, V. Kumar, R. J. Quinlan, J. Ghosh, Y. Qiang, H. Motoda a G. J. McLachlar, „Top 10 algorithms in data mining," Springer, 2008.

[23] C. J. Burges, „A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, zv. 2, %1. vyd.2, pp. 121-167, 1998.

[24] J. Platt, „Fast Training of Support Vector Machines using Sequential Minimal Optimization. In Advances in Kernel Methods: Support Vector Learning," rev. Advances in kernel methods, MIT Press Cambridge, 1998, pp. 185 - 208 .

[25] P. H. Chen, C. J. Lin a B. Schölkopf, „A tutorial on ν-support vector machines," Applied Stochastic Models in Business and Industry - Statistical Learning, zv. 21, %1. vyd.2, pp. 111-136, 2005.

[26] C.-W. Hsu a C.-J. Lin, „A Comparison of Methods for Multiclass Support Vector Machines," IEEE TRANSACTIONS ON NEURAL NETWORKS, zv. 13, %1. vyd.2, pp. 415-425, 2002.

[27] M. Johnatan, C. Mohamed a S. Robert, "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?.

[28] M. S. Sarfraz a O. Hellwich, „Probabilistic Learning for Fully Automatic Face Recognition across Pose," Image and Vision Computing, zv. 28, %1. vyd.5, pp. 744-753, 2008.

New supervised learning approaches for sex identification in archaeology

Vlad-Sebastian Ionescu
Department of Computer Science
Babeş-Bolyai University
Cluj-Napoca, Romania
ivlad@cs.ubbcluj.ro

Abstract—**Most methods used in archaeology for the task of sex identification of human skeletal remains involve bone measurements, DNA analysis and various statistical approaches. We consider machine learning algorithms to have the potential to outperform these classical methods, since the numerical features that can be extracted from bone measurements can be used to train machine learning models to better recognize the necessary patterns for sex identification. A machine learning model based on genetic algorithms is introduced for this task. Our experimental results prove this approach to be superior to the existing results on the data sets that have been used in this research, as shown by comparing this paper's results with previous ones.**

Keywords—**machine learning, genetic algorithms, archaeology, sex classification, prediction**

I. Introduction

Detecting the sex of the skeletal remains found in archaeological sites is of major importance in modern bioarchaeology, since is it useful for obtaining information regarding the social patterns within the society, inter-human relations and family customs etc. Sex identification is an essential for task bioarchaeologists and not a simple one, since the factors for sex selection are not completely known.

We are approaching in this paper the problem of sex identification of human skeletal remains from bone measurements from a machine learning perspective. Many studies in the literature propose various approaches for detecting the sex of human skeletons, most of them are based on bone measurements, DNA and gene analysis or different statistical methods. Except for our previous approach from [1], there are no approaches in the literature that use machine learning in order to identify the sex of human remains. In [1], we have proposed two models: artificial neural networks and self-organizing maps, which proved to be effective for the sex detection task.

A supervised learning approach based on *genetic algorithms* (GA) is introduced in this paper for solving the sex detection problem. The experimental evaluation of the proposed method

is performed on four publicly available data sets. The case studies used in the experiments for evaluating the performance of our models show that our approach outperforms the existing similar approaches from the literature.

To the best of our knowledge, there are no existing approaches that use genetic algorithms to perform supervised learning with the purpose of discovering the sexual category of a skeletal remain.

The rest of the paper is organized as follows. Section II presents the problem approached in this paper, highlighting the motivation of our work, as well as different approaches existing in the literature for sex classification of human skeletal remains. The fundamentals of the *genetic algorithms* used in this article are given in Section III. Section IV introduces the new approach for the determination of sex in human skeletons. Section V presents the settings used for the GA model and the case studies that have been used for testing the new model on. Experimental results of the GA based learning methods on the case studies described in the previous section are given in Section VI, while analysis of the results and comparisons with related work from the literature are presented in Section VII. Section VIII contains some conclusions and indicates future research directions.

II. Sex Identification of Human Skeletal Remains

In this section, the problem of identifying the sex of human skeletal remains will be presented. The section starts by emphasizing the motivation for this research and continues with a survey of the literature that deals with the considered problem.

A. Motivation

According to [2], being able to conduct accurate sex identification from archaeological human remains is essential for studying the sex differences in past civilizations. This ability would lead to a better understanding of the social positions and roles each sex had in that society.

Although the skull and the pelvis are the body parts that most often present sexual dimorphisms, which means sex is usually determined by analysing the skull or the pelvis, other bones can be used as well. Sex identification using only bone

measurements is not completely accurate, especially if dealing with individuals that have not fully matured.

A lot of research has been conducted on important bones in the human body that can be used for sexual identification. For example, forearm bones have been used in [3] and [4] [5] to determine the sex of South African modern human remains and to perform forensic analysis, respectively. In [6], leg bones, such as the femur, have been used to compute relevant metrics with respect to sex identification. It can therefore be concluded that arm and leg bones have the potential to be just as relevant to sex determination as the skull and the pelvis.

We consider it important to note that sexual classification is a delicate task that is highly influenced by the historical period and the geographical origin of the considered skeletal remains. Therefore, a certain data set can only be used successfully if our goal is to accurately classify remains from time periods and geographical areas that are close to our benchmark. This is important due to the fact that populations from different time periods and locations tend to have significantly different stature and bone lengths. Current research in this field also seeks to uncover any intrinsic relations between the features taken into account that might help us to better distinguish between the studied populations, thus enabling researchers to better classify skeletal remains from similar times and places.

Due to the fact that archaeological data is often imprecise and that any eventual hidden relations between features are complex, any kind of processing on such data requires the ability to deal with lack of information, complexity and imprecision. Therefore, computational intelligence techniques [7] such as machine learning [8] algorithms can prove to be effective tools for archaeological data processing.

B. Related work

A study [3] conducted on a considerable number of South African skeletal remains has led to the development of a discriminant function formula which employs measurements of the forearm bones, particularly the radius and ulna. The employment of the function on the studied osteological benchmark provided a $76\% - 86\%$ accuracy rate.

Another proposed discriminant function used in the literature was introduced in [9], using only measurements on the radius obtained from a slightly lighter benchmark of only 63 individual measurements, both male and female. The proposed function correctly classified 35 of the women, representing 97.2% and 25 of the men (92.6%).

In the case of adult skeletal remains, it was statistically proven that in the cases in which the classification is questionable due to several secondary sex indicators in the bone, such as skull size, femoral head size and bone rugosity, the individual under study is more likely to be classified as male rather than female in [10] 12% of cases.

The most popular indicators for the sex, from an osteological point of view, are the cranium and the ox coxae, but in many circumstances these are not enough for a conclusive result and other bones are considered, or even other skeletal

characteristics such as dental material or DNA samples extracted from bone.

A visual based protocol for analysing the ox coxae was introduced by Bruzek [11] and it achieved a 95% classification rate on a benchmark of 402 French and Portuguese adult individuals.

Measurements of the metatarsals, metacarpals and phalanges were used as sexual discriminators in [12] and [13], giving an accuracy of 80% and $84 - 92\%$ respectively. Foot measurements were used in [14] for both sex and stature detection, obtaining accuracies worth mentioning: 95.6% for the right foot and 96.4% for the left foot.

The sex classification task's difficulty drastically increases when incomplete or immature skeletal remains are involved, as this task is close to impossible to achieve with traditional morphometric analysis. In order to overcome this impasse, in addition to the already employed morphological (non-metric) and metrical approaches, molecular studies were introduced to solve the sex classification problem. Each of these methods has its limitations [15], but when employed together they may complement each other to give better classification rates. A very promising approach in this respect uses the amplification of the single-copy amelogenin-encoding gene for sex detection. It was tested on a small number of skeletal remains, 22, correctly classifying 18 out of a benchmark containing osteological material from young children born from 200 to 8000 years ago.

Another technique that uses Polymerase Chain Reaction on preserved DNA was proposed in [16]. This method is highly volatile as even charred or burned remains give good results in the analysis.

Another method based on molecular genetics proposed by Vanharová and Drozdová in [17] can successfully be applied on children or juveniles remains as it is not restricted by physical fragmentation and it provides the clarification needed when traditional classification methods do not bring conclusive results.

The sex classification problem is obviously very interesting from a computational intelligence perspective as several techniques have been applied during the past years on both archaeological and forensic materials.

Neural networks proposed in [18] have been applied to classify archaeological remains from two archaeological databases with a success rate of 81% to 88%.

Backpropagation neural networks were used in [19] to classify individuals with respect to their sex based on measurements of the pelvic and patella bones. Accuracies of $86.6\% - 98.3\%$ were obtained on a data set generated based on statistical values reported in the literature [20].

Apart from the neural network based approaches, variations of decision trees were employed to solve the classification problem. The CHAID (Chi-square automatic interaction detection) model is a type of decision tree technique using the statistical Chi-square test [21] to determine the next best split attribute at each node in the tree. This method was tested on a benchmark of 304 skeletal remains of American, European

and African ancestry dating from 1915 to 1955 and it obtained detection rates between 85% and 85.5%.

III. Background on genetic algorithms

Invented and developed in the 1960s and 1970s by Holland and his students, GAs represent a generic class of algorithms used for solving search and optimization problems. They are inspired by the biological processes of natural selection and evolution, in which the fittest individuals of a population thrive, and the population as a whole adapts to changes in the environment over successive generations [22].

The general idea of a genetic algorithm is to have a list (called the *population*) of chromosomes (called the *individuals*) that encode information about a candidate solution to a problem needing to be solved (for example, in the present case, chromosomes would encode information about a formula that identifies the sex based on a set of given features). Then, the algorithm involves advancing (*evolving*) this population to a new *generation* by applying genetic operators inspired from biology, such as *crossover* and *mutation*. In each generation, the individuals are ranked according to a *fitness function*, such that the chromosomes that encode information that solves the problem better are ranked higher (in our case, those that encode a formula that leads to better accuracy). The algorithm stops once an acceptably fit individual has been evolved [22].

Determining what genetic operators to apply and in what proportion is a difficult problem that has no proven general solution, requiring researchers to deal with it on a per-problem basis.

Due to the applicability of GAs in search and optimization problems, it is reasonable to also study their usefulness in supervised learning scenarios.

IV. A supervised learning model based on genetic algorithms for sex idetification

In this section, the novel approaches for detecting the sex of human skeletal remains using *genetic algorithms* are introduced.

In order to more formally present the new model, let us introduce some notations. For this purpose, let our data set be denoted by $S = \{s_1, s_2, ..., s_n\}$, where each s_i denotes a skeletal remain labeled as female (0) or male (1). Each s_i is called a *training instance*. Each training instance is defined by a set of m bone measurements performed on it (these are the features). Therefore, one can write $s_i = (s_{i1}, s_{i2}, ..., s_{im})$.

The sex classification process takes place in three phases:
1) Data pre-processing.
2) Training.
3) Testing.

As in any eager classification process, the classification *model* (based on Genetic Algorithms) will be built during the training phase, and during testing, the model built during training will be applied for classifying an unseen human skeletal remain.

Figure 1 shows the workflow used for determining the best hyperparameters (settings) for our GA-based model.

The sex classification steps will now be further detailed.

A. Data pre-processing

The first step in building our models is the *data pre-processing* step. This step involves scaling the input data to $[0, 1]$ using the $min - max$ normalization method and then applying a statistical analysis in order to determine those measurements (features) that are most likely to help a machine learning algorithm to correctly identify the sex. For this purpose, the *information gain* (IG) measure [23] is used. Information gain is a statistical property, usually used in building decision trees, that measures how well a certain feature separates the training instances according to their class (male or female in our case). A higher IG value suggests better separation potential for that feature. Note that, for computing the IG for the features, the feature domain must first be discretized by dividing it in ten intervals of equal-size.

B. Training the genetic algorithm based model

The genetic algorithm-based model introduced in this paper is based on the stochastic acceptance [24] model, which has been shown to lead to implementations of genetic algorithms that perform very well in practice on a variety of problems.

The training part of this classifier consists of evolving a real-valued vector w such that, for as many training data instances as possible, the dot product between w and the vector of measurements for our considered features returns a value closer to 0 for female training instances and closer to 1 for male training instances. Note that this evolved formula does not necessarily return a result in $[0, 1]$. Formula (1) shows how a given instance is classified for a certain vector w.

$$IdentifiedSex(s_i) = \begin{cases} male & \text{if } \sum_{k=1}^{m} w_k \cdot s_{ik} \geq t \\ female & \text{otherwise} \end{cases} \quad (1)$$

Thus, this algorithm evolves an acceptable vector under a given accuracy p. A given accuracy p causes the algorithm to stop once the number of incorrectly classified training instances is at most p. For example, for $p = 50$, the algorithm stops once it has evolved a real-valued vector that incorrectly classifies at most 50 training instances.

The parameter t is a threshold value that determines when an instance is classified as male and when as female. Usually, this is avoided by using a bias term in the summation, but experiments have found that specifying a threshold value works better on the test cases at hand.

The fitness function of the algorithm is directly tied to the accuracy, being equal to the number of instances the considered individual incorrectly classifies. Therefore, the lower the fitness value of an individual, the fitter that individual is.

Given this implementation, the stochastic acceptance GA model for this problem involves the following steps: sort the individuals in increasing order of fitness and then copy each individual c to the next generation with probability $\frac{fitness(M)}{fitness(c)}$, where M is the fittest individual. Under this model, the fittest individual will always survive to the next
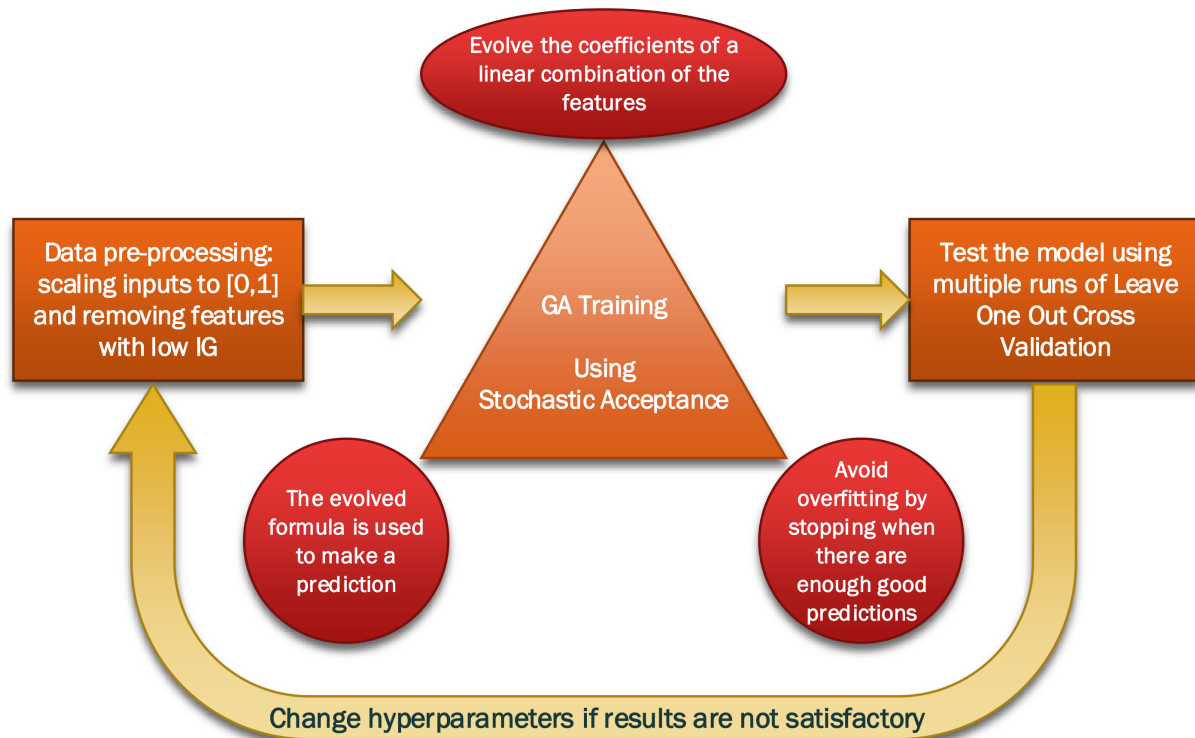
Fig. 1. Flowchart describing the process used to build the best performing model.

generation. After this, parents are randomly selected and accepted with the same probability measure to produce offspring using single point crossover [22]. The final step consists of applying mutation to the new generation.

Once the algorithm has converged to the desired accuracy, the elements of the fittest individual will represent the values for the vector w that will be used for sex identification.

After the algorithm has finished executing on the training set, it produces a real valued vector that correctly classifies a certain number of the training instances (this number depends on the accuracy p that is chosen before the execution). This vector can then be used on the test instances in order to classify these as well, just like in Formula (1).

C. Testing

Testing the model can be achieved by applying a cross validation method. For evaluating the performance of the GA model, five runs of *leave one out (LOO) cross validation* have been used. Cross validation is a technique used for estimating the generalization error of a classifier [25]. In LOO cross-validation, the training data consist of all but one training instances. After the model is trained, it is tested on the instance that was left out. This is then repeated until every training instance has been left out exactly once.

After determining the information gain for the features, as indicated in the *data pre-processing* step (see Section IV-A), features with low IG values are removed. In order to determine what features to eliminate, all features are first sorted in descending order of their information gain. Then, features are

eliminated one by one in increasing order of IG values. The previously described cross-validation is performed at most five times after each new feature elimination. LOO cross validation was run fewer times when convergence was too slow, but the results on the executed runs suggested that the complete five runs would not improve the accuracy. In this way, the best list of features to remove for all the considered case studies was selected, such that the resulting classifier accuracy is maximized.

For evaluating the performance of the proposed classifier, two evaluation measures are used:

1) The *accuracy* (denoted by Acc), which measures the percentage of instances that are correctly classified.
2) The *Area under the ROC curve* measure (denoted by AUC), indicated in the literature [26] [27] as one of the best evaluation measures to compare classifiers. The ROC (Receiver Operating Characteristics) curve is a two-dimensional plot of $1 - specificity$ vs $sensitivity$. In case of classifiers which return the class, the ROC curve has a single point that is linked to the points at $(0,0)$ and $(1,1)$, thus producing a curve, for which the area under the curve (AUC measure) can be computed [28] [27].

Good classifiers have high Acc and AUC, thus in order to obtain better classifiers these measures have to be maximized.

The values of the evaluation measures presented above are averaged over the five runs of LOO cross validation. The obtained values provide a good estimate of the expected

performance of the model.

The GA algorithm will be executed more than one time, because some of the variables are given random values at certain points in the execution. The selection, generation of new individuals and mutation are done with certain probabilities, which give a random character to the execution. This is why the mean and the standard deviation of the measurements of all the executions, for each type of experiment, are also studied.

V. Experimental evaluation

This section describes the experimental evaluation of the GA model (described in Section IV) on four case studies which were built starting from three data sets obtained from the literature [3] [29]. These data sets were previously used for sex identification from the length of the bones.

It is worth mentioning that the GA approach used in this paper was implemented from scratch, without the help of high level third party libraries.

A. GA parameters

The GA model was applied on the case studies mentioned above. The GA settings are as follows:

- *Population size:* 80.
- *Mutation rate:* 30%.
- *Mutation method:* for mutation, a random number from $\left[-\frac{1}{8}, \frac{1}{8}\right]$ was added to a random element of the individual selected for mutation.
- *Population reset:* if the desired accuracy has not been achieved after 300 generations, the GA population is reset. The new population is generated by creating individuals whose elements consist of random numbers from $\left[-\frac{1}{4}, \frac{1}{4}\right]$, just like the initial population.
- *Accuracy:* an accuracy of $p = 45$ was used for the first three cases studies and $p = 55$ for the fourth one. This was chosen in order to keep the execution time to acceptable levels while still obtaining good results: for the data sets of 400 instances (the first three case studies), this forces the GA formula to obtain at least 88.72% accuracy on training data. For the fourth case study of 310 instances, this forces the GA to produce a formula that correctly classifies at least 82.26% of the instances.
- *Threshold*: a threshold value of $t = 0.5$ was used for all case studies.

B. Case studies

The first data set experimented on was extracted from the paper [3] and consists of 200 male and 200 female skeletons from the Pretoria Bone and Raymond A. Dart collections. Ten measurements are considered in [3] for the radius bone and nine measurements are considered for the ulna bone. The human remains from these data sets represent black South Africans from the 19th and 20th centuries, having the birth date between 1863 and 1996.

1) First case study: The first case study consists of human remains identified by ten radial measurements extracted from [3]. There are 10 features describing the instances within this data set. The features represent radial measurements [3]: maximum length of the radius, distal breadth, circumference at the midshaft, sagittal diameter at midshaft (minimum diameter), transverse diameter at midshaft (maximum diameter), vertical radial head height, minimum head diameter, maximum head diameter, circumference of the radial and circumference at the tuberosity [3].

2) Second case study: As the second case study, nine measurements of the ulna bone [3] are considered for the human skeletons. There are 9 features describing the instances within this data set: maximum length of the ulna, maximum length of the ulna measured using the plumbline geniometer method, anterior-posterior diameter (minimum diameter), medial-lateral diameter (maximum diameter), circumference at midshaft, minimum circumference of the ulna, olecranon breadth, minimum olecranon breadth and height of the olecranon [3].

3) Third case study: The third case study consists of the data set which contains both radial and ulnar measurements extracted from [3]. Consequently, in this data set, each skeleton will be characterized by 19 features (measurements): the first ten are the radial measurements (as considered in the first case study from Section V-B1) and the next nine features represent the ulnar measurements (as considered in the first case study from Section V-B2).

4) Fourth case study: The fourth case study considered in the experiments is constructed on a data set which is part of the database available at [29]. This database consists of a set of postcranial osteometric measurements extracted from the Robert J. Terry Anatomical Skeletal Collection. The human skeletons from this data set represent white and black individuals (most of them are americans) from the 19th and 20th centuries, being born between 1822 to 1943.

The data from [29] contains missing values for different measurements, thus, a subset consisting of 310 individuals has been extracted from this database, for which the measurements of *humerus*, *radius*, *ulna*, *femur*, *tibia*, *fibula* bones and the *race* are available. For each bone, the maximum length between the left and right bone is considered. The data set consists of 67 White Males, 107 Black Males, 49 White Females and 87 Black Females.

VI. Results

Results are presented on the four case studies showcased above. The results section concludes with Table I, which summarizes the results of the newly introduced GA-based models and how they compare with each other.

A. First case study

As mentioned in Section V-B, there are 10 initial features used for the the sex classification task for the first case study. After the data scaling, the information gain measure is
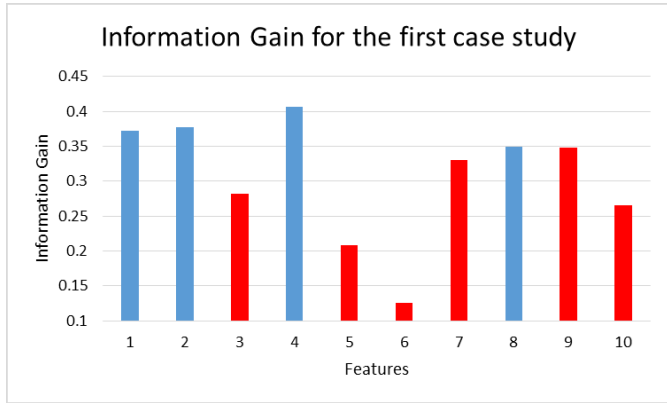
Fig. 2. Information gain for the features from the first case study.
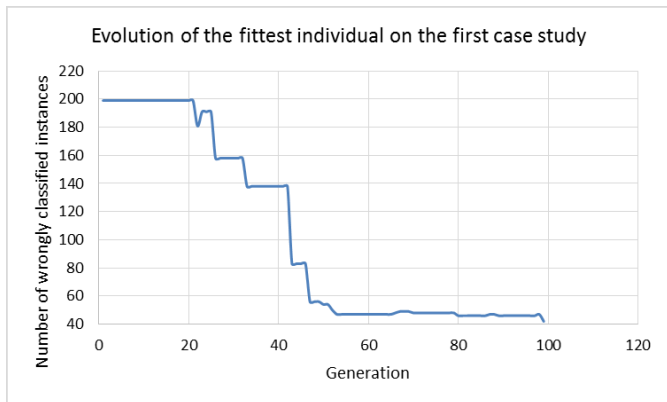
Fig. 3. Evolution of the fittest GA individual for the best training run on the first case study.

Fig. 4. Information gain for the features from the second case study.

$$\begin{pmatrix} 0.41659012285638436 \\ 0.042116974896644094 \\ 0.8921287034020694 \\ 0.17854161990772244 \end{pmatrix} \quad (2)$$

B. Second case study

Considering the second case study, there are 9 initial features in the sex classification task. After the data was scaled, the information gain measure was determined for each feature. The obtained values are depicted in Figure 4.

After applying the same feature elimination method as in the first case study, only the top five features were kept, based on the information gain property. The removed features show up as red in Figure 4.

For the second case study, the GA provided an average accuracy of 0.8775 and an average AUC value of 0.88. The standard deviation of the accuracy values over the five LOO runs was 0.005, with the best of the five runs having an accuracy of 0.88. The standard deviation for the AUC values was also 0.005, with the best value over the five runs being 0.88.

Figure 5 presents the evolution of the fittest individual in one of the GA runs that obtained the best accuracy on the training data in all the runs over the five LOO cross validations. Note that more than one run obtains this value.

This time, convergence to the desired accuracy is about twice as fast.

Formula (3) presents the vector evolved by the fittest individual mentioned above. Its fitness is 43.

$$\begin{pmatrix} 0.28595130905103916 \\ 0.14895761898315588 \\ 0.29456796875214997 \\ 0.15833247386593727 \\ 0.15608989574136944 \end{pmatrix} \quad (3)$$

C. Third case study

For the third case study considered for evaluation, there are 19 features characterizing the skeletal remains (the first ten features are those from the first case study and the next nine

determined for each feature. The obtained values are depicted in Figure 2.

After applying the method described in the methodology (Section IV), a decision has been made to only keep the top four features based on the information gain property. The features removed show up as red in Figure 2.

For the first case study, the GA provided an average accuracy of 0.8715 and an average AUC value of 0.87. The standard deviation of the accuracy values over the five LOO runs was 0.009, with the best of the five runs having an accuracy of 0.89. The standard deviation for the AUC values was also 0.009, with the best value over the five runs being 0.89.

The minimum accuracy value obtained on the five validation runs was 0.86.

Figure 3 presents the evolution of the fittest individual in one of the GA runs that obtained the best accuracy on the training data in all the runs over the five LOO cross validations.

Note that, while convergence to the exact desired accuracy becomes slow quite soon in the execution, a population reset is not necessary for the algorithm to converge.

Formula (2) presents the vector evolved by the fittest individual mentioned above. Its fitness is 42 on the training data.

Fig. 5. Evolution of the fittest individual for the best training run on the second case study.
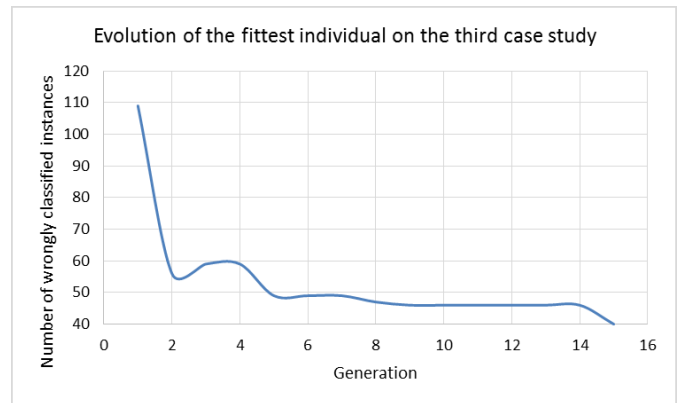
Fig. 7. Evolution of the fittest individual for the best training run on the third case study.
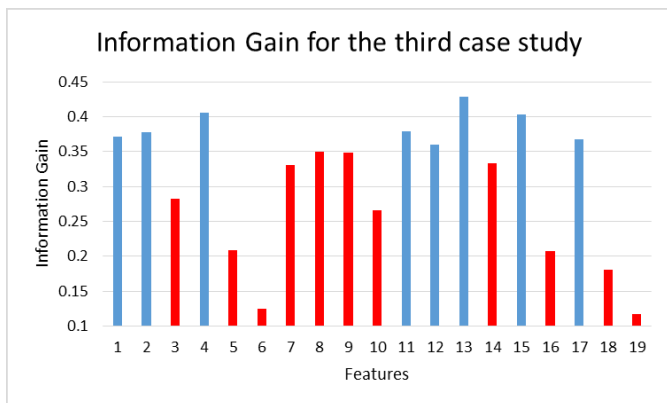
Fig. 6. Information gain for the features from the third case study.
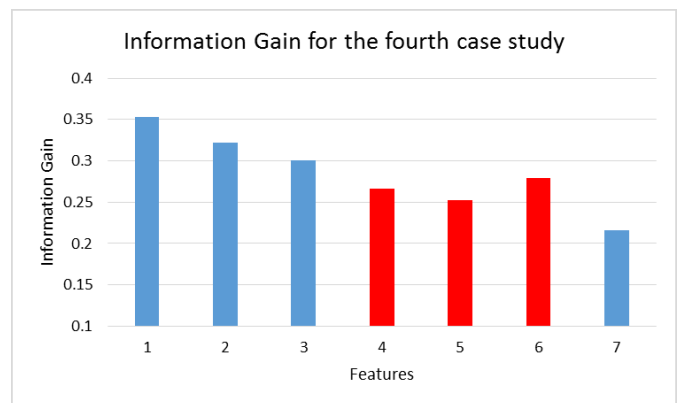
Fig. 8. Information gain for the features from the fourth case study.

features are the ones from the second case study). Like in the first two studies, the feature measurements were scaled to $[0, 1]$ and then the information gain was computed for each feature. This can be seen in Figure 6.

For the third case study, the features computed for removal in both of the previous studies were eliminated first, then, starting from there, more features were eliminated by using the procedure previously described. This way, feature number 8 was eliminated as well, which proved to improve results.

For the third case study, the GA provided an average accuracy of 0.8785 and an average AUC value of 0.88. The standard deviation of the accuracy values over the five LOO runs was 0.008, with the best of the five runs having an accuracy of 0.89. The standard deviation for the AUC values was also 0.008, with the best value over the five runs being 0.89.

Figure 7 presents the evolution of the fittest individual in one of the GA runs that obtained the best accuracy on the training data in all the runs over the five LOO cross validations. This time, this value only occurs once.

Convergence to the desired accuracy is now about four times quicker than in the second case study.

Formula (4) presents the vector evolved by the fittest individual mentioned above. Its fitness is 40, much better than

our minimum required accuracy.

$$\begin{pmatrix} 0.045651048000390965 & 0.15632152486849818 \\ 0.15870263521015154 & 0.19151044356211844 \\ 0.13739229766243344 & 0.15627497164546766 \\ 0.21928416732996503 & 0.09207649043474486 \end{pmatrix} \quad (4)$$

D. Fourth case study

For the fourth case study considered for evaluation, there are 7 initial features characterizing the skeletal remains, six bone measurements (described above) and another feature corresponding to the race. First, the data was scaled to $[0, 1]$ and then the information gain for the features was computed. The information gain values are depicted in Figure 8.

For this case study, the same methodology as before was attempted, but better results were obtained by making an exception for the *race* feature and leaving it in.

For the fourth case study, the GA provided an average accuracy of 0.80 and an average AUC value of 0.80. The standard deviation over the five LOO runs was 0.009, with the best of the five runs obtaining an accuracy of 0.82. The standard deviation for the AUC values was also 0.009.

Figure 9 presents the evolution of the fittest individual in one of the GA runs that obtained the best accuracy on the training
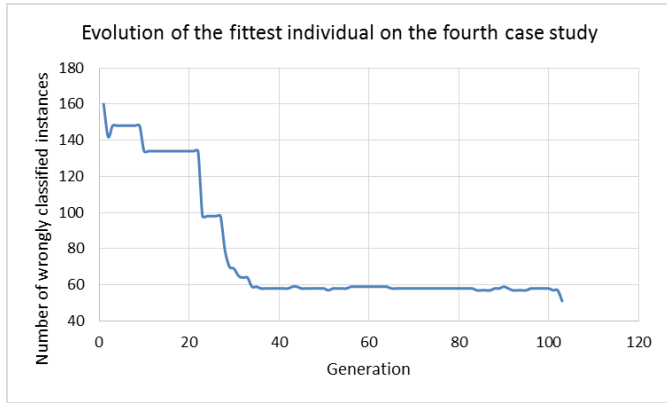
Fig. 9. Evolution of the fittest individual for the best training run on the fourth case study.

Case study	Average accuracy	Average AUC
First	0.87	0.87
Second	0.88	0.88
Third	0.88	0.88
Fourth	0.80	0.80

TABLE I

AVERAGE VALUES FOR ACCURACY AND AUC OBTAINED USING THE GA MODEL.

data in all the runs over the five LOO cross validations. This value occurs more than once for this case study. It can be seen that convergence happens fairly soon, after evolving about 100 generations.

Convergence to the desired accuracy is the slowest in this case. In general however (not only for the fittest individual), convergence is as fast for the fourth case study as for the rest.

Formula (5) presents the vector evolved by the fittest individual mentioned above. Its fitness is 51. Recount that 55 is the maximum allowed number of misclassified instances on this case study.

$$\begin{pmatrix} -0.05939702232996749 \\ 0.7699973541544664 \\ 0.36378287262844655 \\ 0.2286613198843308 \end{pmatrix} \quad (5)$$

E. Summary

Table I summarizes the results obtained for the GA-based model on the four case studies. It can be seen that the best results are obtained on the second and third case studies, and the worst on the fourth case study.

VII. COMPARISON TO RELATED WORK

As shown in Section II-B, most of the approaches existing in the literature for determining the sex of skeletal remains are based on bone measurements, DNA and gene analysis or different statistical methods. As far as we know, there are no approaches that use genetic algorithms for learning to detect the sex of human skeletons frome bone measurements.

The first three data sets used in the experimental part of this paper were also used by the discriminant analysis method

Case study	Model	Mean	Min	Max	Stdev
First	GA	**0.8715**	**0.8625**	**0.8775**	**0.009**
First	DF [3]	0.83	0.81	0.86	0.026
Second	GA	**0.8775**	**0.875**	**0.88**	**0.005**
Second	DF [3]	0.84	0.79	0.87	0.036
Third	GA	0.8785	0.8725	**0.89**	0.008
Third	DF [3]	-	-	-	-

TABLE II

COMPARATIVE RESULTS BETWEEN THE CURRENT MODEL AND EXISTING LITERATURE RESULTS ON THE FIRST THREE DATA SETS.

which was proposed in [3]. In [3], five discriminant functions (DF) were used for the data set considered as the first case study and four functions were proposed for the data set from our second case study. Since only the accuracy is reported in [3] for estimating the performance of the sex detection task, this measure is the one used for comparing the performance of the models.

Table II summarizes the values for the accuracy measures (minimum, maximum, mean and standard deviation of the values) reported by the new GA-based model, as well as for the discriminant analysis method reported in the literature [3] for the problem of sex identification in human skeletal remains. The best obtained values are highlighted. In [3], experiments are performed only on the first and second case studies and the evaluation is made using the $leave-one-out$ method for cross-validation, as in the approaches used for the GA model.

From Table II it can be observed that, considering the average accuracy, the GA model outperforms, both on the first and second case studies, the discriminant analysis method from [3]. Moreover, for both case studies, the standard deviation reported for GA model is less than the standard deviation of accuracies for the method from the literature. This is another argument for the usefulness of this approach.

The absolute best accuracy was obtained on the third case study, which includes measurements for both the radius and ulna bones. These results are also better than what exists in the literature.

The introduced approach has several advantages over methods that require human examination of the skeletal bones and extensive domain knowledge. First of all, observer subjectivity is eliminated, since our method is fully automated. Second, machine learning algorithms are a lot more flexible than simple formulas or algorithms, because they are able to adapt to any kind of related problem (such as sex prediction on different data sets of the same nature) and are able to discover more complex patterns in the data than a human expert is able to. Finally, the new method is easily applicable to different regions of the globe, at most requiring the tuning of a small set of parameters for best results.

A more in-depth comparison of the new approach to other similar existing ones cannot be made, since in the existing approaches the experiments are performed on data sets which differ from the ones considered in this paper and are not publicly available. That is why, a comparison that is based only on the obtained accuracies across distinct data sets is not relevant, since the data sets used in the experiments are not

the same.

VIII. CONCLUSIONS AND FURTHER WORK

In this paper a new supervised learning based method for detecting the sex of human skeletons from bone measurements has been proposed, which is based on genetic algorithms. The experimental evaluation is performed on three open source human skeletal remains data sets and four case studies were conducted in order to validate our methods. The obtained results indicate a good performance of the proposal with respect to existing similar approaches from the literature.

Further work will be done in order to extend the experimental evaluation of the proposed learning models on real data sets [30] to better assess their accuracy and to analyse different mechanisms for feature selection [31] and dimensionality reduction [32]. We will further investigate the effectiveness of instance based learning methods [23] for the problem of sex identification of skeletal remains. Applying relational association rules mining [33] [34] for the considered problem is also a direction with great potential, since relations between different bone measurements in skeletal remains are likely to be relevant, from a bioarchaeological perspective, for the sex identification problem.

REFERENCES

[1] G. Czibula, I. Mircea, D. Miholca, I. Czibula, and R. Moldovan, "Novel data mining approaches for gender detection of human skeletal remains," *Forensic Science International*, p. under review, 2015.

[2] M. Faerman, D. Filon, G. Kahila, C. L. Greenblatt, P. Smith, and A. Oppenheim, "Sex identification of archaeological human remains based on amplification of the x and y amelogenin alleles." *Gene*, vol. 167, no. 1-2, pp. 327–32, 1995.

[3] I. L. O. Barrier, "Sex determination from the bones of the forearm in a modern South African sample," 2007, phD thesis, Univeristy of Pretoria.

[4] O. Celbis and H. Agritmis, "Estimation of stature and determination of sex from radial and ulnar bone lengths in a turkish corpse sample," *Forensic Science International*, vol. 158, no. 2, pp. 135–139, 2014.

[5] E. F. Kranioti and M. Michalodimitrakis, "Sexual dimorphism of the humerus in contemporary cretans–a population-specific study and a review of the literature," *Journal of Forensic Sciences*, vol. 54, no. 5, pp. 996 – 1000, 2009.

[6] K. S. Elizabeth, "A metric method of sex determination of femur and its fragments," *International Journal of Contemporary Medicine*, vol. 1, no. 2, pp. 107–111, 2013.

[7] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd ed. Wiley Publishing, 2007.

[8] N. Nilsson, *Introduction to Machine Learning*. MIT Press, 1998.

[9] Y. Wada, "Discriminant function for sex determination of ancient Iraqis based on radial measurements," *Anthropological Science*, vol. 102, no. 2, pp. 149–158, 1994.

[10] K. Weiss, "On the systematic bias in skeletal sexing," *American Journal Of Physical Anthropology*, vol. 37, no. 2, pp. 239–49, Sep 1972.

[11] J. Bruzek, "A method for visual determination of sex, using the human hip bone," *American Journal Of Physical Anthropology*, vol. 117, no. 2, pp. 157–68, Feb 2002.

[12] D. Troy Case and A. H. Ross, "Sex determination from hand and foot bone lengths," *J Forensic Sci.*, vol. 52, no. 2, pp. 264–270, 2007.

[13] A. Falsetti, "Sex assessment from metacarpals of the human hand," *J Forensic Sci.*, vol. 40, no. 5, pp. 774 –776, 1995.

[14] G. Zeybek, I. Ergur, and Z. Demiroglu, "Stature and gender estimation using foot measurements," *Forensic Sci Int.*, vol. 181, no. 1-3, pp. 54.e1–5, 2008.

[15] M. Bidmos, V. Gibbon, and G. Strkalj, "Recent advances in sex identification of human skeletal remains in South Africa," *South African Journal Of Science*, vol. 106, pp. 1–6, 2010.

[16] J. Kiesslich, F. Neuhuber, H. J. Meyer, M. P. Baur, and J. Leskovarm, "Dna analysis on biological remains from archaeological findings - sex identification and kinship analysis on skeletons from mitterkirchen, upper austria," pp. 147–154.

[17] M. Vanharová and E. Drozdová, "Sex determination of 4000 years old children and juveniles skeletal remains from Eneolithic burial site Hostice 1 za Hanou (Czech Republic) by ancient DNA analysis." *Anthropological Review*, vol. 71, pp. 63–70, 2008.

[18] S. Bell and R. Jantz, "Neural network classification of skeletal remains," in *Archaeological Informatics: Pushing The Envelope. CAA2001*, G. Burenhult and J. Arvidsson, Eds. Archaeopress, Oxford, pp. 205–212.

[19] I. Afrianty, D. Nasien, M. Kadir, and H. Haron, "Determination of gender from pelvic bones and patella in forensic anthropology: A comparison of classification techniques," in *2013 1st International Conference on Artificial Intelligence, Modelling and Simulation (AIMS)*, Dec 2013, pp. 3–7.

[20] M. Akhlaghi, A. Sheikhazadi, A. Naghsh, and G. Dorvashi, "Identification of sex in iranian population using patella dimensions," *Journal of Forensic and Legal Medicine*, vol. 17, no. 3, pp. 150 – 155, 2010.

[21] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*, ser. Wiley Series in Probabilities and Statistics. Wiley, 1996.

[22] M. Melanie, *An Introduction to Genetic Algorithms*. The MIT Press, 1999.

[23] T. M. Mitchell, *Machine learning*. McGraw-Hill, Inc. New York, USA, 1997.

[24] A. Lipowski and D. Lipowska, "Roulette-wheel selection via stochastic acceptance," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 6, pp. 2193 – 2196, 2012.

[25] G. Wahba, Y. Lin, and H. Zhang, "GACV for support vector machines, or, another way to look at margin-like quantities," *Advances in large margin classifiers*, pp. 297–309, 2000.

[26] N. Lavrac, B. Kavsek, P. A. Flach, and L. Todorovski, "Subgroup discovery with cn2-sd." *Journal of Machine Learning Research*, vol. 5, pp. 153–188, 2004.

[27] T. Fawcett, "An introduction to ROC analysis," *Pattern Recogn. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.

[28] V. López, A. Fernndez, S. Garca, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics," *Information Sciences*, vol. 250, no. 0, pp. 113 – 141, 2013.

[29] R. Terry, "Terry collection postcranial osteometric database," 2015, http://anthropology.si.edu/cm/terry.htm.

[30] "Institute of Interdisciplinary Research in Bio-Nano-Sciences," 2015, http://bionanosci.institute.ubbcluj.ro/.

[31] J. Xiao, Y. Xiao, A. Huang, D. Liu, and S. Wang, "Feature-selection-based dynamic transfer ensemble model for customer churn prediction," *Knowledge and Information Systems*, vol. 43, no. 1, pp. 29–51, 2015.

[32] X. Zhang, Y. He, L. Jiao, R. Liu, J. Feng, and S. Zhou, "Scaling cut criterion-based discriminant analysis for supervised dimension reduction," *Knowledge and Information Systems*, vol. 43, no. 3, pp. 633–655, 2015.

[33] G. Serban, A. Câmpan, and I. G. Czibula, "A programming interface for finding relational association rules," *International Journal of Computers, Communications & Control*, vol. I, no. S., pp. 439–444, June 2006.

[34] J. Sahoo, A. Das, and A. Goswami, "An effective association rule mining scheme using a new generic basis," *Knowledge and Information Systems*, vol. 43, no. 1, pp. 127–156, 2015.

Web 2.0 services for informing elderly people: Web for Health

Dragan Peraković*, Marko Periša*, Ivan Cvitić*

* University of Zagreb, Faculty of Transport and Traffic Sciences,
Department of Information and Communication Traffic,
Zagreb, Republic of Croatia, EU
dragan.perakovic@fpz.hr, marko.perisa@fpz.hr, ivan.cvitic@fpz.hr

Rosana Elizabeta Sente**, Maja Radošević**, Petra Zorić**

* * Graduate Student at University of Zagreb, Faculty of Transport and Traffic Sciences
Zagreb, Republic of Croatia, EU
sente.rosana@gmail.com, maja.radosevich@gmail.com,

Abstract – **This research on user requirements defines functionalities of the Web 2.0 systems for the targeted user group (elderly, people with impaired vision, hearing, speech, and neurodegenerative diseases - TUG) in order to improve the quality of their lives. For this purpose, there is a proposal of Web 2.0 technologies (HTML5 and PHP) whose functionalities allow dynamism and adaptability of information to the TUG. Architecture of solution is based on Cloud Computing (CC) technology which provides 24/7 support for provision of service and safety. Functionalities defined in this paper, are in fact a form of an assistive technology, which allows a registration of the TUG and enables a necessary adjustment of websites according to the degree of deficiency. Due to the functionalities listed within this paper, developed application solution allows the TUG to spend their third age in a dignified manner, providing them with the possibility to have more interaction with people and the opportunity to participate in various activities in their fields of interest.**

Key words: HTML5, Cloud Computing, assistive technology, eHealth

I. INTRODUCTION

Human life time has lengthened in the last twenty years, which indicates a large population of elderly people. It is a known fact that older people have difficulties in adjusting to the hectic world around them and mostly to the use of the functionalities based on newly developed technologies.

The research [1] is aimed at older people who have mental health problems and want to stay at home. A significant number of people is in need of greater care and technology is one of the things that can help them self-manage their health. This requires a large number of mutually synchronized devices. The application solution would decrease the number of needed devices, because everything would be available in one place. Based on [2], it can be concluded that a small number of older people use Internet and that is because they do not have it implemented at home or because they find it unnecessary. Content that is available via Internet is considered unattractive because it's based on providing services of interests to young people. Therefore, it is necessary to understand the needs and interests of the elderly in order to create a compelling content. Because some of the users do not have a constant access to the Internet, a development of application based on HTML5 programming language would give them the privilege to use it without the need of constant Internet connection. The article [3] has examined how older people respond to the introduction of new technologies. They think that the usage of new technology will not be much of a help for them in their everyday life, primarily because the learning of how to use new technologies requires a lot of time. Therefore, they think they need help in acquiring the new knowledge, and the application solution must provide them with sufficient information of its appliance.

Due to previous facts, several researches were carried out in order to gather information about the needs of older people. It was concluded that there is still no satisfactory hardware and software solutions to match the high-quality forms of the assistive technologies. Therefore, the proposed application solution contains functionalities that can provide greater autonomy and awareness of the TUG. The implementation of the web application and its development allows TUG to customize the user interface according to their actual needs with an aim to provide accurate and real-time information services. Considering that the technology is constantly evolving, the learning and the appliance of new technologies is often considered repulsive so it is necessary to create the proper environment and a simple visual presentation to enhance TUG's desire for education. The proposal of the CC technology as the operational basis of the *Web for Health* solution is conducted on a previous research and based on the accuracy of the use of real-time information [4]. Wireless technologies of collecting and providing information on the status of user and his further movement depend on specification and technical capabilities of the service which can provide the user with desired useful information [5].

The aim of this research is to create a Web 2.0 service which will be based on new web technologies and suitable for targeted deployment. Based on the previous research and studies, it is possible to make a selection of services that are

proved as useful to the TUG and also represent a basis for functionalities of the application solution.

II. RESEARCH METHODOLOGY

To assure that functionalities of the service are fully customized to the TUG, important information were gathered by conducting a survey. The survey included questionnaires and interviews of elderly in 10 nursing homes in the city of Zagreb. From the total number of 208 respondents, 56% were older than eighty years. Other data about the age of respondents is shown in Figure 1.
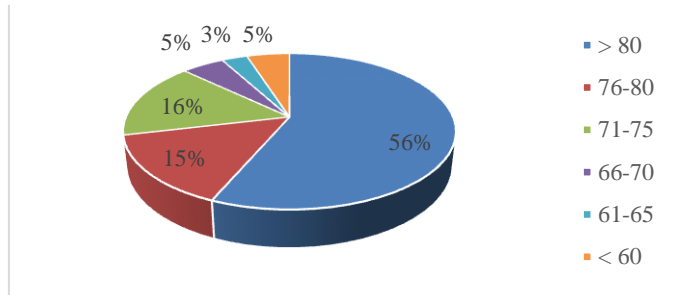
Figure 1.   Structure of the test sample (i.e. users) according to age

Figure 2 shows the reasons for the use of services based on new technologies. Out of total, 42% of respondents listed that they find the application solution most helpful in terms of health problems, 23% of respondents find it most helpful for the orientation and mobility, and 16% for the communications and entertainment (19% of the respondents have listed other reasons).
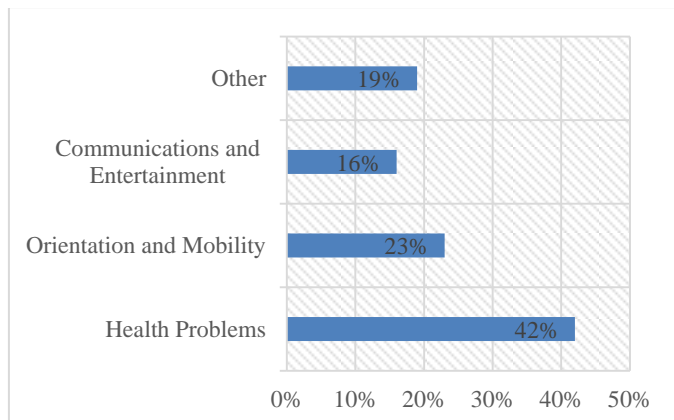
Figure 2.   The purpose of using Web for Health services (i.e. interests of the TUG)

Considering the fact that a large percentage of respondents defines a health care as their most important need, it was necessary to conduct an additional research on their health problems. Visual impairment is present within the 49% of the respondents, and the hearing loss within the 38% as well as the presence of neurodegenerative damage. These results provide the information on how to adjust the registration for people with cognitive impairment as well as how to adjust the application solution. The TUG rarely use mobile devices or computers to solve such problems, but they're willing to learn

how to use the application solution that could be customized to their needs with quality instructions.
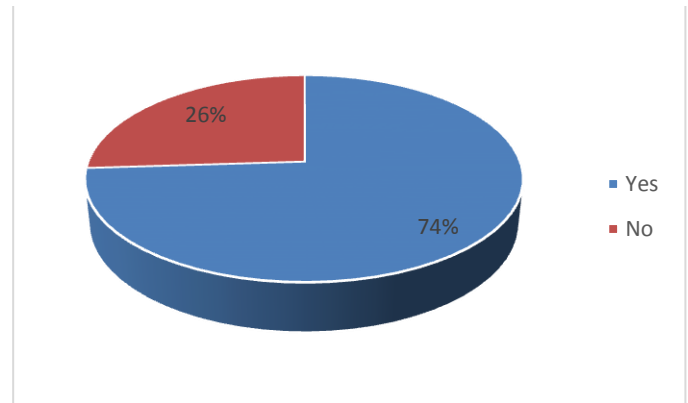
Figure 3.   The percentage on the use of mobile terminal devices within the TUG

A percentage of targeted users who use a mobile terminal device is surprisingly high and it even ranges to 74% (Figure 3). But, out of the total number of respondents only 22% use a personal computer. While conducting the research on the purposes of using a mobile terminal device or a personal computer it is concluded that the 23% of respondents use their devices to access the Internet and the 21% use their devices for the purpose of health care. The results of this conducted research have greatly contributed to the adjustment of proposed application solution. There was also a research about technologies that make the best match for the form of application solution, conducted by a simple communication with the TUG in order to find out what kind of content and layout of the application is considered suitable. Through the interview of the individuals who are in a nursing home, it was concluded that the best solution would be if there is a person who will show them how to use the application or a tutorial that will provide a voice guidance throughout the application. The concept of this service does not require any demanding action and provides a tutorial that guides the user through every step of it so it could truly be easy to use. If the user forgets how to use a service he can restart tutorial at any time.

III. PROPOSED TECHNOLOGIES FOR *WEB FOR HEALTH SOLUTION*

The accessibility of application solution is carried out and assured by the current information and communication technologies, the optimization for search engines, the offer of alternative solutions and tools and the modern design that was developed in HTML5 environment. Application solution must be developed according to the guidelines for the development of accessible web applications published in 2008 under the name of WCAG 2.0. The appearance of application solution is adjusted to TUG, based on instructions for the operability of the web applications. A nine criteria is taken into account [6]. The standardization is made for the purpose of abbreviated keyboard using six recommended shortcuts as well as the definition of the direction of navigation that goes from left to right. Informative architecture is composed of two most important types of information; the main and the useful. Flashing content and pop-ups are not readable for the visually

impaired, therefore, are not used within the application design. Each page content that is in text, visual or audio form contains its short description in tags. A need for a visual adaptability is also taken into consideration, therefore the users are able to adjust the size of the letters, select the color and choose a background color. The information about which color and size TUG finds most appropriate for the use was also collected, and it was taken into account when developing application solution.

Registration system makes one of the most important components of the application solution. It includes user defined

users, the registration system is inevitable process of defining the purpose of the usage and the functionalities of application solution.

The registration process is designed in the form of a small windows where the main window contains a series of steps for the user. The user is not allowed to pass onto the next step until he finishes the present step. If the user cannot see well, he opens a window of defined issues related to his damage, and after he answers all the questions, he passes to the following possible deficiency. On this basis, the application adapts itself
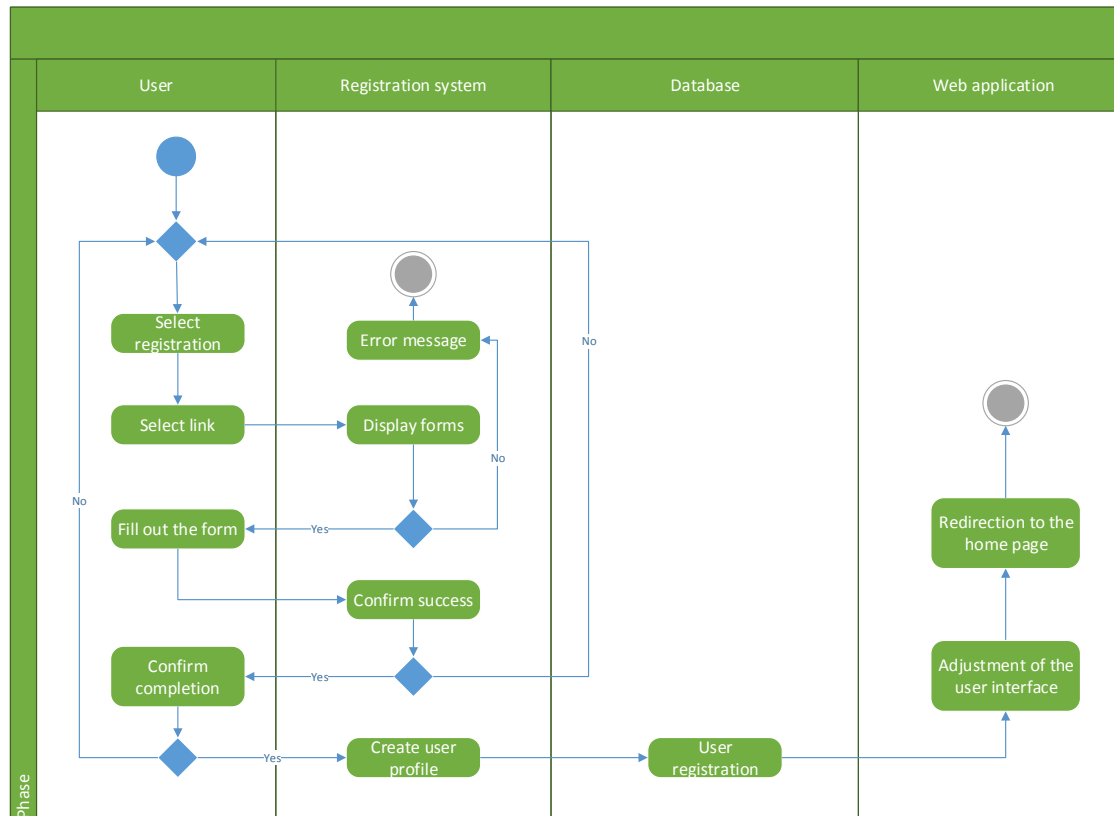
Figure 4. The registration process

desires and needs, as much as it is possible. Registration is required as it contains questions whose answers directly affect the layout of the application solution. The Figure 4 shows the process of registration. The authentication system is activated by accessing the page and it opens a dialog box for entering a name and a password. The registration process actualizes the integration of applications and databases with the authentication system.

Users who actively interact with the web page are divided into three groups and each group has a different color of username for the purpose of easy recognition when using the application solution. The first group consists of administrators who have access to all elements of the page. The second group consists of end users, or people who are not a part of development team but use applicative solution to inform and communicate. The last group consists of volunteers whose page content is limited in order to prevent any confusion of the end-users. Because of the present differences between the

to the user. The issues related to problems with concentration, dyslexia, epilepsy etc. are also included. The steps of application customization are written within the PHP script, and they can, for example, define the background color and the content that will not initiate epileptic seizures. When the assessment of the user's deficiency is made, he fills in a questionnaire about his personal data. The required data such as the user's name must be entered, but it cannot be seen by other users. The reason for this principle is to increase the privacy of the users and thus reduce the occurrence of false identity and provide comfortable use of the application. To avoid confusion, the entire registration is defined by simple questions that motivate users to learn and to use modern software solutions. After completing the registration process, applicative solution adapts to the user's needs. If the user is not satisfied with the choice of colors or just want a change, he is always able to choose the background color as well as the font size and color.

The background is made of selected MySQL backend database for web content organization. The application is based on the CC solution and it defines its architecture. The reason for choosing PHP technology is that it provides access to a variety of platforms, and the development time in the PHP environment is much shorter compared to ASP.NET. Short codes within the application program provide faster operability of application solution and directly affect on the speed and the

is important to note that users are not provided with the infrastructure or the development platform but with finalized application solution. The service provider is in charge of the improvements, maintenance and development of application solution distributed in the *Software as a Service* (SaaS) environment [10], [11].
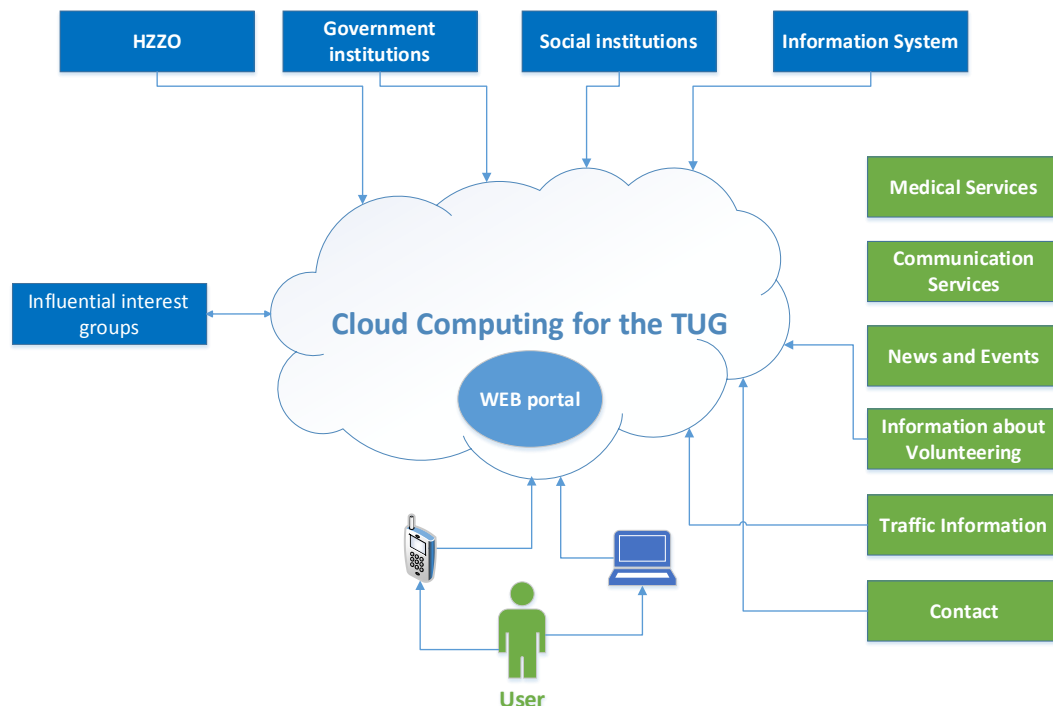
Figure 5. Architecture of the Web4Health application solution based on CC technology

availability of information to the user. PHP has better connectivity and communication with MySQL - TV & MSSQL and is compatible with a variety of databases which enables easier appliance of assistive technologies [7], [8].

HTML5 development environment was chosen because it has great advantages compared to a *native app* [9]. Developing and adapting in a *native app* requires a separate work on every operating system (OS), while in HTML5 it's not necessary due to the fact that the same applies to all versions of OS's. Applied technology is distributed through the conceptual architecture of the application solution which is based on the CC technology as shown in Figure 5. Elaborated recommendation of the CC environment from the user's point of view includes services of informing and social networking. Since the customer wants to increase the quality of life, the application solution also includes external links to social and government institutions that have been added to provide users with additional information [4].

The reason for the use of CC environment are many advantages from the user's point of view, such as very simple management, automatic update, facilitated cooperation and communication among users, data compatibility, worldwide availability, and data security through its backup and update. It

IV.   FUNCTIONALITIES OF THE APPLICATIVE SOLUTION BASED ON THE RESULTS

According to the results, it is noticed that the majority of users have problems with visual or hearing disabilities. Because of these difficulties, application solution includes functionality that will adapt the text depending on deficiency. Textual alternatives are assured for all elements that are not textual so that they can be converted into a needed form, such as the texts for the visually impaired, braille, speech, symbols or simplified language. Each user has an optional choice of visual, audible, tactile or combining form to read the content. Full developed application solution includes the user registration system that allows customization of interfaces depending on the user's disabilities and needs.

Figure 6 shows the segmentation of the Use-Case diagram of functionalities within the application. Certain functionalities present the fields of interest for certain stakeholders. Communication services enable virtual communication through a public chat groups and a private chat. That functionality is the field of Interest for the TUG. Informative functionalities are medical information services, the Events and news, the Traffic information services and the external links and they are a field of interest for influential interest group, as well as for the TUG.

Application functionality of Volunteers represent a list of volunteers and services they are willing to provide to the TUG. The TUG can communicate through the Contact of the service provider, i.e. via e-mail.
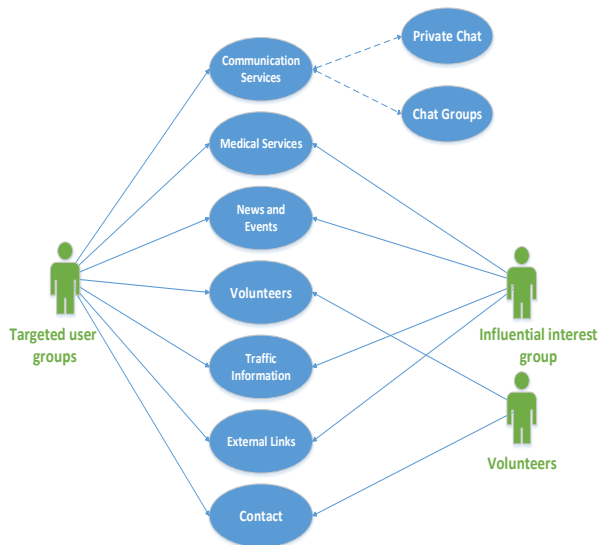
Figure 6. Use-Case diagram of Web4Health functionalities

Medical information services provide the user with reminders and lists of health institutions. The reminder consists of a calendar and an alarm. The calendar is made for the purpose of writing down the events on a specific dates, such as visiting the doctor, the pharmacy, but also as the reminder on the general events like birthdays and family visits. During each use of the application, the home page shows the scheduled events for the present day. The alarm is designed for the purpose as any other, as an alarm clock, but also as a reminder that at certain times of the day the user has to take a medicine or perform some scheduled chores. The alarm can be defined by the daily time for taking medications. The calendar automatically synchronizes with the alarm reminding the user to perform a defined task/chores. This option greatly helps individuals that are suffering from Alzheimer's disease, because it reduces user's dependence on the medical staff, i.e. increases TUG's independence.

List of health institutions is divided into a list of physiotherapists, doctors, pharmacies and hospitals and it includes their addresses. The system is implemented to show user's current location and provide directions to the nearest or/and needed medical facilities so that the user can easily find needed facility even if it is placed in an unfamiliar part of the town.

As a large number of respondents have expressed a desire to communicate with people of their age, it was necessary to create a communication service. Each user has his own profile and can use two different types of service defined within communication service. There is a public chat group, based on the principle of the forum, where users can talk about their current issues. There are administrators who are responsible for the inspection and the maintenance of the content. The second part consists of private conversations that provide the user with

the ability to communicate with people of their choice without publicly sharing conversational content with others.

The service of events and news is created for establishing a better connection between the users. The TUG is frequently attending group dances, theater, drawing classes and similar group activities. The application defines events within the user's location area and provides information about places to visit.

To avoid isolated gathering of the elderly, a service of volunteering gives a possibility to volunteer and to participate in the lives of TUG and spend some time with them. Also, it is desirable to avoid socializing based only on a virtual environment, so this service makes the application very popular to use. Every volunteer is verified to determine whether the person is capable to satisfy certain customer requirements.

If the users have any questions, there is a part of the application solution for this purpose. Contact service enables communication of users and administrators and provides users with answers in the shortest possible time. Also, a certain part of the page shows frequently asked questions as well as corresponding answers.

Users can create their travel information with the aim of effective real time informing and better navigation through the transport network. They can create their own routes of interest and points of view, and review their content anytime they want. Creating of personal routes for the movement through the city traffic provides the TUG with a sense of security.

V. ACCESSIBILITY OF DESIGN AND SECURITY MEASURES

The users have different user equipment so it is necessary to make the web application accessible to all users. Therefore, it is necessary to define the natural language, and given that the users being questioned are the citizens of Croatia, the language of the web site is defined as Croatian. Since there are citizens of foreign nationality, application solution supports English language as well. Natural language enables the search engine to find the key words. The applicative solution is designed to be applied by the people with certain deficiency, so it was taken into account that animated and flickering content makes it difficult to read and the creation of such content is therefor avoided [12].

In the case of presentation of animated content, implementation scripts for creating animation enables each user to exclude or skip this effect [13]. Considering that the application solution should be operable on various mobile terminal devices and on a desktop computer, an adjustment was made to provide its operability with a large number of user's equipment. It is extremely necessary to visually separate navigation bar, main site area, registration area and other areas/parts of site so users can easily navigate. It is taken into a consideration that a blind person cannot see the content, therefore, the option of read screen text allows users to read the contents of the applicative solution and to use the application through voice commands or specially adapted keyboard for the blind. Each section of the site is specially appointed so that visually impaired people would know when to move from one part of the content to another.

The protective measures starting at the registration are intended for a smooth and safe operation of the application. The measures are implemented through three basic procedures: identification, authentication and authorization. Considering that the TUG is not aware of the potential attacks that can be executed, there are measures for the protection of the electronic mail. In case of hacking the user's account, protective measures disable access to personal data of users, such as codes that are stored in a database. Stored procedures are used for database protection to ensure that the SQL query is entered by the user who has permission to enter the stored procedures. The result of this protection is in high-reliability queries. Stored procedures reduce the burden on the server when calling a standard SQL query. In the case of the attempted attack to user's account through registration, the detection of malicious activities is performed by the client and the server. JavaScript (client) retrieves and writes down the information about the IP address, username and password. Then the client sends data to a server that communicates with the database from which data is taken and verifies them. If the data from a database and JavaScript do not match, the user is alerted. The user account will be temporarily blocked if the wrong password is entered three times in a row. An e-mail is sent to the user and the account is blocked until the user confirms that it's a real person.

VI. CONCLUSION

This research has examined the possible application of new technology development in order to improve the quality of life of the targeted user group (TUG). HTML5 and PHP web technologies proved themselves as a good choice for development of application solutions due to their characteristics and functionalities. With the help of this web tools, it is possible to make a specific type of user registration with a dynamic representation of the content and the applicative solution. Specific registration can eliminate everyday difficulties of the TUG when using Internet services. The choice of CC architecture has proved as a quality selection in a combination with HTML5 technology providing users with constant access to data at any time and from any place, regardless the terminal device they use. The appliance of mobile terminal devices within the TUG is increasing and so is their need for the integration with new technologies and the services they provide.

This application solution is adapted to the TUG in a way to satisfy users' requirements that are, based on the conducted research, proved to be significant for the development of quality service. New information and communication technologies have greatly contributed to the development of this particular application solution. Further research and development can create other effective solutions that will be able to provide accessibility of the computerized contents and services to the elderly. The significance and the contribution of this research is reflected in a possibility of affective usage of information and communication technologies and services. The effectiveness is shown in a better way of informing the user about his everyday necessities and that can be of a great significance for the TUG.

REFERENCES

[1] Center for Technology and Aging, "Technologies to Help Older Adults Maintain Independence: Advancing Technology Adoption", July 2009.

[2] Independent Age, "Older people, technology and community", 2011.

[3] Pew Research Center, (2014, April 3) *Older Adults and Technology Use* Available at: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/

[4] M.Periša, D.Peraković, S.Šarić: „Conceptual Model of Providing Traffic Navigation Services to Visually Impaired Persons", Promet - Traffic & Transportation 26, 3; 209-218, 2014.

[5] M.Periša, D.Peraković, A.Bilić-Prcić: "Possibilities of Applying ICT to Improve the Safe Movement of Blind and Visually Impaired Persons", Cutting Edge Research in Technology, InTech, ISBN 978-953-51-4272-0, 2015.

[6] M. Periša, D. Peraković, V. Remenar, "Guidelines for Developing e – Learning System for Visualy Impaired vol. 2", International Conference on Computers Helping People with Special Needs, 2012.

[7] R. D. Sanchez (2014, October 22) *Is the future of PHP brighter than JSP and ASP?,* Available at: https://www.quora.com/Web-Applications/Is-the-future-of-PHP-brighter-than-JSP-and-ASP

[8] K. Mikoluk (2013, November 27) *PHP vs. ASP.NET: Costs, Scalability and Performance,* Available at: https://blog.udemy.com/php-vs-asp-net/

[9] Socitm (2013) *Comparin Native Apps With HTML5: What is the right approach for your organisation*, Available at: http://www.governmentm.com/

[10] J. W. Rittinghouse, J. F. Ransome, "Cloud Computing Implementation, Management, and Security" CRC Press, 2010, pp. 50 – 53

[11] E. A. Marks, B. Lozano, "Executive's Guide to Cloud Computing" John wiley & sons Inc., 2010

[12] W3C, (1999, May 5) *Web Content Accessibility Guidelines 1.0,* Available at: http://www.w3.org/TR/WCAG10/

[13] W3C, (2008, December 11) *Web Content Accessibility Guidelines 1.0,* Available at: http://www.w3.org/TR/WCAG20/

Exploiting geo tools as education opportunities

Medjon HYSENAJ
Department of Geography
University of Shkoder
Shkoder, Albania
mhysenaj@unishk.edu.al

Abstract—**The reason to induce this study is the comprehensive belief to exploit GIS opportunities as constructive pillar to student performance. The offered platform improves both student collaboration and individual strategy highlighting their skills and potentialities. The assigned project claims to reduce the gap between theoretical assignments and real world issues. By so we aim to create an optimized methodology toward a sustainable development process. Spatial tools as executable units and geodatabases as storage units are the future of modern technologies whose embracement leads the analyst into a higher level of control. Due to vast implementation in multiple fields and life routines GIS technology comes as an attractive discipline toward students' curiosity, easily to be absorbed and subject to further development. The explosive growth of the GeoWeb and geographic information contributed by users through various application programming interfaces has made GIS powerful media for the general public to communicate [10].**

Keywords- WGA, spatial technology, GIS, academic, method

I. INTRODUCTION

It is perceived that universities must fulfill student requirements. In general according to students' requests academicals entities structure their curricula. Still the intention remains to offer the students high probability of a fast enrolling in a working environment. The offered educational programs are required to represent and furthermore support the current situation of the country.

In times of technological change they frequently reflect the impact of the opportunities afforded by evolving technology and the changing labour demands of the economy in which they sit [9]. In Albania issues such as natural hazards, pollution concerns, business development, logistic management, etc strictly related to spatial extension require sustainable knowledge about gis technologies. In the country there is lack of GIS specialists, that is due to few number of GIS courses offered by academicals institutions.

Contrasting growing numbers for geo-specialists request reveal the uncertainty of students approach toward this discipline. Hence through projects and sustainable support we must invert the trend with the scope to create a new generation of individuals able to apply and transmit GIS knowledge.

Maps have always been the referring point from which has found support any analyzes developed related to earth phenomena. Spatial technologies such as Geographical Information Systems, Remote Sensing and Global Positioning Systems have turned into the pillar of many issues and human concerns. The "G" part of GIS is a two-dimensional or three-dimensional map representing real world phenomena, the "I" stands for the database supporting descriptive information, ending with "S" covering the model created by the assemble of hardware, software, methods and people [3].

Using GIS as a decision-making tool is a smart way of gathering all the things you already know and placing them in a single spot so you can see the entire picture [1]. Hence derive the primary key of success for such a complex geographical system.

The right to "mix" into a single environment unlimited and diverse variables able to create the most complete panoramic of the studied issue. GIS tends to manage the ability to adapt multiple layers from multiple sources, to succeed by arranging geodatabases and statistical data into spatial extensions, conceive facilities by executing analytical queries.

GIS is quite usual to be absorbed without drawback from all levels of followers due to its massive deployment into every kind of institutional entity whether public, private or NGOs. Due to constant application of spatial technology in working environments' it comes quite natural to transmit and analyze GIS from a vast range of aspects.

With the development of communications and grid technology, mobile GIS and grid GIS technology have entered people's vision; mobile GIS combine communications technology to GIS to achieve Mobile Location Services.; Grid GIS based on grid technology, break all closure marginal and make GIS fully integrate into the Internet environment, is a completely distributed architecture [8].

II. MATERIALS AND METHODS

The paper goal is to analyze a GIS based platform in order to create a self-learning environment by design. Deliberately we aim to improve apprentice's conscience by developing their skills in the management of spatial tools. A GIS is a tool for supporting a wide range of techniques of spatial analysis, including processes to create new classes of spatial objects, to analyse the locations and attributes of objects, and to model using multiple classes of objects and the relationships between them [11].

Learners get assigned specific tasks which require minimal GIS knowledge to be performed independently without interfering to the rest of the group. The collected data meet into a unique geo-database where overlapping information has been handled by the application code.
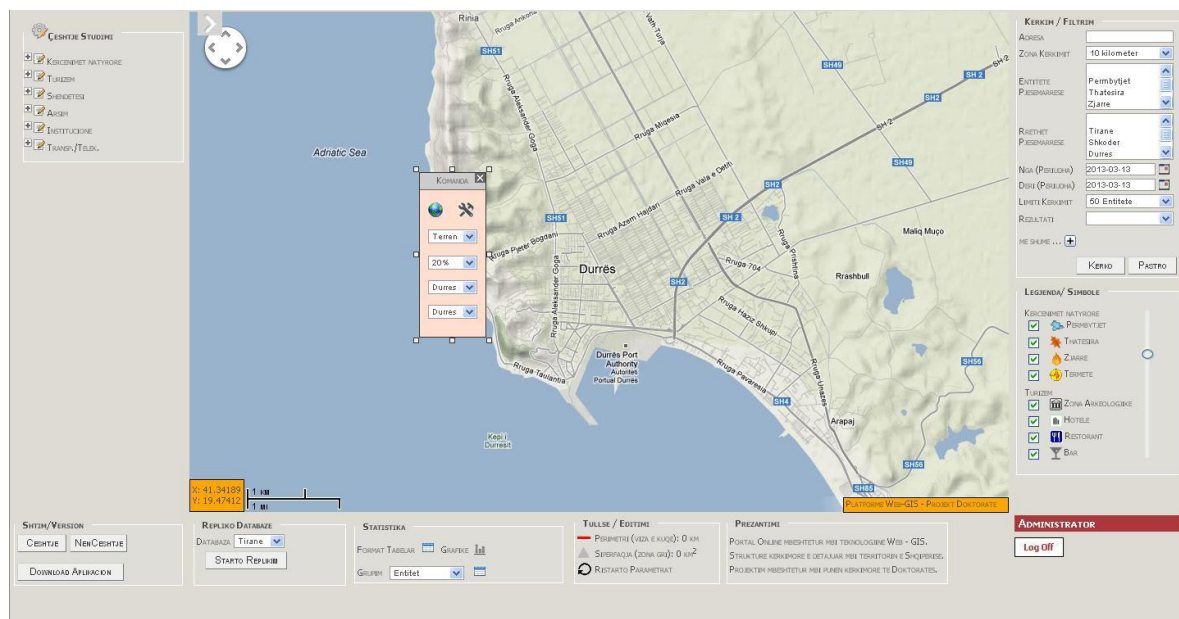
Figure 1.   Modularity backstage

Students performance is monitored constantly starting from their data input to the level of accuracy and update the selected information has been chosen. Still any error or misunderstanding can be recovered any time facilitating the operational procedure. The intention is to create a synchronous group with a satisfactory level of collaboration. Due to the online environment the platform relies on, it enables the extension of working hours beyond the institution limit, creating a pragmatic application fully adoptable in a vast variety of facilities and free from any time constrain.

GIS is a multi concentrated discipline including a wide range of profiles that turns in the eyes of the students into an attractive and stimulatory trend to be followed and practiced. Experience show us that people specialized in GIS field create larger opportunities to get involved into projects and institutional entities. Geo technologies allow students to study local to global phenomena and incorporates fieldwork [2]. They connect students to the real world, out bounding the institutional theoretical boundaries to the outside reality. By so we increase student awareness and promptitude toward the assigned task.
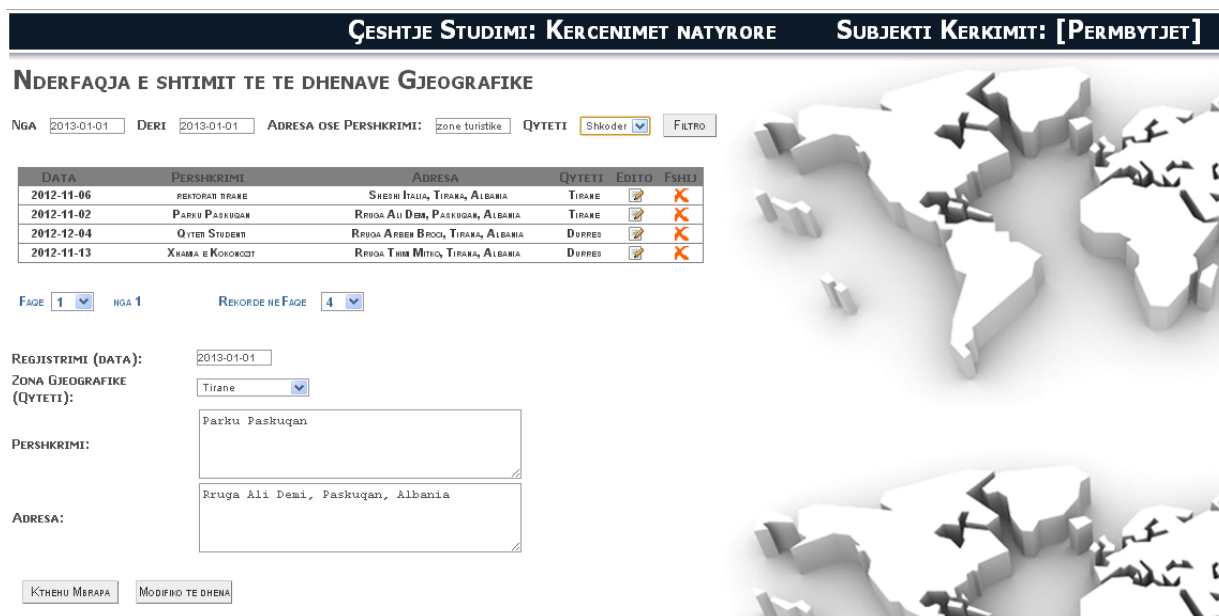
Figure 2.   Spatial data management module

To carry out the project students get assigned into small groups, usually no more than the fields of the studied area. Figure 3 show the login process, where each student receives unique credential data, further can be personalized. The thematic chosen in this paper is Tourism. Five fields mean that five students are required to carry on the project. The spatial area should not be much extended; this would require larger groups and multiple students for a single field due to the vast amount of geographical data to be processed. Taking into account the geographical area of the country, we assume that fragmenting the territory at district level is the most adaptable spatial extension. Hence it is required the student to have sustainable knowledge about the assigned area, by so the assignation process must be carefully conceived.

Further the analyst manages his own geodatabase. Each storage unit is autonomous from other units with unique attributes. This creates an optimal environment for the student. Each row reflects the geographical coordinated of the inserted data followed by the entity attributes.

The advantage offered by the application consist in that the student cooperate with each other as for geographical allocation concerns, still compile their personal environment where commands like insert, edit, delete or any other executed query are applied without interference to each other work. As such it turns comprehensive to benefit from a clearly and distinct evaluation process of both collaborative and individual student performance.

The platform can be exploited online or easily downloaded to be subject of further investigation. The modularity structure of WGA creates the perfect condition to develop student performance and creativity. On the other side students are encouraged to enrich the platform with new modules according to the revealed issues during the projects development. GIS has the potential to enhance spatial thinking itself [5]. A new understanding, or a new narrative, or a new solution to a problem, could emerge from a map because we can so easily expand our frame of reference, which is the basis for spatial relations and spatial cognition [6]. The possibility to work both online and offline is a stimuli for a larger approach toward the platform.
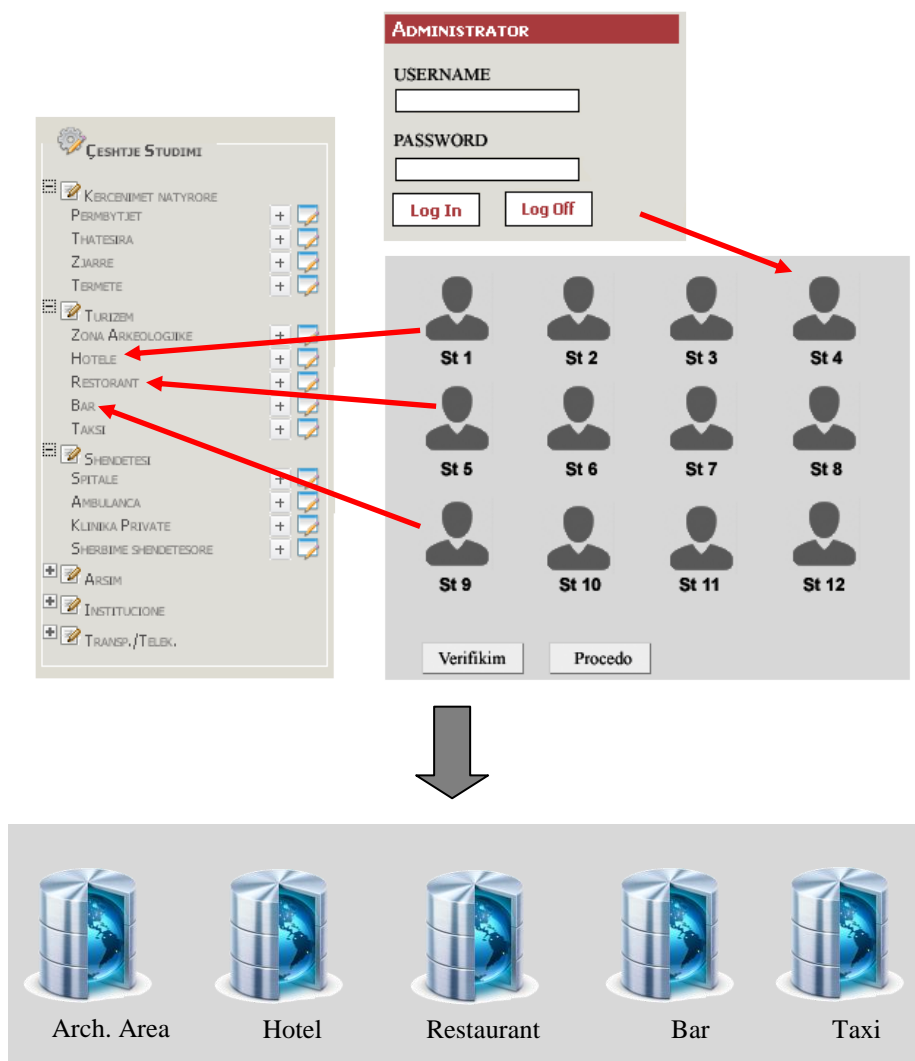
Figure 3.   Class management

III. RESULTS

Geographical information systems (GIS) are computer systems developed for the collection, storage and processing of information referenced to some form of location coordinates, with this location information usually being a key element of any analysis [7]. Today we perceive GIS technology as a powerful mainstream in various web-based maps such as Google Earth, position allocation devices such as Global Positioning System, remote sensing technologies such as satellite imagery, etc. The field of GIS is rapidly encompassing previously disparate fields, incorporating tool and skill sets into a unification of technology and science [4]. In such an eager market GIS specialist can quickly find themselves into comfortable positions, followed by a large number of duty assignment possibilities due to its chameleon singularity to be adapted and integrated into vast issues.

The paper present a web-based platform build for informative and communication purpose. Its modular structure allows the usage for academicals purpose. The application aims to improve students' knowledge on spatial technology. The case study present the process of compiling data for tourism issues through a selective group of students, each with specific and distinct assignments with the scope of evaluating their individual and collaborative skills.

The possibility to download and modify the application code enables students to overpass the line that represent simple users allowed to insert and modify geographical data by throwing the application into a higher level of complexity.

REFERENCES

[1] T. McCaffrey. (2011). Building a University of the Future, ArcNews, University of Calgary. [Online]. Available: http://www.esri.com/news/-arcnews/.

[2] J. Kerski. (2002). Developments in Technologies and Methods in GIS In Education. [Online]. Available: http://josephkerski.com/data/.

[3] J. Andrew, A. Demirci and J. Kerski, International Perspectives on Teaching and Learning with GIS in Secondary Schools, Milson, 2012, pp. 68-73.

[4] R. Thomas, Applications of GIS in the K-12 Science Classroom, Proceedings of 20 th Annual ESRI User Conference, pp. 102-120, 2000.

[5] National Research Council, Learning to Think Spatially: GIS as a Support System in the K–12 Curriculum (Washington, DC: National Academies Press), 2006, pp. 125-130.

[6] D. Stuart, Roles for GIS within Higher Education, Journal of Geography in Higher Education, University of Redlands, USA, pp. 8-10, Oct. 2009.

[7] J. David, E. Kenneth, J. Nicholas and D. David, Teaching Geographic Information Science and Technology in Higher Education, Wiley Publishing House, US, 2011, pp. 3-10.

[8] B. Fuling and Sh. Wang, Problem, Innovation and development of GIS higher education in our country, ISPRS Congress, Oct. 2008, pp. 269-272.

[9] P. Forer and D. Unwin, Enabling progress in GIS and education, 1999, pp 747–756.

[10] S. Daniel and M. Goodchild, The convergence of GIS and social media: challenges for GIS science, International Journal of Geographical Information Science, pp. 1737-1748, Nov. 2011.

[11] M. Goodchild, Geographical information science, National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA 93106, U.S.A, 1990, pp. 189-200.

Spatial evaluation for data inconsistency

Medjon Hysenaj
Department of Geography
University of Shkoder
Shkoder, Albania
mhysenaj@unishk.edu.al

Elira Hoxha
Department of Mathematics, Statistics and Applied Informatics
University of Tirana
Tirane, Albania
elira.hoxha@unitir.edu.al

Abstract—**The paper aims to emphasize the direct influence visual impact release on error checking compared to routine evaluation process. The comparison through two different methodologies relies on the belief that image perception requires minimum time evaluation, higher probability in error detection, better management of the decision-making process, compared to standard analytical environment. Talking about visual impact it is obvious that we deal with spatial extension and logically with mapping evaluation. Hence we tend to reorder concepts by disseminating strict policy to deploy map usage in academicals areas, government institutions, research entities, etc.**

Keywords- human perception, spatial technology, GIS, academic, methodology

I.   INTRODUCTION

We could stare a table of numbers all day and never see what would be immediately obvious when looking at a good picture of those same numbers [4]. Data quality can be assessed through data accuracy (or error), precision, uncertainty, compatibility, consistency, completeness, accessibility, and timeliness as recorded in the lineage data [2]. Large amounts of data expressed through statistical variations turn to be more difficult to interpret and analyze. It is expected that the level of complexity for error evaluation increase sharply with data acquisition.

The world is evolving at a rapid pace and in this environment, "information is considered power", and this is what GIS performs best, "manipulates information in favor of better decision-making" [7]. Since the first steps GIS technology has contribute to bring problem perspective into a different level of management. Map researchers developed continues efforts to present digital maps as a potential source of solution for multiple issues.

Through digital mapping the researcher create a new approach exploiting the human spatial cognition abilities. Cognition of geographic information deals with human perception, memory, reasoning, problem-solving, and communication involving earth phenomena and their representation as geospatial information [8].

When it comes for error evaluation we notice two types of cognition, spatial and numerical. Numerical cognition involves anything that deals with numbers and figures. Here we mention any tabular data registered in complex storage units such as database management systems or spreadsheet applications. Numerical cognition requires specialized people to perform evaluation task and error detection.

It means that apart from personal knowledge about the environment which the current issue is being analyzed the researcher need to have substantial applicative skills in the field of software management.

As for spatial cognition we claim anything concerned with the acquisition, organization, utilization, and revision of knowledge about spatial environments [9]. Anything related to spatial environment leads directly to mapping evaluation and as a result the need to exploit spatial tools into a geographical information system.

Spatial cognition is the last step of a long process of data manipulation, strictly connected to cognitive agents to act and interact in space intelligently and to communicate about spatial environments in meaningful ways [9]. That's why the advantage of spatial cognition is that it doesn't necessary need Gis experts to perform the evaluation. In contrast with numerical cognition the researcher is basically required to own sustainable knowledge about the studied area.

II.   MATERIALS AND METHODS

GIS gain much of their power from being able to collate and cross-reference many types of data by location; hence they can integrate many discrete datasets within a single system [3]. Hence multiple people can work on the same project, using multiple sources. By so the possibility for data error to recur is highly probable. Meaning during mapping process the user must follow a continuous and rigorous checkup. The goal is to denote that errors recur

The analytical process comes to be much easier and approachable from the user's side not only to distinguish but also to define the areas that need correction or the type of errors that have been made. The fact that the geographical data is expressed through digital mapping and not rough database rows increases the possibilities of perception that the map on the left is the correct one; meanwhile the one on the right contains corrupted data [5]. The understanding of the analytical perception of the human choice for the current situation comes as a result of several factors:

According to Peng citation we can divide inconsistency into spatial inconsistency, temporal inconsistency, attribute inconsistency and inconsistency among any combination of space, time and attribute [1]. Spatial inconsistency refers to problems that emerge due to errors in the distribution of entities into a specified spatial extension. Temporal inconsistency is associated with the period which the geographical entities reveal or belong to.

At last we claim attribute inconsistency related to mistakes developed on data elaboration mainly expressed through charts, diagrams, or geographical schema. We proceed to analyze each inconsistency form through the country territory paradigm.

Case one takes into account the most polluted shoreline zone in the country coastline. Coliform bacteria have been used to evaluate the general quality of water. Two other groups of bacteria that are present in feces are: fecal streptococci (FS) and Clostridium (FC) [10].

Our investigation consists by taking in consideration the values obtained by the Public Health Institute which has set a limit rate of 100. The attribute data the map reveals are structured improperly. As we see from the geo-database almost all checkpoints exceed the limit, which means that necessarily the chart should have expressed high values for FS and FC which in fact doesn't seem to happen.

Case two refer to spatial inconsistency giving an example of population density for the country at a comparative level. The analytical process comes to be much easier and approachable from the user's side not only to distinguish but also to define the areas that need correction or the type of errors that have been made [5].

The fact that the geographical data is expressed through digital mapping and not rough database rows increases the possibilities of perception that the map on the left is the correct one; meanwhile the one on the right contains corrupted data [5]. The understanding of the analytical perception of the human choice for the current situation comes as a result of several factors:
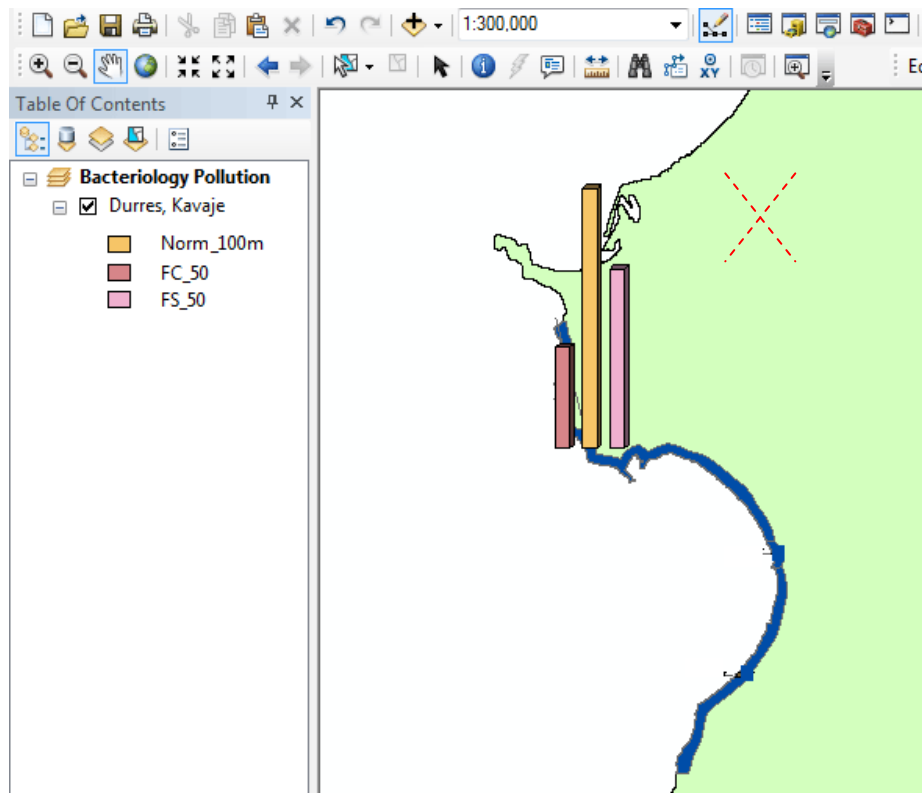
Figure 1. Attribute (Logical) inconsistency, Albania Coastal Pollution, Source: M. Hysenaj; ModTech 2015

FC-50,FS-50 values		
Ch. P.	**FC-50**	**FS-50**
1	1073	669
2	906	687
3	679	435
4	640	591
5	455	424
6	281	206
7	247	194
8	199	161
9	186	142
10	171	135
11	162	124
12	162	107
13	154	118
14	151	104
15	142	118
16	140	93
17	130	95
18	121	92
19	118	88
20	115	72
21	114	81
Average	271	226
Norm/100 ml	**100**	**100**

Ch. P.	22	23	24	25	26	27	28	29	30	31
FC-50	54	91	95	112	112	159	318	325	404	460
FS-50	20	38	39	56	50	56	256	177	296	326

[1] The user may have personal information (knowledge) about the country territory;

[2] The user may have a partial knowledge about the country; still he assumes that generally the highest population density lean near the capital and the surrounding areas, also toward coastal zones and that generally the density decrease moving from the center to the suburbs;

[3] Smaller the area extension, imply higher population density;

Image, generate such assumptions and define such error, by performing the data control, through standard database. GIS claims the ability to associate tabular information with map presentation which turns to be a powerful tool compared to routine databases. That's why through spatial tools analyze we developed the skill to locate possible errors and data inconsistency. The key stands in the relation GIS creates between spatial technology and human perception.

Case three refer to temporal inconsistency. We analyze the administrative division of the country territory between two different periods. Albania is divided into 12 administrative counties. These counties were further divided in 36 districts. Recently, the government introduced a new administrative division to be implemented in 2015.

Districts so far considered as direct sub-units of counties with be suppressed in circumstances where municipalities are reduced to 61 in total, which mean that municipalities with become the new sub-units of counties and the term district with no longer exist. Hence there is the possibility that in its "infancy" the administrative map territory of the country may turn into a source of confusion.
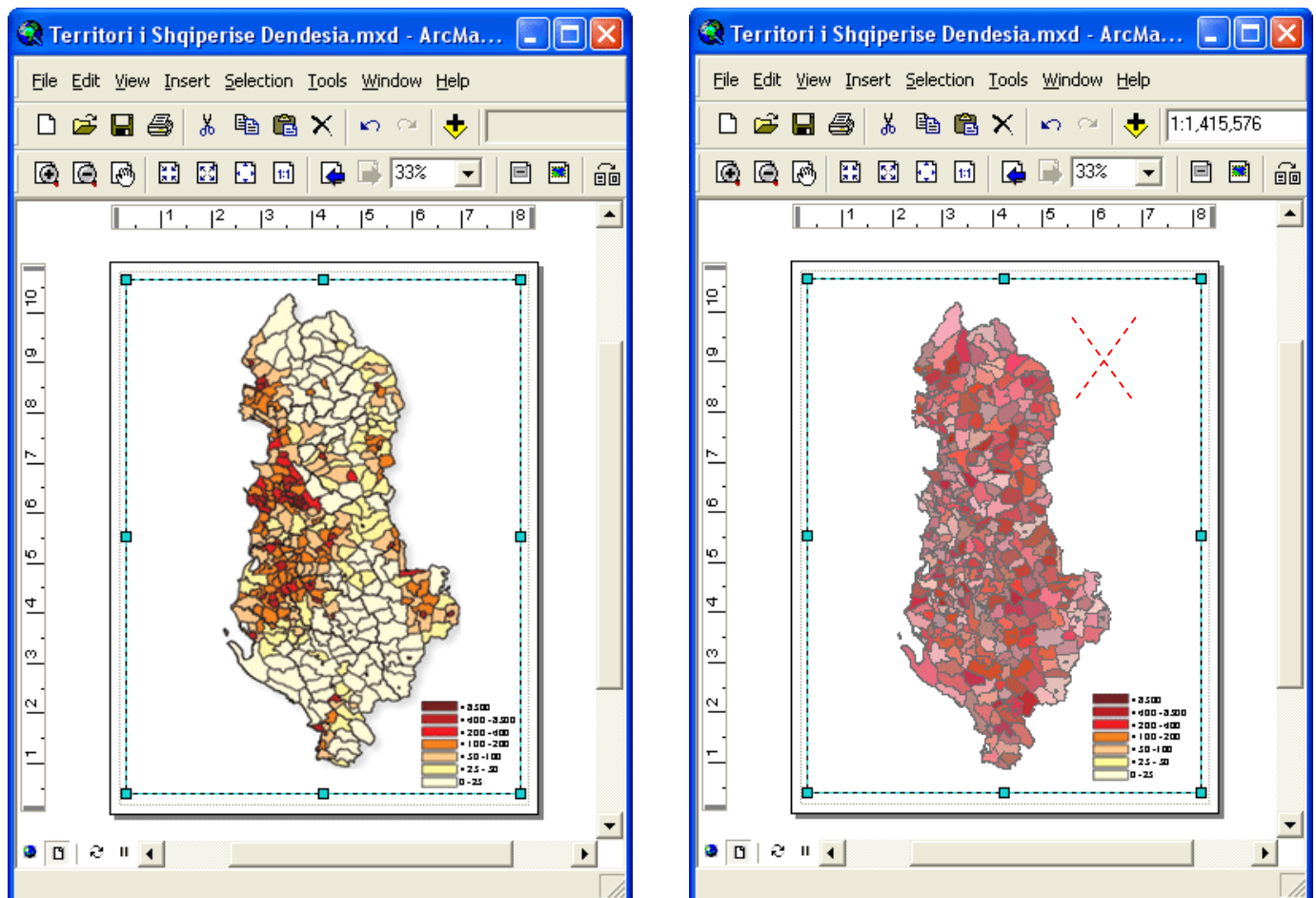
Figure 2. Spatial inconsistency, Verified issue: Population density

Figure 3. Temporal inconsistency, Verified issue: Administrative division

III. RESULTS

Maps are becoming a determinant issue with a developed ability to transform numerical and statistical information into "visual" perspective, object to a much easier analysis and manipulation process [5]. The way people conceptualize space is an important consideration for the design of geographic information systems, because a better match with people's thinking is expected to lead to easier-to-use information systems [6].

The way people identify entities within a spatial extension turns to be a powerful tool in error detecting process. Every time analyze is performed there will always be the possibility for error to happen. First of all we have to be good enough to notice something is wrong. The matter is how we are going to handle the verified situation. Categories of people believe that relying issue evaluation into a spatial environment could be much more prolific than proceeding with numerical analysis.

The examples shown in this paper reveal three types of data inconsistency, spatial, temporal and attribute. We refer to errors related to the country territory to offer a full picture of the offered environment. The goal of the paper is to cultivate researcher's conviction that visual evaluation will always offer larger opportunities compared to numerical evaluation.

Anyway geographical information systems are the ideal platform that involves both mapping presentation in-front also numerical presentation in-back. The offered possibility to integrate into a single platform statistical data associated with geographical entities should be considered a great advantage for data analyzers. It means that researchers can benefit from both options.

REFERENCES

[1] G. Peng, and M. Lan, Error detection through consistency checking, 2000.

[2] J. Chen, and P. Gong, Practical GIS: Building and Maintaining a Successful GIS, Science Press, Beijing, p.186, 1998.

[3] E. Kenneth, and J. Donald, Error, Accuracy, and Precision, Department of Geography, University of Texas at Austin, 1995.

[4] F. Stephen, Data Visualization for Human Perception, 2014.

[5] M. Hysenaj, and R. Barjami, Web GIS Albania Platform, an Informative Technology for the Albanian Territory, Informatica 36, pp. 431-439, 2012.

[6] S. Freundschuh, and M. Egenhofer, Human Conceptions of Spaces: Implication for GIS, Transactions in GIS: 361-375, 1997.

[7] M. Hysenaj, Geographical Information Systems, 2011.

[8] R. Daniel, and S. Freundschuh, Cognition of Geographic Information, A research agenda for geographic information science (pp. 61-91), 2005.

[9] C. Freksa, Spatial Cognition, An AI Perspective, pp. 1122-1128, ECAI, 2004.

[10] M. Hysenaj, Coastline planning and management through Digital Mapping Systems, ModTech 2015.

Possibilities of using Location-based Services in the Public Bicycle Systems

Dragan Peraković, Siniša Husnjak
Department of Information and Communication Traffic
University of Zagreb, Faculty of Transport and Traffic
Sciences
Zagreb, Republic of Croatia, EU
dragan.perakovic@fpz.hr, sinisa.husnjak@fpz.hr

Saša Poustecky
Graduate Student at University of Zagreb, Faculty of
Transport and Traffic Sciences
Zagreb, Republic of Croatia, EU
sasicaxp@gmail.com

Abstract—**The problem of the availability and the distribution of bicycles of the traffic models based on conventional stations can be addressed through the introduction of information services based on location. Tracking the location and the movement of each bike can provide better availability of services to the public bike system. The introduction of such forms of information and communication services in public bicycle systems can reduce the cost of introducing and maintaining the system, and create a full range of indirect services providing added value for all participants. This research focuses on the possibilities of using location-based services in the public bicycle systems, and the benefits they provide to the users and to the telecommunications operator. The analysis of the impact of location-based services to the transport process of public bike system provides an overview of the management system of public bicycles. This paper has identified the elements and participants in the value chain of delivery of information and communication services and content based on location as the possibility of their application in the system of public bicycles.**

Keywords- public bicycles system; location-based services; information and communication services

I. Introduction

Transport infrastructure in big cities has been always primarily subordinated to the car traffic, but the growing trend of the cost of living, especially transportation, begins to force a growing number of city residents to swap their daily commuting by private car to some forms of public transport. Public Bicycle System reduces traffic congestion and alleviates problems created by a lack of parking spaces, reduces air pollution and complements the public transport.

As the use of the system increases users are faced with the inability to return the bike, because the station in which they want to return the bike is already full, as well as the occasional lack of bicycles in the individual stations. While trying to solve the problem by the resettlement of bicycles on various locations in the city by truck, without a rational approach to system management problems with the availability of bikes and the possibility of return are inevitable, as well as the user dissatisfaction.

Location-based services provide the user with the possibility of the mobile terminal device location and are nowadays increasingly used in planning to improve service quality and significantly reduce operating costs.

Data collection and the analysis of user data are important stages in the creation of high-quality services. The findings of certain patterns of behavior can facilitate segmentation of users in groups. By tracking the position of mobile terminal devices installed in the bikes it is possible to predict the movement of the user, and thus through the segmentation of users increase the quality of existing services and provide new and improved contents.

This paper will show the benefits of user location technology, which can be provided by the operator of information and communication services, and the possibilities of the use of obtained information in the system of public bicycles. Previously mentioned user services can provide additional benefits, as well as the total value for all participants of the value chain. The paper will give a proposal for improving the services of existing public bicycle systems with better management and deployment of bicycles and identify potential user requirements for new communication services and content.

The purpose of this paper is successful application of modern technologies in order to design information and communication service applicable to the transport system, providing the users of the transport system with reliable information and additional benefits and the telecom operators with the added value by improving the management of transport systems.

II. Location based services

Location-based services are defined as services that use geographical information to serve mobile users, i.e. it is every application that is using the information on the current position of the mobile terminal [3].

A. LBS ecosystem

LBS are at the top of the architecture of information and communication system so they are not affecting services at lower levels and they vary according to the type of services offered to subscribers, network operator and service providers. It is difficult to determine the exact architecture of the reference model uniquely describing the components of LBS services. A simplified model is shown in Figure 1.
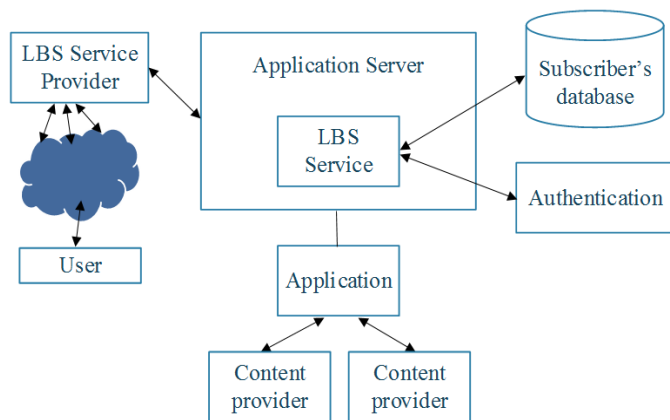
Figure 1.   Simplified eco-system of location based services [1, 13]

The basic characteristic of location services is a location information, which can be obtained by different methods of positioning, and varies according to the precision, the place of determined location and the necessary network infrastructure. In addition to the methods in which measured parameters of the signal source are based on a network (i.e. Network-based), most widely used methods are based on the interaction of mobile device and satellite global positioning systems (i.e. Handset-based) [2].

The term of location-based services covers a large area including many services such as navigation and orientation, information services, monitoring of children and disabled persons, increased subscriber's safety, calls for help, advertising, charging for services depending on the current position, intelligent transport systems, management, improvement of mobile network's performance [2, 13]. For example, a company engaged in advertising via LBS has to obtain information from the mobile provider about the location of its users. Keeping track of the change of user's location can help creating his thorough profile that can be used later for marketing, but also for forming different types of control.

The research of market's possibilities allows retrieval of the information about the market's potential and possible market's share of some services, information about the current sale and prediction of the future sales, as well locating the sales points (directions of distribution) [4]. Collection and analysis of customer data are important stages in the creation of high-quality services, and the finding certain patterns of behavior can facilitate segmentation of users in groups. By tracking the position of mobile devices, it is possible to predict the movement of a user and thereby determine his position within the segmentation.

B. Demand for accuracy of LBS environment

An important question for the development of optimal LBS system is the question of precision of the location-based service in order to provide useful information. Positioning accuracy is one of the four basic, previously mentioned security parameters of positioning. In practice, the accuracy of positioning is represented as a horizontal assessment of positioning errors for each positioning sample.

Scheme of the positioning accuracy levels which are thought to be necessary for different kinds of applications related to location-based services are shown in Table 1 [5].

TABLE I.   THE EXAMPLE OF THE DEMAND FOR ACCURACY FOR EACH APPLICATION

Application	Demand for accuracy
News	Low
Routing	High
Traffic information	Low
Advertising	Middle to low
Vehicle navigation	Middle to high
Personal navigation	High
Fleet management	Low
Vechicle tracking	Middle to high
Property tracking	High
Children surveillace	Middle to high
Electronic toll collection	Middle to high
Public management system	Middle to high
Location sensitive collection	Middle to high

Demands for the position accuracy are not the same in all areas. Rural and uninhabited areas will not require a high level of accuracy while high accuracy in urban areas will be crucial because of the potentially large concentration of services.

III.   PUBLIC BICYCLES SYSTEM

Public bicycle system brings great benefits to each city: reduces congestion and alleviates problems created by a lack of parking spaces, reduces air pollution and thus has a beneficial effect on human health, encourages physical activity of the population and complements the public transport. The introduction of the system of public bicycles and its maintenance does not create high costs for the cities and it delivers significant benefits.

The basic principle of almost all systems in the world is almost identical: users pay a membership fee, take a bike that was locked in the bike racks or electronic docking station and return the bike to any available station in the system. The basis of the system consists of a relatively dense network of tens to hundreds of separate stations for parking or hiring of these specially designed bicycles [6].

Costs and financing are the key concerns of the public bicycle system. The main costs from the operational point of view can be divided into two main categories: infrastructure and introduction, and labor costs. Depending on the system configuration, implementation costs in large systems can increase the unit price per bike, due to the cost of building the station with a greater capacity [7].

TABLE II. THE EXAMPLE OF THE MAINTAINANCE COSTS FOR THE PUBLIC BICYCLES SYSTEM [7]

Maintainance costs	Share in the total costs
Distribution of bicycles	30%
Maintainance of bicycles	22%
Maintainance of stations	20%
IT support	14%
Administration	13%
Alteration (bicycles, stations)	1%

System configuration without cycling stations can greatly reduce the cost of implementation and maintenance compared to models based on conventional stations.

IV. THE APPLICATION OF LOCATION BASED INFORMATION SERVICES IN THE PUBLIC BICYCLES SYSTEM

The main objective of the concept of public bicycles is an establishment of the sustainable transport. To avoid the problem of the bicycles shortage in one location and the crowded station in another, the optimization of the relocation of public bicycles is aided by the use of new technologies in order to determine the position of bicycles and their movement in a given period.

A. *Advanced information and communication services and technologies*

The user of a transport system does not like the insecurity on his way. If the user can be provided with personalized, real-time information on traffic and public transport, or solutions that help him make decisions about the way to reach his destination, he will have a lot less insecurity in the process of travel.

In order to simplify and cheapen the system of public bicycles, it is necessary to change the conventional way of using the method of taking the bicycle and returning it to the station. This can be achieved with built-in locking devices within the bicycles and with positioning units.

Such an approach access can greatly simplify the use of public bicycles as a service (Figure 2). It enables locating the bicycle on the local area map providing the user with an overview of the bicycles that are offered for the use [8].

Figure 2. Simplified presentation of information and communication service of the use of public bicyleces [9]

Each bicycle can be reserved for a limited time, taken over on the bicycle station location and unlock for the use. After using the bicycle can be locked anywhere and as such will be ready for the next user.

B. *Determining the position of the bicycle*

The position of a bicycle is not much of a use if it is not connected to a map of the road network. In this way, a user can be provided with a quick information for this road segment, as well as with the ability to determine the path between the current and previous position.

Each time when the user is provided with the bicycle position, the system calculates the area around the location in order to find possible segments of the road (i.e. paths), then it sorts out the sets of road segments corresponding to the area of specific geographical characteristics. These road segments must go through a selection process before choosing the one that best corresponds to the position of the bicycle. The segment of road where the bicycle is located has a similar direction and the shortest distance from the location of the bicycle.
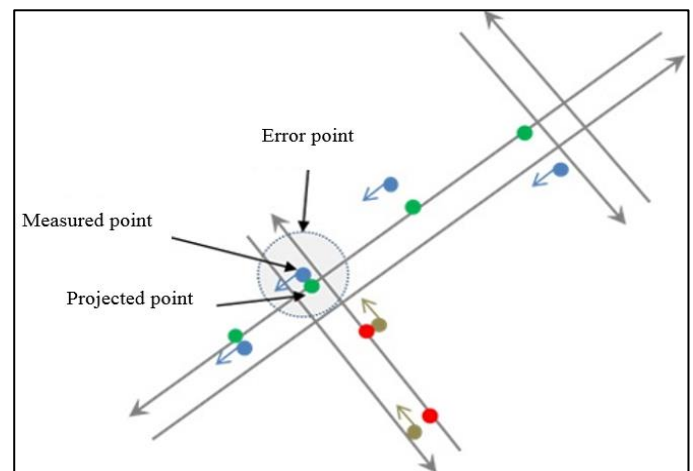
Figure 3. Adjustment of the GPS location and the road network map [10]

To calculate the speed of the travel, it is necessary to know bicycle's path between the two recorded times. The reconstitution of the speed data provides the information on the speed of the road segments where there is no data. This ensures that users continuously receive reliable information on the speed.

C. *Composition of the system*

The modern view on the information and communication services is user-oriented, i.e. it is moving in a direction of user requirements to access services using network capacity. In the implementation of solutions of a system, participants express their needs and show their interest and contribute with their knowledge and skills. Participants (i.e. stakeholders) are persons, groups or institutions whose needs and interests can contribute to the solution of a problem [11].

The end user is the main participant in the value chain generating the demand for a particular service. The user requests a service of using the bicycle and uses an application offered by the system via the mobile network.

The mobile operator (i.e. mobile service provider) is a legal or physical person who provides or is authorized to provide public communications services in the mobile network or to

rent public communications mobile network's resources and connected equipment. In the context of information services based on the location, the mobile operator provides the service of user location, and monitors and maintains information and communication systems involved in the process of locating users.

The service provider offers a system of public use of the bicycles in response to markets demand or the need for more efficient public transport. It deals with the management system, the deployment and maintaining of the bicycles.

The service provider creates a system, as well as an application for the use of the system. The application manages the information on the position of bicycles, the ability of the reservation, the pause and release of the bicycle.

Supplementary services providers are participants who are indirectly engaged in participation in the public bicycle system, and that as partners or subcontractors providing the various types of information, services and products based on sales and marketing activities. Provided information may be those of traffic or the current time, as well as the prognosis for a shorter time period that is of interest to the end user of the system.Various types of services and products can be advertised and associated with the location, i.e. the movement of the user, as well as with patterns of user behavior.

The value chain is quite fragmented with stakeholders who are in various ways involved in the design and development of the system, as well as the supply of services and products, shown in Figure 4.
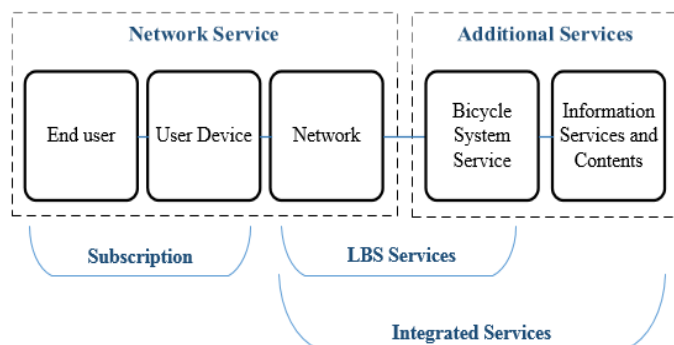
Figure 4.   The value chain of the use of information location service of the use of the public bicycles system [12]

The user is in a subscriber relationship with the mobile operator. On the other hand, the mobile network operator has the equipment and the technology for positioning, providing the service provider of the cycling system. The user provides a cycling system with a request for a service, and therefore user requests location-based services. While using these services the user creates the value for himself, but to make the service viable he must simultaneously create the value for service providers and mobile network and the provider of the cycling system service.

Accordingly, the provider of the bicycle system service system offers advertising service, supplementary services and products connecting the user with supplementary service providers, as well as the offered products. The providers of

supplementary services and products are interested in increasing their own market, so they reduce the cost of using the bicycle system, and thereby create added value for users as well as for the service provider of the bicycle system through the charges for advertising.

Ultimately, provider of the bicycle system service has an additional interest in collecting data on the user habits and profiles, which can be offered to other parties in order to improve their business, as well as other companies engaged in direct marketing.

D. System architecture and the possibility of use

The system consists of a control center, public bicycles, users, applications available via mobile phone, access to the Internet network or fixed points in the city. The control center is mainly responsible for the background management of the system functions, including management of the information on bicycles, management of the equipment and the movement of bicycles and management of user's information (Figure 5).
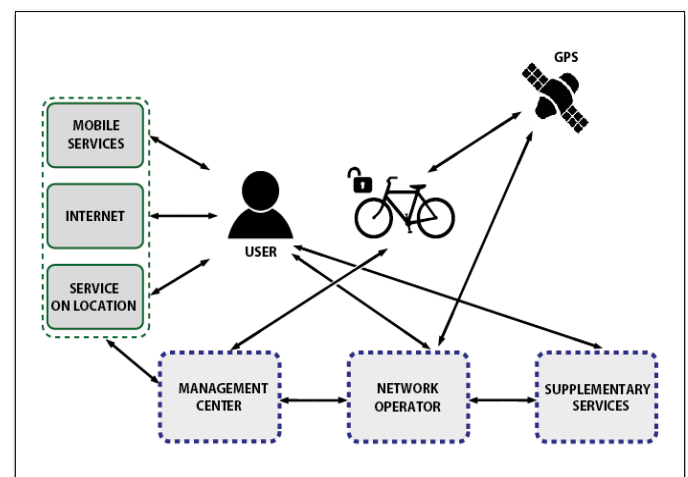
Figure 5.   System architecture and communication overview

The system is designed to give users the freedom to find and use any nearby located bicycle using the application on their mobile device (Figure 6).

Using the application system, the user sends a request for creation of a user account. After logging into the system, the user gives a request for locating and finding the nearest bicycle. By opening mobile applications' browser, the user has an insight of all the available bicycles in his location area considering the size of the wanted search area. The system can offer different variants of service due to the selected bicycle and entered parameters of the selected path:

- free of charge for a certain initial period of use,

- more affordable if the destination location is defined by the user,

- more affordable if the offered traveling route is selected by the user,

- more affordable or free of charge if using commercial and retail services,

- bonus time for the Defects Notification / failure on a bicycle,

- bonus free time for providing information about traffic and the like.
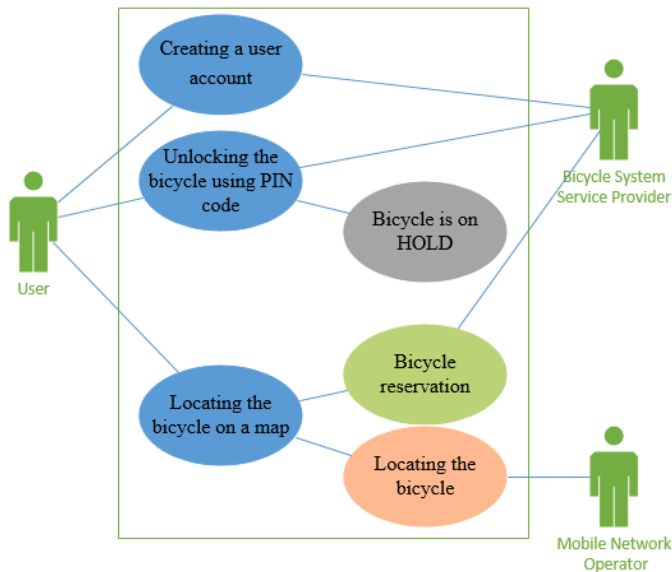
Figure 6.   Use-case diagram [12]

The system also can offer optimal solutions for a wanted destination point- a combination of using a number of bicycles or other forms of public transport (city bus, city rail, taxi etc.), and will charge a certain amount only in the case of use after a certain period of time or in the case of unfavorable destination location. The user reserves a selected bike (reservation is waiting in a shorter period), and it takes for the use by unlocking the PIN number provided by the application. The user can make a short stop when it is needed by pressing the "hold" on the bicycle locking system. This ensures that other users of the system can't take over an occupied bicycle over that stopping period.

By the end of the bicycle use, a user logs off from the system and automatically activates the locking system on the bike displaying the bicycle on the free-to-use list. In case the user does not log out of the system and the "hold" option isn't turned on, after a short time, the system will charge a certain amount from the account of the user. This prevents negligently leaving the bikes. If the user started his travel by defining the destination point or he use a particularly favorable travel route, the bike is all the time present in the list of applications with the current position in real time and an estimated time of arrival at the destination. In this way, other users can evaluate the feasibility of waiting for the incoming bike. Reservation of that bike runs until the arrival on the destination position previously defined by the first user.

The provider of the bicycle system service can cooperate with supplementary service providers to offer their customers the use of various services. On the basis of mutual agreement, various kinds of services can be offered, which will be advertised and offered to customers based on their position.

Such services can be directly related to travel by cycling, such as information about the traffic conditions or short-term weather forecasts, but they can also be associated with the vast range of services that could be interesting to the user. In addition to services, the object of advertisement can be physical products that can be bought at nearby stores located along the user's travel route. If the customer has choose to use the offer of services or products offered by the system, he achieves certain benefits while traveling by bicycle so that in some cases, travel may be cheaper, and in some cases even free, regardless of time or length of use.

The system monitors the user's position in real time and based on a specific itinerary, entered destination location, as well as predictions based on the current position or route segments, offers such services and products. By continually monitoring the behavior of individual users the system can make a segmentation of users and offer specific services and products that are based either on location or on the established patterns of behavior of the individual user. If a user, for example, is passing through a specific route and buys a specific product each day, the system informs him about such and similar offer of products in the area, and interested merchants can offer discounts in case the user decides to use their services.

V.   Conclusion

Information and communication technologies have become an important driver of everyday life and economic activity, offering a wide range of tools that can create new opportunities. The majority of people in Europe today use computers and mobile devices as a common everyday activity for a variety of purposes.

By increasing the capabilities of existing systems, as well as through the design and construction of the new global positioning systems, the offer of new information and communication services based on location is constantly growing. The introduction of information and communication services based on the location in the systems of public bicycles greatly reduces the cost of construction and maintenance of such systems considering that the tracking of the location of each bicycle provides the use of system models without stations for bicycles.

Also, the monitoring service of each bicycle in real time creates a dynamic image of cyclists and, therefore, assures better quality management system, as well as additional benefits for users. By upgrading the existing public bicycle service systems with better management and better customer approach, it is possible to identify potential user requirements for new information and communication services and contents.

The integration of new information and communication technologies in the transport system of public bicycles can encourage and increase the use of this form of economic and environmentally friendly transport mode, and thus contribute to the development of sustainable transport.

References

[1]   C. Pontikakos, T. Glezakos, T. Tsiligiridis, Location-based Services: Architecture Overview, Agricultural University of Athens, 2005.

[2] CARNet CERT, Options of the Application and Misuse of the User Positioning in GSM Networks, LSS-Faculty of Electrical Engineering, Zagreb, 2008.

[3] S. Husnjak, Development of Application for Mobile Devices in the Function of the Reminder Based on User Location, University of Zagreb, Faculty of Traffic and Transport Sciences, Master's Thesis, Zagreb, 2010.

[4] M. Jurčević, The Basics of Marketing and the New Concepts (lessons from the course Technology Marketing and Management), University of Zagreb, Faculty of Traffic and Transport Sciences, Zagreb, 2014.

[5] J. Schiller, A. Voisard. (2015, june). Location-Based Services. Available at: http://books.google.hr

[6] M. Gledec, The Introduction Program for the Bicycle Service of Zagreb, ISIP – MG d.o.o., Research and Projects in Transport, Zagreb, 2009.

[7] Optimising Bike Sharingin European Cities-OBIS, Zagreb, 2011.

[8] Univelo Boston - Hubway. (2015, July). Application. Available at: https://itunes.apple.com/au/app/univelo-boston-hubway/id659521552?mt=8

[9] Donkey Republic Bike Share Service, Copenhagen, Denmark. (2015, July). Available at: http://www.donkeyrepublic.com/

[10] Map-matching GPS Points to the Road Network. (2015, July). Available at: http://gis.stackexchange.com/questions/143909/map-matching-gps-points-to-the-road-network

[11] Guidelines for Planning Bicycle Traffic in the Urban Areas, Mobile 2020-EU Project, 2012.

[12] S. Poustecky, The Application of Information Services Based on Location in the Public Bicycles System, University of Zagreb, Faculty of Traffic and Transport Sciences, Master's Thesis, Zagreb, 2015.

[13] D. Perakovic, I. Jovovic, D. Sobota, The model of System for Marketing Promotion based on User Location, The 4ᵗʰ International Virtual Scientific Conference on Informatics and Management Science – ICTIC 2015, Thomson, Slovakia, pp. 133–138, March 2015.

Suitability of the digital forensic tools for investigation of cyber crime in the Internet of Things and Services

Jevgenijus Toldinas, Algimantas Venčkauskas,
Šarūnas Grigaliūnas

Computer Science Department
Kaunas University of Technology
Kaunas, Lithuania
jevgenijus.toldinas@ktu.lt, algimantas.venckauskas@ktu.lt
sarunas.grigaliunas@ktu.edu

Robertas Damaševičius, Vacius Jusas

Software Engineering Department
Kaunas University of Technology
Kaunas, Lithuania
robertas.damasevicius@ktu.lt, vacius.jusas@ktu.lt

Abstract – **The Internet of Things and Services (IoT&S) is a novel-networking paradigm, which allows the communication among various types of physical objects and people over the Internet. Operations and application models of IoT&S are different from the traditional networks have brought great challenges and opportunities to digital forensic technology. In this paper, we: 1) discuss specific of cybercrime in the IoT&S; 2) analyze digital evidence in various types of electronic devices; 3) analyze a taxonomy of free, commercial and open source computer forensic tools; 4) discuss suitability of the traditional digital forensic tools in the Iot&s investigation domain; and 4) summarize analyze results in tables, provide conclusion.**

Keywords- Internet of Things and Services, cyber crime, computer forensic

I. INTRODUCTION

There is no generally accepted definition of the term the Internet of Things (IoT). Generally speaking, the Internet of Things creates the ability for physical objects, which were previously often unconnected and without computing power, and people to remotely interact via the Internet. One of the threats arising from this is that, whereas people often consciously log into computers and even smartphones, they may not be aware of how they are connected to the IoT environment. The concept of the Internet of Everything, or Internet of Things and Services (IoT&S), is understood as a next evolutionary stage of the Internet of Things. It is characterised by the convergence of people, processes, data, and objects with a view to combining communications between machines (M2M), between people and machines (P2M) and between people (P2P) to deliver new or enhanced services, provide improved and broader contextual awareness, and allow for better informed and faster decisions.

Meanwhile cybercrime itself is a growing problem. Trends suggest considerable increases in the scope, sophistication, number and types of attacks, number of victims and economic damage. There are two important factors worth highlighting in the IoT&S context: *Crime-as-a-Service* and *anonymisation*. The *Crime-as-a-Service* business model drives the digital underground economy by providing a wide range of commercial services that facilitate almost any type of cybercrime. The anonymisation techniques used in parts of the Internet allow users to communicate freely without the risk of being traced. These are perfectly legitimate tools for citizens to protect their privacy. However, the features of these privacy networks are also of primary interest to criminals that abuse such anonymity on a massive scale for illicit online trade in drugs, weapons, stolen goods, illegal distribution and use of digital content, dissemination of illegal information and anonymous texts (defamatory information or propaganda), forged IDs, child sexual exploitation, etc.

Operations and application models of the IoT&S are different from the traditional networks have brought great challenges and opportunities to digital forensic technology.

In this paper, we analyze current digital forensics tools and possibilities of their application for investigation of cybercrime in IoT&S.

II. DIGITAL EVIDENCE IN VARIOUS TYPES OF ELECTRONIC DEVICES

The goal of any given forensic examination is to find facts, and to recreate the truth of an event via these facts. The examiner reveals the truth of an event by discovering and exposing the remnants ("footprints" or artifacts) of an event that have been left on the system [1].

Noblett et al. [2] define computer evidence of physical items such as chips, boards, central processing units, storage media, monitors, and printers that can be described easily and correctly as a unique form of physical evidence. Information items such as logging, description, storage, and disposition can be described easily and correctly as a unique form of informational evidence.

The challenge to computer forensic science is to develop methods and techniques that provide valid and reliable results while protecting the real evidence—the information—from harm [2]. Holder et al. [3] define four types of electronic devices and in the published FBI guide for first responders proposed potential evidence for those types: computer system, handheld devices, peripheral devices, and computer networks.

Computer system typically consists of the main base unit, sometimes called a central processing unit (CPU), data storage devices, a monitor, a keyboard, and a mouse. It may be a standalone or it may be connected to a network. Storage devices enclose hard drives, thumb or flash drives, memory cards. *Potential evidence* most commonly found in files that are stored on hard drives and storage devices and media. User-created files, user-protected files, computer-created files and other data areas are the examples of this type.

Handheld devices (personal digital assistants, electronic organizers) is a small device that can include computing, telephone/fax, paging, networking, and other features. It is typically used as a personal organizer. A handheld computer approaches the full functionality of a desktop computer system. Some do not contain disk drives, but may contain PC card slots that can hold a modem, hard drive, or other device. They usually include the ability to synchronize their data with other computer systems. *Potential evidence* includes address book, appointment calendar information, documents, e-mail, handwriting, passwords, phone book, text messages, and voice messages.

Peripheral devices such as modems, routers, printers, scanners, docking stations. *Potential evidence* relate to the devices themselves. In addition, for routers in the configuration files.

Computer networks: two or more computer systems linked by data cables or by wireless connections to enable them to share resources and data. Often include printers and data-routing devices such as hubs, switches and routers. The networked computers and connected devices themselves may be evidence that is useful to an investigation or prosecution. The data they contain may also be *valuable evidence* and may include software, documents, photos, etc.

The sources of potential digital evidence in electronic devices are as follows:

- The device and its components;

- The function(s) it performs or facilitates;

- Software; documents; photos, image files, e-mail and attachments, databases, financial information, internet browsing history, chat logs, buddy lists and event logs;

- Information stored on the device regarding its use, e.g., incoming and outgoing phone and fax numbers and recently scanned, faxed or printed documents;

- Identifying information associated with the computer system, e.g., Internet protocol (IP) and local area network (LAN) addresses, broadcast settings, and media access card (MAC) or network interface card (NIC) addresses.

III. A TAXONOMY OF COMPUTER FORENSIC TOOLS

In recent years the number of computer crime has been growing rapidly. The problem how to gather any information from a computer after an incident is becoming an important issue. There are a lot of free or commercial forensic tools on the market. We should answer questions which tool, where and when forensic investigator should be used. To answer the question the taxonomy of the forensic tools is analyzed.

Francia and Clinton [4] presents the design and implementation of an experimental Computer Security and Forensic Analysis (CSFA) laboratory and the tools associated with it. Forensic tools can be divided into three main categories (imaging, analysis and visualization) and thirteen subcategories [5].

A. *Imaging category.*

Imaging volatile memory (included on PDAs and mobile phones) - at the earliest stage, it is important to identify the nature of the power requirements for the device and to avoid removing any batteries, at least until all information resident in volatile memory has been extracted.

Disk and file imaging - in cases where the investigation involves actual seizure of a computer, disk imaging takes place once the computer has been seized according to warrant and is properly in custody. Once the imaging is complete, further analysis can then take place on the duplicate(s) with an assurance that the integrity of the original evidence is not compromised.

Write blockers are used to make absolutely certain that any unknown or unexpected disk or the file write (a write unknown to and unexpected by the investigator) is blocked. This ensures that no writes to the disk or file that is about to be imaged can take place before the imaging.

Integrity code generators and checkers are used to characterize files by calculating a fingerprint for each file.

B. *Analysis category.*

Ambient data recovery and searching of raw disk data for text strings, by sectors. The term ambient data refers to those areas on disk that are not accessible at the logical or application level. The term actually encompasses a number of separate areas on disk where forensically useful information may reside and from which it may be recoverable. One of the most important is the so-called file slack space, which refers to the space left over in the last cluster allocated to a file. Data residing in file slack space is potentially of forensic value.

File recovery and file system integrity checking tools.

File conversion (i.e., conversion of proprietary files into text files or vice versa, or between proprietary formats, to facilitate further processing).

Data filtering by date last modified and other file properties such as file or application type such as e-mail, graphics, word processing, spreadsheets, or presentation files.

Search tools sophisticated search engines with fuzzy logic capability. There are not only text searching utilities but also those tools that carry out advanced file recovery through raw, binary search of a hard disk or image file(s) for hexadecimal file headers.

Data mining tools are used to sort and analyze seemingly unrelated entities within datasets. Data mining seeks to process large volumes of data by attempting to identify or extract

meaningful relationships that can be used to establish and predict trends and patterns both past and future.

C. Visualization category.

Time-lining is concerned with time-tagging or the association of timestamps with each event or the data item of interest.

A link or relationship analysis explores the inherent relationships existing within a set of data by identifying links or relationships between each entity resident within the dataset. Link or relationship analysis depicts the relationships by graphically visualizing each entity and its interconnecting relationship(s) with other entities.

The National Institute of Standards and Technology (NIST) has taken a big step in helping to classify and search for a needed tool [6]. NIST propose a computer forensic tool taxonomy and search by their functionality.

Cloud services are aimed to provide easy, scalable access to applications, resources and services, and are fully managed by a cloud services provider (1 tool).

Deleted file recovery is used for examining file system metadata to identify deleted files and to attempt their reconstruction or recovery (10 tools).

Disk imaging is developed to create an exact copy of a computer system hard drive (16 tools).

Email parsing is used to automatically extract data you need from email messages (11 tools).

File carving is a forensics technique that recovers files based merely on the file structure and content and without any matching file system metadata. File carving is most often used to recover files from the unallocated space in a drive. Unallocated space refers to the area of the drive which is no longer holding any file information as indicated by the file system structures like the file table (3 tools).

Forensics boot environment is a forensically sound bootable CD/USB to acquire electronic media or conduct forensic analysis (3 tools).

Forensic tool suite: Mac Investigations (3 tools); Windows Investigations (5 tools).

GPS forensics techniques and tool sets specifically designed to examine GPS devices in a manner consistent with the best practices of handling digital evidence (1 tool).

Hardware write block is a specialized type of a computer hard disk controller made for gaining read-only access to computer hard drives without the risk of damaging the contents of the drive (14 tools).

Hash analysis is the first tool of choice in investigating large volumes of data, it is also routinely used to validate data integrity and identify known content. A hash function takes an arbitrary string of binary data and produces a number, often called a digest, in a predefined range. Ideally, given a set of different inputs, the hash function will map them to different outputs (10 tools).

Image analysis is used to extract all data and metadata hidden in digital images. Metadata may be written into a digital graphics file that will identify who owns it, copyright and contact information, which camera created the file, along with the exposure and descriptive information such as keywords about the photo, making the file searchable on the computer and/or the Internet. The camera writes some metadata and some data saved by the photographer and/or software after downloading to a computer (2 tools).

Infotainment & vehicle forensics acquires data from vehicle systems both at a logical and physical level using a combination of hardware and software (1 tool).

Instant messenger allows examiners to see messaging threads in their native format with support for data from SMS, MMS, Skype, iMessage, iChat, Kik, TextPlus, TextFree (3 tools).

Media sanitization/drive re-use is one key element in assuring confidentiality (6 tools).

Memory capture and analysis is used to acquire data available in volatile memory. This data include: information about processes, open files and registry handles, network information, passwords and cryptographic keys, unencrypted content that is encrypted (and thus unavailable) on disk, hidden data, worm and rootkits written to run solely in memory are all potentially stored there (12 tools).

Mobile device acquisition and analysis (21 tools).

P2P analysis because criminals take advantage of its speedier transmission and direct communication to share pornographic material (it is a federal offense to knowingly transmit child pornography, and many state laws forbid distribution of adult pornography to minors), share copyrighted material such as songs and movies, commit identity theft, commit credit card fraud, access private data stored on computers. When a target computer is seized, a digital forensic examiner needs to rapidly identify the kind and number of files that have been shared through P2P technology (4 tools).

Remote capabilities / remote forensics – the ability to acquire or analyze live remote data. Connect to and search live remote systems, collect running processes, files, RAM, etc.; or the ability to acquire or analyze remote data post mortem: search, preview, or acquire remote hard drive, remote media that has been seized but the analyst is accessing it over the network, i.e., remote lab capability (2 tools).

Social media analyzing tools are used to effectively address social media content from the leading social media networking sites such as Instagram, Facebook, Twitter and LinkedIn (5 tools in this functionality).

Software write block can be either tailored to an individual operating system or can be an independent boot disk (5 tools).

Steganalysis tool designed to extend the scope of the traditional computer forensic examinations by allowing the examiner to scan suspect media or forensic images of the suspect media for known artifacts. Artifacts may be identified by scanning the file system as well as the registry on a Microsoft Windows system (4 tools).

String search implements a string search algorithm that takes a query and a search universe and returns a response (7 tools).

Web browser forensics is used to search for the evidence left by the Web browsing activity. After retrieving data such as cache, history, cookies, and download list from a suspect's computer, it is possible to analyze this evidence for the visited Web sites, time and frequency of access, and search engine keywords used by the suspect (3 tools).

Windows registry analysis is used because Windows registry contains lots of information that can be used as evidence or helpful for forensic analysis (4 tools).

Twenty-five functionalities focused on the artifacts for evidence collection from various sources and total 156 tools classified.

Open source digital forensic [7] referenced tools organized into the following categories.

Bootable environments - used to boot a suspect system into a trusted state (3 tools).

Data acquisition - used to collect data from a dead or live suspect system (26 tools).

Volume system - used to examine the data structures that organize media, such as partition tables and disk labels (15 tools).

File system - used to examine a file system or disk image and show the file content and other metadata (20 tools).

Application - used to analyze the contents of a file (i.e. at the application layer) (54 tools).

Network - used to analyze network packets and traffic. This does not include logs from network devices (11 tools).

Memory - used to analyze memory dumps from computers (3 tools).

Frameworks - used to build custom tools (8 tools).

Eight functionalities are focused on the artifacts for evidence collection from various sources and total 140 tools referenced (the same tool is referenced in more than one category). Table I illustrates computer forensic tools taxonomy proposed categories and subcategories and number of tools referenced.

TABLE I.   COMPUTER FORENSIC TOOLS TAXONOMY

Taxonomy author	Categories	Subcategories	Number of referenced tools
Francia and Clinton	Imaging	4	n/a
	Analyzing	7	
	Visualization	2	
NIST	25	n/a	154
Open Source	8	n/a	140

IV. LICESED AND OPEN SOURCE COMPUTER FORENSIC TOOLS

Internet Evidence Finder (IEF) [8] is a digital forensics software solution used to find, analyze and present digital evidence found on computers, smartphones and tablets. IEF helps to recover evidence like instant messaging chats, social media artifacts and web history to support a wide variety of cases and investigations from child exploitation, to terrorism and intellectual property theft. Free Magnet Forensics Company tools allows examiner:

- to decrypt the Dropbox filecache.dbx and config.dbx files, which are encrypted SQLite databases;

- to check local physical drives on the system for encrypted volumes created by TrueCrypt, PGP®, Bitlocker®, and other full disk encryption products;

- to assist in putting more meaning around the Google Maps tile files left behind when viewing maps or getting directions on the Google Maps website;

- to take a list of URLs and saves scrolling captures ("snapshots") of each page.

Licensed Magnet Forensics Company tools are:

- IEF - designed for the use in a lab environment on a forensics workstation. It recovers 260+ unique Internet artifacts from Windows or Mac file systems;

- IEF Triage - designed for the use on live systems in the field. It recovers 260+ unique Internet artifacts from Windows or Mac file systems. The tool runs directly from the USB thumb drive on a target computer, and is particularly valuable for running a quick search on-scene, taking a live RAM capture, and/or checking for disk encryption;

- IEF Bundle (includes IEF license and IEF Triage) - designed for investigators (or forensic teams) who work in a lab environment and in the field on live systems. It enables to get both our standard and triage licenses together on a single USB dongle for added convenience and costs savings (use both IEF editions without having to pay for two separate licenses).

Forensic Toolkit (FTK) [9] is a court-cited platform of digital investigations built for speed, stability and the ease of use. It provides comprehensive processing and indexing up front, so filtering and searching is faster than with any other product. FTK can be set up for distributed processing and incorporate web-based case management and collaborative analysis. FTK is database-driven which provides the stability necessary to handle large data sets. In addition, FTK components are compartmentalized allowing processing workers to continue processing data even if the GUI stalls. By processing and indexing data up front and leveraging a powerful dtSearch engine, as well as a full-featured regular expression engine, FTK produces fast and accurate results. The FTK offers:

File and Disk Encryption Support - with proper credentials you can decrypt technologies, such as Credant, SafeBoot, Utimaco, PGP, Guardian Edge, Sophos Enterprise and S/MIME and more. FTK can also decrypt hundreds of file types. It will decrypt files during processing using provided passwords, or encrypted files can be selected within FTK and

sent to the built-in Password Recovery Toolkit® module for password recovery;

- Advanced Gallery View for Images and Video with Automatic Identification of Explicit Images quickly identifies critical image and video files. In addition, FTK identifies pornographic images automatically, which is an invaluable feature for law enforcement. It not only recognizes flesh tones, but also shapes and image orientations that could be of pornographic nature;

- Superior Email Analysis - FTK supports 18 different email types, including Notes NSF, Outlook PST/OST, Exchange EDB, Outlook Express DBX, Eudora, EML (Microsoft Internet Mail, Earthlink, Thunderbird, Quickmail, etc.), Netscape, AOL and RFC 833;

- Single-Node Enterprise gives the possibility to get a full remote analysis and incident response capabilities of the AD Enterprise. It previews, acquires and analyses hard drive data, peripheral device data, and volatile/memory data;

- Volatile and Memory Analysis enumerates all running processes, even those hidden by rootkits, and display associated DLLs, network sockets and handles in context. Search memory, automatically map hits back to a given process, DLL or piece of unallocated space, and dump the corresponding item. VAD tree analysis exposes registry artifacts in memory, parsing and displaying handle information. Supports Windows® 32- & 64-bit, Apple®, UNIX® and Linux®;

- Internet Artifact Analysis provides broad browser support with SQLite parsing and includes 40 Internet artifact carvers for popular web applications, including Facebook, Google Drive ("Docs"), Google Chat, ICQ 7M, Skype, DropBox, Torrent and many more;

- Data Visualization for Automated Timeline Construction and Social Analysis - the Visualization technology in FTK displays data in timelines, cluster graphs, pie charts and more.

Covert Forensic Imaging Device (CFID) [10] designed for forward deployed military, intelligence, and law enforcement personnel who need a simple, small, portable and inconspicuous solution for imaging, cloning and wiping data from portable media such as USB and SD Cards. The CFID offers a battery powered multipurpose solution for the field and lab use. It takes four clicks on the simple, fat-finger friendly touch screen to begin an E01 or dd style image of a suspect media device. The CFID also employs haptic feedback (vibration) to alert the operator when an imaging process is complete. The CFID performs forensic acquisitions while ensuring the following industry standard conditions are met:

- MD5 and SHA (1,256) check-sums are computed;

- Source device is write-blocked to ensure no data has been ever modified on it;

- Images are created in standard DD or E01 formats to comply with any major analysis tool;

- Block size can be set for increased performance during the image process from one sector, to 1 MB;

- Compression can be enabled to optimize destination drive for space when imaging large drives.

BlackLight [11] is a multi-platform forensic analysis tool that allows examiners to quickly and intuitively analyse digital forensic media. BlackLight is capable of analysing data from Mac OS X computers, iOS devices (iPhone, iPad, iPod Touch) and Windows computers. It is compatible with all leading logical and physical forensic image formats. Examiners may use BlackLight as a time-saving data triage tool, or as an advanced forensic examination tool depending on circumstances. The BlackLight features:

- Details view provides a visual device configuration and usage such as device type, iOS / OS version, serial number, UDID, and IMEI. Also include artefact summary statistics for documents, emails, movies, calls, voicemail, account information of the device user and common Internet account information for applications such as Twitter and iCloud; recent usage history, including dialled phone numbers, last running applications, and most recent web-based location searches;

- Signature file filter view includes examiner-defined filter options to quickly pinpoint relevant data within large data sets. Filter criteria include: file name, kind, size, or extension, date of creation, modification, or access, picture metadata attributes, including GPS coordinates and camera (iOS device) type, positive and negative hash set filtering;

- Media view has built-in support for all commonly used picture and video file types, and it includes several helpful and examiner-oriented analysis features, such as: built-in GPS mapping, proprietary skin tone analysis algorithm, video frame analysis;

- Social view provides examiners with comprehensive support for the most popular social media applications. As the proliferation of social media has infiltrated popular culture, understanding the data from these apps has become integral to an overwhelming number of investigations. BlackLight will automatically parse and display data from Facebook, Twitter, LinkedIn, Foursquare;

- The messaging view allows examiners to see messaging threads in their native format with support for data from SMS, MMS, Skype, iMessage, iChat, Kik, TextPlus, TextFree.

OSForensics [12] allows the examiner to identify suspicious files and activity with hash matching, drive signature comparisons, e-mails, memory and binary data. It allows to quickly extract forensic evidence from computers using advanced file searching and indexing and enables this data to be managed effectively. The OSForensics features:

- Search by filename, size and time, within file contents, through email archives from Outlook, ThunderBird, Mozilla, etc;

- Recover and search deleted files, uncover recent activity of website visits, downloads and logins;

- Collect detailed system information;

- Password recovery from web browsers, decryption of office documents;

- Discover and reveal hidden areas in your hard disk;

- Browse Volume Shadow copies to see past versions of files; Verify and match files with MD5, SHA-1 and SHA-256 hashes;

- Find misnamed files where the contents don't match their extension;

- Create and compare drive signatures to identify differences;

- Timeline viewer provides a visual representation of system activity over time;

- File viewer that can display streams, hex, text, images and meta data;

- Email viewer that can display messages directly from the archive;

- Registry viewer to allow easy access to Windows registry hive files;

- Raw disk viewer to navigate and search through the raw disk bytes on physical drives, volumes and images;

- Web browser to browse and capture online content for offline evidence management;

- SQLite database browser to view the and analyse the contents of SQLite database files;

- ESE DB viewer to view and analyse the contents of ESE DB (.edb) database files, a common storage format used by various Microsoft applications;

- Prefetch viewer to identify the time and frequency of applications that been running on the system, and thus recorded by the O/S's.

Steganography Analysis and Research Center (SARC) [13] has developed state-of-the-art steganography detection and extraction capabilities that address the needs of the computer forensic examiners and information technology security personnel in law enforcement, government agencies, military organizations, the intelligence community, and the private sector. The SARC provides the largest commercially available repository of the steganography application and develops the most advanced steganalysis tools, techniques, and procedures to find and extract hidden information. The SARC steganalysis products and features:

- StegAlyzerAS is a steganalysis tool designed to extend the scope of traditional computer forensic examinations by allowing the examiner to scan suspect media or forensic images of the suspect media for known artifacts of over 1,225 steganography applications. Artifacts may be identified by scanning the file system as well as the registry on a Microsoft Windows system. StegAlyzerAS allows for identification of files by using CRC-32, MD5, SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512 hash values stored in the Steganography Application Fingerprint Database (SAFDB). SAFDB is the largest commercially available steganography hash set. Known registry keys are identified by using the Registry Artifact Key Database (RAKDB). RAKDB is the only commercially available steganography registry key database;

- StegAlyzerSS is a steganalysis tool designed to extend the scope of traditional computer forensic examinations by allowing the examiner to scan suspect media or forensic images of the suspect media for over 55 uniquely identifiable byte patterns, or known signatures, left inside files when particular steganography applications are used to embed hidden information within them. Automated extraction algorithms unique to StegAlyzerSS can be used to recover hidden information. StegAlyzerSS extends the signature scanning capability by also allowing the examiner to use more traditional blind detection techniques for detecting whether information may be hidden within potential carrier files;

- StegAlyzerFS is a steganalysis tool designed to perform rapid field triage on the suspect media on computers to detect the use of steganography to conceal information. Often it is necessary quickly identify potential evidence of the concealed information while at the scene. If the information was hidden using a steganography application, currently deployed computer forensic triage tools will not detect it. A suspect computer can be booted from the StegAlyzerFS device and results can be obtained in a matter of minutes. StegAlyzerFS detects any of the files associated with over 1,225 applications in SAFDB. In addition, StegAlyzerFS detects over 55 uniquely identifiable byte patterns, or known signatures, left inside files when particular steganography applications are used to embed hidden data within them;

- Sensitive data leakage is of the utmost concern to corporate management. Data Loss Prevention (DLP) solution providers offer products with a wide range of functionality and capability. However, none of these products detect inner use of steganography. StegAlyzerRTS is a commercially available network security appliance capable of detecting digital steganography applications and the use of those applications in real-time. StegAlyzerRTS offers a "drop-in, turn-key" capability that will not affect network throughput. StegAlyzerRTS detects insider

downloading steganography applications by comparing the file fingerprints, or hash values, to a database of a known file, or artifact, hash values associated with over 1,225 steganography applications. StegAlyzerRTS also detects insider use of steganography applications by scanning files entering and leaving the network for known signatures of over 55 steganography applications. StegAlyzerRTS detects insider theft of sensitive information hidden inside other seemingly innocuous files, which sent to an external recipient as an e-mail attachment or posted on a publicly accessible web site.

Infotainment and Vehicle System Forensics (iVe) [14] is a vehicle system forensic tool that acquires user data from vehicles, and allows forensic examiners and investigators to analyse it. Vehicle Infotainment systems store a vast amount of the user-related data such as recent destinations, favourite locations, call logs, contact lists, SMS messages, emails, pictures, videos, and the navigation history of the vehicle. iVe directly interfaces with vehicle systems via specially-designed hardware. It can acquire a full or partial binary image and decode the data. It can recover deleted information from either image type. iVe can decode and parse data such as:

- Vehicle/system information - serial number, part number, original VIN number, build number;

- Installed application data - Weather, Traffic, Facebook, and Twitter. Connected devices - phones, media players, USB drives, SD cards, wireless access points;

- Device Information - device IDs, calls, contacts, SMS, audio, video, images, access point information;

- Navigation data - track logs and track points, saved locations, previous destinations, active and inactive routes;

- Events - doors opening/closing, lights On/Off, Bluetooth connections, Wi-Fi connections, USB connections, system reboots, GPS time syncs, odometer readings. iVe currently supports Ford, Lincoln, Mercury, Buick, Cadillac, Chevrolet, and GMC vehicles from 2007-2015.

XRY [15] is a software application designed to run on the Windows operating system which allows the examiner to perform a secure forensic extraction of data from a wide variety of mobile devices, such as smartphones, GPS navigation units, 3G modems, portable music players and the latest tablet processors such as the iPad. XRY is a purpose-built software-based solution, complete with all the necessary hardware for recovering data from mobile devices in a forensically secure manner. Logical and physical analysis tools available for supported devices can produce a combined report containing both live and deleted data from the same handset. XRY Complete functions:

- SIM card reading and cloning,

- Mobile device logical and physical examinations,

- GPS devices physical examinations,

- Memory card logical and physical examinations,

- Hex viewer,

- Hash algorithms,

- File signature analysis,

- Selective extraction of data.

Automated Image and Restore (AIR) [16] is a graphical user interface front-end to dd/dc3dd. This tool (Fig. 4.8) is designed to make the task of creating forensic images of the digital media easier for investigators and incident response personnel. AIR provides a convenient interface to execute the dd set of commands. License: GNU General Public License version 2.0 (GPLv2) for Unix OS, programming language Pearl. Features include:

- Auto-detection of IDE and SCSI drives, CD-ROMs, and tape drives;

- Choice of using either dd or dc3dd;

- Image verification between source and copy via MD5 or SHA1/256/384/512;

- Image compression/decompression via gzip/bzip2;

- Image over a TCP/IP network via netcat/cryptcat;

- Supports SCSI tape drives Wiping (zeroing) drives or partitions;

- Splitting images into multiple segments;

- Detailed logging with date/times and complete command-line used.

The Sleuth Kit® (TSK) [17] is a library and collection of command line tools that allow a user to investigate disk images. The core functionality of the TSK allows you to analyse volume and file system data. The plug-in framework allows you to incorporate additional modules to analyse file contents and build automated systems. The library can be incorporated into larger digital forensics tools and the command line tools can be directly used to find evidence. The source code in TSK distributed under several licenses. Each source code file identifies the license that applies to its contents. Some of the files in TSK core (non-framework) have roots in The Coroner's Toolkit (TCT) and distributed under the IBM Public License. These files are limited to the file system code and mainly for the FFS and Ext2 file systems. Files that have been created since the fork are released under the Common Public License. This includes all other files in the library. Note that the Common Public License is a generic form of the IBM Public License. The framework code is distributed under the Common Public License. TSK also distributes a striped down copy of GNU binutils strings, which has a GPL 2 license. Features include:

- Analyses raw (i.e. dd), Expert Witness (i.e. EnCase) and AFF file system and disk images;

- Supports the NTFS, FAT, UFS 1, UFS 2, EXT2FS, EXT3FS, Ext4, HFS, ISO 9660, and YAFFS2 file

systems (even when the host operating system does not or has a different endian ordering);

- Tools can be run on a live Windows or UNIX system during Incident Response. These tools will show files that have been "hidden" by rootkits and will not modify the A-Time of files that are viewed;

- List allocated and deleted ASCII and Unicode file names;

- Display the details and contents of all NTFS attributes (including all alternate data streams);

- Display file system and meta-data structure details;

- Create time lines of file activity, which can be imported into a spread sheet to create graphs and reports;

- Lookup file hashes in a hash database, such as the NIST NSRL, Hash Keeper, and custom databases;

- Organize files based on their type (for example all executables, jpegs, and documents are separated). Pages of thumbnails can be made of graphic images for quick analysis.

The Sleuth Kit is written in C and Perl and uses some code and design from TCT. The Sleuth Kit has been tested on: Linux, Mac OS X, Windows, CYGWIN, Open & FreeBSD, and Solaris.

Autopsy [18] is a digital forensics platform and graphical interface to The Sleuth Kit and other digital forensics tools. Autopsy offers the same core features as other digital forensics tools and offers other essential features, such as web artefact analysis and registry analysis that other commercial tools do not provide.

Digital Evidence & Forensic Toolkit (DEFT) [19] is the Linux distribution. DEFT is made up of a GNU / Linux and DART (Digital Advanced Response Toolkit) suite dedicated to digital forensics and intelligence activities.

This distribution is currently used during the laboratory hours of the Computer Forensics course held at the University of Bologna and in many other Italian universities and private entities. It is also one of the main solutions employed by law enforcement agencies during computer forensic investigations.

DEFT 8.1 most important package and tool list:

- File Manager with the status of a disk mount;

- Full support for Bitlocker encrypted disks;

- The Sleuthkit 4.1.3;

- Digital Forensics Framework 1.3;

- Full support for Android and iOS 7.1 logical acquisitions;

- Java Decompiler GUI;

- Skype Extractor 0.1.8.8;

- Maltego 3.4 Tungsten;

- A new version of the OSINT browser.

Digital Forensics Framework (DFF) [20] is a free and Open Source computer forensics software built on top of a dedicated Application Programming Interface. It can be used by both professional and non-expert people in order to quickly and easily collect, preserve and reveal digital evidences without compromising systems and data. DFF has three versions: DFF (free), DFF Pro (commercial) and DFF Live (commercial). DFF Pro is an enhanced edition built from the free and Open Source project. Get new features earlier, professional support and deploy DFF with confidence for your digital investigations. DFF features:

- Preserve digital chain of custody (Software write blocker, cryptographic hash calculation);

- Access to local and remote devices;

- Read standard digital forensics file formats (Raw, Encase EWF, AFF 3);

- Virtual machine disk reconstruction; Windows and Linux OS forensics;

- Quickly triage and search for metadata;

- Recover hidden and deleted artefacts;

- Volatile memory forensics.

DFF Pro computer forensics software has advanced features:

- Automation engine;

- Malware detection;

- Engine connector;

- Windows event logs management;

- Hash database reduction;

- HTML reporting.

DFF Live designed to provide a comprehensive and portable digital investigation arsenal. In addition to the most recognized Free and Open Source digital forensics software DFF Live do:

- Collect volatile data before shutting down a system;

- Boot on a dedicated Linux forensics distribution;

- Perform forensics analysis of systems and local storage devices;

- Preserve evidences in a secured storage area.

V.   SUITABILITY OF THE TRADITIONAL DIGITAL FORENSIC TOOLS IN THE IOT&S INVESTIGATION DOMAIN

IoT&S is the global network connecting any smart object [21]. Objects that are not smart themselves are embedded with smartness and communication capabilities using technologies

such as Radio Frequency Identification (RFID), sensors and other forms of the embedded computing [22].

All tools analysed in this paper and referenced in NIST taxonomy are suitable for evidence investigations in computer systems. Some of them BlackLight, DEFT, IEF are able to recover mobile forensic artefacts from iOS (iPhone, iPad, iPod Touch) and Android powered smartphones and tablets, including GPS coordinates. Others (StegAlyzerAS, StegAlyzerSS, StegAlyzerFS, StegAlyzerRTS) have developed for steganography detection and extraction capabilities.

There are tools developed for specific purpose only as follows:

- Imaging, cloning and wiping data from portable media such as USB and SD Cards (CFID);

- Making the task of creating forensic images of digital media easier for investigators and incident response personnel (AIR).

The NIST taxonomy does not include peripheral devices (modems, routers, scanners, printers, docking stations). Computer networks (two or more computer systems linked by data cables or wireless connections to enable them to share resources and data) limited to the categories "Email parsing", "Instant messenger", "P2P analysis", "Social media", "Web Browser Forensics". Only one IEF tool is classified in the "Cloud Services" category and in the "P2P Analysis" category. IEF support networks are GigaTribe, BitTorrent, Ares, eMule and Gnutella. In the category "Social Media" BlackLight and IEF, support social networks, such as Facebook, Twitter, LinkedIn, Foursquare, MySpace, Google+, and Flickr.

Category "Web Browser Forensics" support three tools: BlackLight, IEF and OSForensic. Category "Instant messenger" support two tools: BlackLight and IEF. A number of tools support device type of computer networks due to reference in "Email parsing" category, those tools are: BlackLight, DEFT, DFF, FTK, IEF and OSForensic. StegAlyzerRTS also detects the insider use of steganography applications by scanning files entering and leaving the network for known signatures.

The IoT&S domain areas not only include computer systems, handheld devices, peripherals devices, computer networks, cloud computing, virtualization but also other types of the IoT&S such as sensors, RFID devices, etc. Forensics in the IoT&S will therefore encompass forensics in all mentioned areas. We did not find tools referenced in the NIST taxonomy for digital forensic investigations in others types of the IoT&S such as sensors network, RFID, embedded and artificial intelligence devices.

The iVe tool referred to in the NIST taxonomy may be classified as other types of the IoT&S because it used forensic investigations for vehicle systems.

XRY Complete can take place in classification as a tool for others types of the IoT&S because it performs a secure forensic extraction of data from a wide variety of mobile devices, such as smartphones, GPS navigation units, 3G modems, portable music players and the latest tablet processors such as the iPad.

Table II highlights suitability of the analysed tools for forensic investigation in the IoT&S by types of electronic devices.

TABLE II.   COMPUTER FORENSIC TOOLS TAXONOMY

Tool Name	Traditional forensics and NIST taxonomy category	Suitability for IoT&S domain				
		Computer system	*Handheld Devices*	*Peripheral devices*	*Computer networks*	*Other IoT&S devices*
Autopsy	Deleted File Recovery Hash Analysis String Search	+	-	-	-	-
Automated Image and Restore	Disk Imaging	+	-	-	-	-
BlackLight	Deleted File Recovery Email Parsing File Carving Forensic Tool Suite (Mac Investigations) Forensic Tool Suite (Windows Investigations) Hash Analysis Image Analysis (Graphics Files) Instant Messenger Mobile Device Acquisition and Analysis Social Media String Search Web Browser Forensics Windows Registry Analysis	+	+	-	+	-
Covert Forensic Imaging Device (CFID)	Disk Imaging	+	-	-	-	-
Digital evidence & forensic toolkit (DEFT)	Deleted File Recovery Email Parsing File Carving Forensic Tool Suite (Windows Investigations) Hash Analysis Memory Capture and Analysis Software Write Block String Search Windows Registry Analysis	+	+	-	+	-
Forensic Toolkit (FTK)	Deleted File Recovery Email Parsing File Carving Forensic Tool Suite (Windows Investigations) Hash Analysis Memory Capture and Analysis String Search	+	-	-	+	-
Internet Evidence Finder (IEF)	Cloud Services Email Parsing Forensic Tool Suite (Windows Investigations) Instant Messenger Memory Capture and Analysis Mobile Device Acquisition and Analysis P2P Analysis Social Media Web Browser Forensics	+	+	-	+	-
Infotainment and Vehicle System Forensics (iVe)	Infotainment & Vehicle Forensics	-	+	-	-	+
OSForensics	Deleted File Recovery Email Parsing File Carving Forensic Tool Suite (Windows Investigations) Hash Analysis Memory Capture and Analysis String Search Web Browser Forensics Windows Registry Analysis	+	-	-	+	-

TABLE II. COMPUTER FORENSIC TOOLS TAXONOMY (CONT.)

Tool Name	Traditional forensics and NIST taxonomy category	Suitability for IoT&S domain				
		Computer system	*Handheld Devices*	*Peripheral devices*	*Computer networks*	*Other IoT&S devices*
Steganography Analyzer	Steganalysis Artifact Scanner Steganalysis Field Scanner Steganalysis Real-Time Scanner Steganalysis Signature Scanner	+	-	-	+	-
The Sleuth Kit	Deleted File Recovery Hash Analysis	+	-	-	-	-
XRY Complete	Deleted File Recovery Hash Algorithms Mobile Device Acquisition and Analysis GPS Devices Physical Examinations Memory Card Logical and Physical Examinations File Signature Analysis 3G Modems and Portable Music Players Analysis	-	+	+	-	+

Batalla and Krawec [23] propose a novel architecture and mechanisms of the ID (IDentifier) layer for the IoT&S purpose. Such a solution solves the registration of the IoT&S as well as the search and delivery of the information related to them. ID layer embedded in the network level instead of traditional overlay solutions where users gain easier access to the IoT&S resources and native support for multicast. The authors introduce a hierarchical addressing of objects following their physical location, which contains the services offered, by objects or groups of objects. The appropriate hierarchical tree sends the requests of data by the applications as well as the information from objects and/or services (data). The nodes in the tree are able to cache the data and store for further requests during the validity period of the data. Sensors may go on sleep mode for saving energy consumption, whereas sensing data are still available for applications.

VI. CONCLUSIONS

Operations and application models of the IoT&S are different from the traditional networks have brought great challenges and opportunities to digital forensic technology.

There is a huge number of available computer forensic tools from standalone packages to complex integrated tools, developed for wide range crime investigations. The National Institute of Standards and Technology (NIST) proposes computer forensic tool taxonomy and search by their functionality.

The NIST computer forensic tool taxonomy does not include the definition of functionality and classification for the IoT&S such as sensors, embedded computing devices and others with communication capabilities using RFID technologies.

Existing traditional computer forensic tools are insufficient for cyber-crime investigation in IoT&S.

In order to investigate cyber-crime on the IoT&S new, innovative digital forensics readiness and investigation methods are needed to be developed.

REFERENCES

[1] Altheide, C., Carvey, H. Digital Forensics with Open Source Tools. 2011, Syngress, 225 Wyman Street, Waltham, MA 02451, USA. ISBN: 978-1-59749-586-8.

[2] Noblett, M.G., Politt, M.M, Presley, L.A.. Recovering and Examining Computer Forensic Evidence. Forensic Science Copmmunications. October 2000 - Volume2 – Number - 4, (2015, 01 26) Available at http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/oct2000/computer.htm/#Top%20of%20article

[3] Holder, E. H., Robinson, L. O., Rose, K. Electronic Crime Scene Investigation: An On-the-Scene Reference for First Responders. U.S. Department of Justice Office of Justice Programs 810 Seventh Street N.W. Washington, DC 20531, 2001.

[4] Francia, G., Clinton, K. Computer forensics laboratory and tools. Journal of Computing Sciences in Colleges. Volume 20, Issue 6, June 2005, 143-150.

[5] Altheide, C., Carvey, H. Computer and Intrusion Forensics. 2003, ARTECH HOUSE, INC., 685 Canton Street Norwood, MA 02062, USA. ISBN 1-58053-369-8.

[6] NIST, Computer Forensic Tool Catalog. (2015, 06 18) Available at http://www.cftt.nist.gov/tool_catalog/index.php

[7] Open Source Digital Forensic. (2015, 07 18) Available at http://www2.opensourceforensics.org/tools

[8] Magnet Forensics. Internet evidence finder (IEF). (2015, 07 28) Available at http://www.magnetforensics.com/

[9] AccessData Group. Forensic Toolkit (FTK). (2015, 07 28) Available at http://accessdata.com/solutions/digital-forensics/forensic-toolkit-ftk

[10] SCG Canada Inc. Covert Forensic Imaging Device (CFID). (2015, 04 08) Available at http://www.scgcanada.com/

[11] BlackBag Technologies. BlackLight. (2015, 03 18) Available at https://www.blackbagtech.com/

[12] PassMark Software Pty Ltd. OSForensics. (2015, 03 18) Available at http://www.osforensics.com/

[13] Steganography Analysis and Research Center (SARC). (2015, 06 02) Available at http://www.sarc-wv.com/

[14] Berla. Infotainment and Vehicle System Forensics (iVe). (2015, 06 02) Available at https://berla.co/ive.html

[15] Micro Systemation. XRY. (2015, 06 02) Available at https://www.msab.com/xry/what-is-xry

[16] Automated Image and Restore. (2015, 07 28) Available at http://sourceforge.net/projects/air-imager/ http://www.linuxlinks.com/article/20110114032833760/AutomatedImageandRestore.html

[17] Sleuth Kit. (2015, 07 12) Available at http://www.sleuthkit.org/sleuthkit/

[18] Autopsy. (2015, 07 12) Available at http://www.sleuthkit.org/autopsy/

[19] DEFT Association. DEFT. (2015, 07 12) Available at http://www.deftlinux.net/

[20] Digital Forensics Framework (DFF). (2015, 07 12) Available at http://www.digital-forensic.org/

[21] IoT-A Internet of Things – Architecture. (2015, 07 12) Available at http://www.iot-a.eu/public/terminology

[22] Oriwoh, E.., Jazani, D., Epiphaniou, G., Sant, P. Internet of Things Forensics: Challenges and approaches. 9th Int. Conference Conference on Collaborative Computing: Networking, Applications and Worksharing, 20-23 Oct. 2013, pp.608-615.

[23] Batalla, J.M., Krawiec, P. Conception of ID layer performance at the network level for Internet of Things. Pers Ubiquit Comput (2014) 18:465–480.

Rapid evaluation of objective function for efficient optimization of the TEAM22 SMES electromagnetic device

Laurentiu Duca, Cornel Popescu
Department of Computers
Polytechnic University of Bucharest
Bucharest, Romania
laurentiu.duca@cs.pub.ro, cornel.popescu@cs.pub.ro

Daniel Ioan
Department of Electrical Engineering
Polytechnic University of Bucharest
Bucharest, Romania
daniel@lmn.pub.ro

Abstract— **Considering that global efficiency of electromagnetic devices optimization essentially depends of the evaluation speed of the objective functions, this paper focus on rapid evaluation of them. It is considered a typical optimization problem: the TEAM22 benchmark problem. The computer evaluation of the objective function for its corresponding electromagnetic device is optimized by using the pthreads library. As a consequence of the parallel optimization, the running time of the optimization procedure decrease with more than 40%.**

Keywords — electromagnetic device, objective function, rapid evaluation, pthreads library

I. INTRODUCTION

Optimization of electromagnetic devices typically consists in search of their geometrical dimensions and/or shape of their component parts as well as the field sources such that some objectives are reached (e.g. cost function to be minimized) and some restrictions are satisfied. Any competitive device should satisfy a series of technical restrictions, meanwhile minimizing its costs. However in many engineering projects, the objective function has a technical meaning and not an economical one. At this moment, there is not a general method able to solve all kinds of optimization problems applied to the large variety of electromagnetic devices. There are many different particular applications and each requires its own specific optimization procedure. The main difficulty in the optimization of electromagnetic devices consists in the evaluation of the objective function, which is a high complexity procedure. It has many independent variables and many points of local extremum. Moreover, each evaluation of the objective function implies numerical solving of an electromagnetic field problem which requires important computation resources, CPU time and memory. In the past years, the electronic technology evolution, based on the downscaling of integrated components generated a sustained increasing of the chip complexity accompanied by the increasing of the working frequency. The transistor shrinking implies a higher power dissipation and the difficulty to extract the generated heat. In this context, the solution to continue the increasing of the computational power is represented by use of the parallel multi-core architectures. They consequently require using of the parallelism also in the algorithm design. Therefore, the parallelism became a basic concept in all advanced High Performance Computing

Systems. Studying techniques and high performance systems is very important due to the fact that many problems with industrial complexity cannot be solved with classic sequential techniques because of very high computing time. Modeling, analysis and optimization of complex electromagnetic devices are part of this category of problems which requires high performance techniques and systems in their solving [1].

II. THE TEAM22 BENCHMARK PROBLEM

The TEAM22 benchmark problem aims the optimization of an electromagnetic device [2].

Two coaxial coils operating under superconducting conditions (Fig. 1) and carrying current with opposite direction offer the opportunity to store a significant amount of energy in their magnetic fields while keeping the stray field within certain limits. An optimal design of the system should couple the desired value of energy (180MJ) to be stored with a minimal stray field.
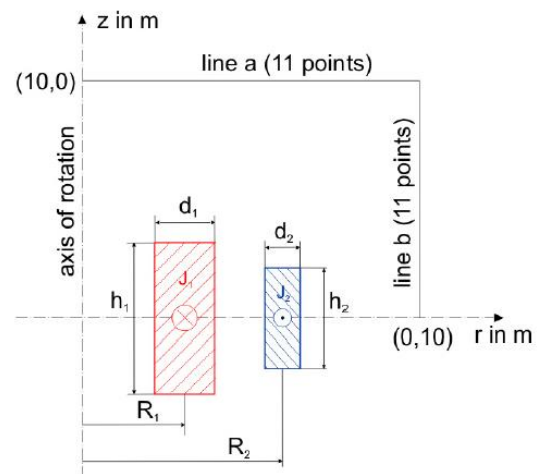

Figure 1.   TEAM22 problem configuration [2]

The two objectives are mapped into a single objective function with eight parameters: heights (h1, h2), radii (R1, R2), thickness (d1, d2), and current densities (J1, J2) of both coils. The restrictions of these parameters are presented in table 1.

TABLE I.   PARAMETERS RESTRICTIONS OF THE TEAM22 PROBLEM

Parameters restrictions of the TEAM22 problem								
	R1 [m]	R2 [m]	h1/2 [m]	h2/2 [m]	d1 [m]	d2 [m]	J1 [MA/m^2]	J2 [MA/m^2]
min	1.0	1.8	0.1	0.1	0.1	0.1	10.0	-30.0
max	4.0	5.0	1.8	1.8	0.8	0.8	30.0	-10.0

Besides these restrictions, the optimization problem has to take into consideration that the superconducting material should not violate the quench condition and the solenoids should not overlap each other. The quench condition consists in a relation between the maximum value of the magnetic induction in the coils and the current density (see Fig. 2).
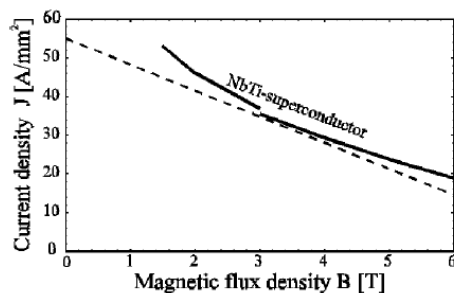

Figure 2.   Critical curve of an industrial superconductor [2]

The optimization problem is reformulated as a problem with six parameters, because for a given geometry and a stored energy, the values of the current densities can be computed by deterministic quadratic optimization [3].

III. STATE OF THE ART

Although in literature are present various algorithms [4],[5] that optimize the TEAM22 SMES device by choosing the best parameters after repeated evaluation of the objective function, there are no results in trying to improve the objective function code by using parallel techniques, in order to run faster. In [6], the TEAM22 objective function is approximated by a fuzzy neural network, but there are considered only three parameters, the other being fixed. In [7] the approximation is made by a PLP neural network, but the results obtained with the optimization GA on the PLP are one order of magnitude worse than the results in [5] on the normal objective function.

IV. OPTIMIZATION OF THE TEAM22 OBJECTIVE FUNCTION

A. Objective function statistics

One of the serial implementations in the C language of the objective function was made in [8]. The evaluation method of the objective function is based on the Biot-Savart-Laplace formula in which the elliptic integrals are computed by using numerical integration and the King algorithm. In order to optimize the serial implementation of the objective function we analyzed the source code by using the gprof tool which showed the profile of the objective function. The gprof simplified flat profile in the point (RSp1=1.366, DSp1=0.598, HSp1=1.167•2,

RSp2=1.825, DSp2=0.291, HSp2=1.551•2) is shown in table 2. The gprof tool also offers the graph who shows which functions call or are called by any function and how many times are the calls done. A simplified version of the call graph listed by gprof is presented in Fig. 3 (in brackets is shown the number of function calls).

TABLE II.   GPROF SIMPLIFIED FLAT PROFILE

gprof simplified flat profile	
Function name	Calls number
King()	27892
InterLin()	189706
ASpira()	116680
ABaraDepRec()	23336
BSpira()	3520
BzSpira()	2650
ABaraDep()	389
BzBaraRec()	174
BBaraDep()	48
BzBaraApColt2()	4
ProperEnergy()	2
MutualEnergy()	1
GeomOF()	1
team22ObjectiveFunction()	1

The total run time of the objective function on a system with Intel Xeon X5650 processor running at 2.67 GHz is between [16.00, 72.00] ms. The gprof tool has the sampling time step of 10 ms, which fails in creating a time profile. The time profile for each function is made by using, inside the code, system calls that provide time measurement.

---

team22ObjectiveFunction()→evaluate(1)→GeomOF(1)→MutualEnergy(1), ProperEnergy(2), BBaraDep(48), BzBaraApColt2(4)

MutualEnergy()→ABaraDep(153)

ProperEnergy()→ABaraDep(236)

ABaraDep()→ABaraDepRec(23336)→ASpira(116680)→InterLin(181784), King(25788)

BzBaraApColt2()→BzBaraRec(174)→BzSpira(2650)→InterLin(1360), King(1970)

BBaraDep()→BSpira(3520)→InterLin(6562), King(132)

Figure 3.   The simplified gprof call graph

B. Objective function optimization

The most interesting functions to be parallelized are MutualEnergy(), ProperEnergy(), BBaraDep(), ABaraDep(), BzBaraApColt2(). From these functions, specific code sequences that can be parallelized were identified in MutualEnergy(), ProperEnergy(), BzBaraApColt2() and ABaraDep(). In this functions, the sequences which contained for-loops where were calls of children functions, were replaced with parallel calls of those children functions by using the pthreads library, which implements the parallel programming

model where the threads share the same address space within a single process. The parallel version of the TEAM22 objective function is available at [5] and is described below.

The runtimes obtained for the parallelized functions in the point (RSp1= 1.366, DSp1=0.598, HSp1=1.167•2, RSp2= 1.825, DSp2=0.291, HSp2=1.551•2) are presented in table 3. Those times were obtained by running the objective function after compiling with gcc with the optimization option O3. The number of function calls for which were measured the time is the one from the function GeomOf(), because starting from this function the program branches.

TABLE III.   RUNNING TIMES FOR PARALLELIZED FUNCTIONS IN THE POINT

Running times for parallelized functions in the point (RSp1= 1.366, DSp1=0.598, HSp1=1.167•2, RSp2= 1.825, DSp2=0.291, HSp2=1.551•2)			
Function	Calls number	Time for original function (ms)	Time for parallelized function (ms)
MutualEnergy()	1	7.651	6.103
ProperEnergy()	2	17.434	9.393
ABaraDep()	1	0.092	0.558
BBaraDep()	48	0.583	-
BzBaraApColt2()	4	0.450	-

The parallelization of ABaraDep() function increased the running time of this function because its children functions are short functions called many times (see table 2 and Fig. 3) and the time necessary to create and join the parallel threads is not negligible. Because of the same reason, the functions BBaraDep() and BzBaraApColt2() were kept as original and not parallelized.

```
            for(i=0;i<=n;i++) {
   r0=rsp-dsp/2+i*dsp/n;
   for(j=0;j<=m;j++) {
            z0=zsp+j*hsp/(2*m);
            d[0]=r0-(rsp-dsp/2); d[1]=(rsp+dsp/2)-r0;
            h[0]=z0-(zsp-hsp/2); h[1]=(zsp+hsp/2)-z0;
            r0A=0;
            if(i>0) {
                    r0A+=AbaraDep(r0,z0,r0-d[0]/2,z0-h[0]/2,
d[0],h[0],DstEtalon2,interp,EI, KI, Step, From);
                    if(j<m)
                            r0A+=AbaraDep(r0,z0,r0-d[0]/2,
z0+h[1]/2,d[0],h[1],DstEtalon2,interp,EI, KI, Step, From);
                    }
            if(i<n) {
                    r0A+=AbaraDep(r0,z0,r0+d[1]/2,z0-h[0]/2,
d[1],h[0],DstEtalon2,interp,EI, KI, Step, From);
                    if(j<m)
                            r0A+=AbaraDep(r0,z0,r0+d[1]/2,z0+h[1]/2,d[1],
h[1],DstEtalon2,KI,Step, From);
                    }
            EP+=ArieRel(i,j,n,m)*r0A;
    }
}
```
Figure 4.   The code from ProperEnergy() that was subject to parallelization

Because the parallelization of MutualEnergy() brought approximately the same time as the original version, in the final version was kept the original code of this function. The parallelization of the ABaraDep() loop function calls from ProperEnergy() brought a significant less time than the original serial implementaion of ProperEnergy(). The code from ProperEnergy() that was subject to parallelization is listed in Fig. 4.   Each call from ProperEnergy() to ABaraDep() which

was done from two for loops in the original code was parallelized by using pthreads, and each thread made a parallel call of the ABaraDep() function. The call to ABaraDep() was made indirectly by using callABaraDep() which was called from pthread_create(). The parameters were transfered by using the global vector vabara[] indexed by the thread id. The code is shown in Fig. 5.

```
pthread_t vthreads[PTHREAD_THREADS_N];
int vtid[PTHREAD_THREADS_N];
typedef struct {
double r0;double z0;double rsp;double zsp;double dsp;double hsp;
double dstEtalon;BOOL interp;
double *EI; double *KI; double Step; double From;
} ABaraDepType;
ABaraDepType vabara[PTHREAD_THREADS_N];
double abararet[PTHREAD_THREADS_N];

void* callABaraDep(void *params)
{
   int tid = *((int*)params);
   abararet[tid]=ABaraDep(vabara[tid].r0,vabara[tid].z0,vabara[tid].rsp,
      vabara[tid].zsp,vabara[tid].dsp,vabara[tid].hsp,
      vabara[tid].dstEtalon,vabara[tid].interp,vabara[tid].EI,
      vabara[tid].KI, vabara[tid].Step, vabara[tid].From);
   return NULL;
}
double  ProperEnergy(double  rsp,double  zsp,double  dsp,double  hsp,  double
DstEtalon,int interp, double *EI, double *KI, double Step, double From)
{
  double EP,r0A,r0,z0,Arie,d[2],h[2]; int i,j,n,m;
  ///////////////////////////////////
  /**/ double DstEtalon2=DstEtalon/2; /**/
  ///////////////////////////////////
  n=(int)ceil(dsp/DstEtalon); m=(int)ceil(hsp/(2*DstEtalon));

  EP=0; Arie=(dsp*hsp)/(n*m);

  // pthread_create
  int tid = 0;
  for(i=0;i<=n;i++) {
     r0=rsp-dsp/2+i*dsp/n;
     for(j=0;j<=m;j++) {
        z0=zsp+j*hsp/(2*m);
        d[0]=r0-(rsp-dsp/2); d[1]=(rsp+dsp/2)-r0;
        h[0]=z0-(zsp-hsp/2); h[1]=(zsp+hsp/2)-z0;
        //r0A=0;
        if(i>0) {
                tid++;
                vtid[tid] = tid;
                vabara[tid].r0=r0; vabara[tid].z0=z0; vabara[tid].rsp=r0-d[0]/2;
                vabara[tid].zsp=z0-h[0]/2; vabara[tid].dsp=d[0];
                vabara[tid].hsp=h[0];vabara[tid].dstEtalon=DstEtalon2;
                vabara[tid].interp=interp; vabara[tid].EI=EI; vabara[tid].KI=KI;
                vabara[tid].Step=Step; vabara[tid].From=From;
                if(pthread_create(&vthreads[tid], NULL, &callABaraDep,
&vtid[tid]))  {
                        perror("pthread_create");
                        exit(1);
                }
...
```
Figure 5.   Parallelizing ABaraDep()calls from ProperEnergy()

The return values of ABaraDep() which were summed in the r0A variable in the original code, in the parallel version were saved in a global vector and, after thread joining, the values from the global vector were summed in r0A. Thread joining is shown in Fig. 6.

```
...
  // pthread_join
  tid = 0;
  for(i=0;i<=n;i++) {
    r0=rsp-dsp/2+i*dsp/n;
    for(j=0;j<=m;j++) {
           z0=zsp+j*hsp/(2*m);
           d[0]=r0-(rsp-dsp/2); d[1]=(rsp+dsp/2)-r0;
           h[0]=z0-(zsp-hsp/2); h[1]=(zsp+hsp/2)-z0;
           r0A=0;
           if(i>0) {
                        tid++;
                        // wait for completion
                        if (pthread_join(vthreads[tid], NULL))
                                    perror("pthread_join");
                        r0A+=abararet[tid];
                        //r0A+=ABaraDep(r0,z0,r0-d[0]/2,z0-
h[0]/2,d[0],h[0],DstEtalon2,interp,EI, KI, Step, From);
                   if(j<m) {
                        tid++;
                        // wait for completion
                        if (pthread_join(vthreads[tid], NULL))
                                    perror("pthread_join");
                        r0A+=abararet[tid];
                        //r0A+=ABaraDep(r0,z0,r0-
d[0]/2,z0+h[1]/2,d[0],h[1],DstEtalon2,interp,EI, KI, Step, From);
                   }
           }
...
```

Figure 6.   Thread joining in parallelized ProperEnergy()

V. RESULTS

For performance testing of the parallelized version against the original version it was considered the implementation of the SPSO (Standard Particle Swarm Optimization) algorithm in the C language for the TEAM22 objective function (see [5]). SPSO is an algorithm described in [9] and has the scope to find the minimum point of the objective function.
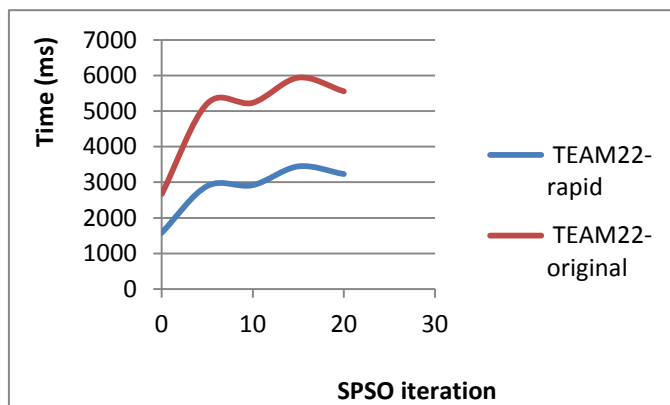

Figure 7.The running time of the SPSO algorithm iterations

SPSO consists in a population of candidate solutions (which are points in the search space) which evolves (changes) at each iteration of the optimization SPSO algorithm. For testing, it was used a population of 128 particles. At each iteration, each particle from the population it was evaluated by calling the objective function in the respective point. The total times of running the objective function for all of the 128 particles of the every iteration were saved and a comparison was made between the original and the rapid (parallelized)

objective function. The results are shown in Fig. 7. The parallelized objective function implies running times for the SPSO iterations with at least 40% less than the SPSO on the original TEAM22 objective function.

VI. CONCLUSION

In this paper were used high performance techniques for rapid evaluation of objective functions for efficient optimization of electromagnetic devices. Using the parallel pthreads library, it was optimized the objective function of the electromagnetic device for a famous problem in the electromagnetic field, the TEAM22 benchmark problem. For this problem the running time of the objective function was reduced with more than 40% of the running time of the original function.

The electronic technology evolution, based on the downscaling of integrated components, generated a sustained increasing of the chip complexity and the difficulty to extract the generated heat. The solution to increase the computational power is represented by use of the parallel multi-core architectures and the parallelism in the algorithm design.

REFERENCES

[1] A. Muller, A. Moldoveanu, V. Asavei and D. Dascalu, A 3 D Smith Chart For Active and Passive Microwave Circuits and Visual Complex Analysis, Scientific Bulletin of UPB, vol. 74, 2012, pp. 181-188.

[2] P.G. Alotto, U. Baumgartner, F. Freschi, M. Jaindl, A. Kostinger, Ch. Magele, W. Renhart and M. Repetto, SMES Optimization Benchmark: TEAM Workshop Problem 22. Available at: http://www.compumag.org/jsite/images/stories/TEAM/problem22.pdf

[3] D. Ioan, G. Ciuprina and A. Szigeti, Embedded stochastic-deterministic optimization method with accuracy control, IEEE Transactions on Magnetics, vol. 35, 1999, pp. 1702-1705.

[4] R. Barbulescu, A.-S. Lup, G. Ciuprina and D. Ioan, Intelligent Particle Swarm Optimization of Superconducting Magnetic Energy Storage devices, International Symposium on Fundamentals of Electrical Engineering, 2014, pp. 1-4.

[5] A. Duca, L. Duca, G. Ciuprina, E. Yilmaz and T. Altinoz, PSO algorithms and GPGPU technique for electromagnetic problems, International Workshop on Optimization and Inverse Problems in Electromagnetism (OIPE), 2014. Digest available at: http://www.oipe2014.nl/uploads/File/BookOfDigests.pdf

[6] K. Rashid, J.A. Ramirez and E.M. Freeman, Hybrid optimization in electromagnetics using sensitivity information from a neuro-fuzzy model, IEEE Transactions on Magnetics, 2000, Vol.36, pp. 1061-1065.

[7] D.A.G. Vieira, W.M. Caminhas and J.A. Vasconcelos, A hybrid approach combining genetic algorithm and sensitivity information extracted from a parallel layer perceptron, IEEE Transactions in Magnetics, 2005, Vol. 41 , pp.1740-1743.

[8] G. Ciuprina, Electromagnetic field study in non-linear medium, Phd Thesis, University Politehnica of Bucharest, 1999.

[9] M. Zambrano-Bigiarini, M. Clerc and R. Rojas, Standard Particle Swarm Optimisation 2011 at CEC-2013: A baseline for future PSO improvements, IEEE Congress on Evolutionary Computation, 2013, pp. 2337-2344.

Type of rating potential customer benefiting from the promotion company

Elmira Kushta

Department of Mathematics, Faculty of Technical Sciences
University of Vlora
Vlora, Albania
kushtamira@gmail.com

Markela Muça

Department of Applied Mathematics, Faculty of Natural Sciences
University of Tirana
Tirana, Albania
markela.muca@fshn.edu.al

Fatmir Memaj

Department of Statistics, Faculty of Economy
University of Tirana
Tirana, Albania
fmema@yahoo.com

Abstract—**Logistic regression analysis is a known statistical technique which is used to analysis the role of qualitative variables (variable outcomes). In this article, we show an implementation of this method (LR) for purposes of evaluating the probability that a customer will use the client card. We care about the costs and impact as if a customer has the credit card in predicting the probability of customer card use. It indicated that credit card customers are 6 times more favorable to the customer than not using the card, regardless of costs that do. Also, it is shown that a customer is more likely to use customer card when costs rise despite has or no credit card.**

Keywords: logistic regression, odds ratio, chances, binary data analysis

I. Introduction

Regression methods have been an integral part of any analysis of data related to the description of the relationship between a response variable and one or more explanatory variables. Logistic regression model is a form of regression which is used when the dependent variable (Y) is dichotomous (binary) and qualitative values i.e., 2 levels e.g., with or without customer card, etc., and the independent variable (X s) can be numerical, categorical or mixed.

This technique is applied in different research area, particularly in the medicine (e.g., Antonogeorgos G. et al., 2009), socio-logical sciences (e.g., HL Chuang, 1997) and in the education (state graduation results) (e.g., Sadri Alija, et al., 2011, Muca M., et al., 2013).

The purpose of an analysis that uses this method is the same as that of any technical building models in statistics, so to find the best and more reasonable model to describe the connection between a variable result and a set of variables independent. In logistic regression after adjusting the model, the emphasis is to evaluate and interpret is o the coefficients and their values, where conformity assessment methods are more of a technical nature.

The estimated coefficients for the independent variable present growth (i.e. the boundary change) a function of the dependent variable per unit change in the independent variable. Thus, interpretation involves two issues:

1- Determining the functional relationship between the dependent variable and independent.

2- Determining the unit change in the independent variable.

II. The Logistic Regression Equation

The logistic regression model has the same form of the regression model, which is used when the dependent variable Y is qualitative (binary) with two values or categorical with more than two values and the independent variables can be quantitative, qualitative or mix. In multiple regression analysis, the mean or the expected value of y, E(Y), is calculated from equation (1)

$$E(Y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \tag{1}$$

While, in the logistic regression is shown (Chao-Ying Joanne Peng and Tak-Shing Harry So, 2002) the connection between E(Y) and $x_1, x_2, \cdots, x_k$ that is described by nonlinear equation (2)

$$E(y) = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)} \tag{2}$$

If the dependent variable y is coded as 0 and 1, the value of E(y) in equation (2) provides the probability that y = 1 given a particular set of values for the independent variables $x_1, x_2, \cdots, x_k$ (see appendix). Because of the interpretation of E(y) as a probability p, the logistic regression equation is given in equation (3).

$$E(Y)=P(Y=1/\underset{\sim}{X}=\underset{\sim}{x})=p(y)=\frac{\exp(\beta_0+\beta_1 x_1+\cdots+\beta_k x_k)}{1+\exp(\beta_0+\beta_1 x_1+\cdots+\beta_k x_k)} \quad (3)$$

III.   DATA. LINEAR REGRESSION ANALYSIS LOGISTICS LR

It is believed that the total costs in a supermarket and the fact that a customer has a credit card or not, are two important variables in predicting if a client who has client card will use it. Therefore, in this study we used a sample which consists of information for n = 100 clients (50 have credit card and a credit card 50 have not). The set of data consists of a dependent variables and forecast variables (predictor variables). For every customer the total costs marked 1000 ALL and {0,1} are listed whether or not the customer has credit card (see Table 1). Logistic regression model will be used to assess what kind of potential customer is to benefit from the promotion of the company.

TABLE I.   DESCRIPTION OF FEATURES FOR THE DATABASE COMMUNITY

Variables	Description
Use client card	{0 → don't use and 1 → use}
Total costs	Are in " thousand ALL (Leke)"
Use credit card	{0 → don't use and 1 → use}

Figure 1, shows that customers can be grouped as follows:

- Customers who do not use credit cards (78%, 39 from 50) spend less (to 10 thousand) than those who buy with credit card.

- Consumers, middle class, who spend 10-20 thousand cash purchase 16%, 8 from 50 and 78% purchase credit card, 39 from 50.

- Customers who spend from 20-28 thousand cash purchase 6%, 3 from 50 and 14% purchase credit card, 7 from 50.
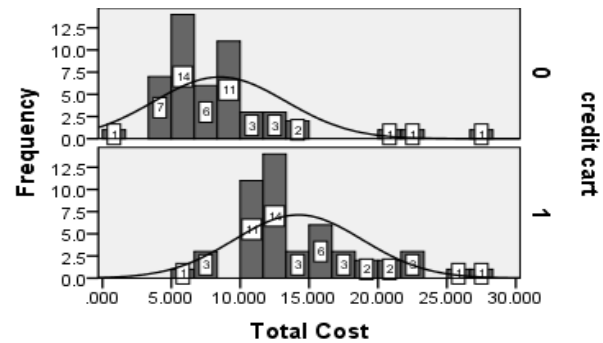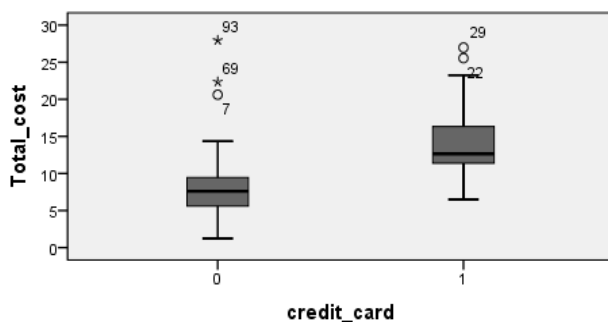
Figure 1.   Graphical presentation according to the type of payment.

IV.   LOGISTIC REGRESSION ANALYSIS

The linear logistic regression model was adapted into the set of the data presented in Table 1. In Table 2 are represented the results of RL method performed in SPSS.

$$P(Y=1/\underset{\sim}{X}=\underset{\sim}{x})=\frac{\exp(-1.965+0.035\text{Total\_Cost}+1.794\text{Credit\_Cart})}{1+\exp(-1.965+0.035\text{Total\_Cost}+1.794\text{Credit\_Cart})}$$
$$(4)$$

According to the model, the natural logarithm of odds that a client use client card is positively associated with both variables (cost and Credit card) with p = .467 and p = .001 respectively. In other words, more money a client has spent, the greater will be the chance of being a user (that a client used client cart). Exp(B) for the total cost (expenditure) is equal to 1.036, which means when the total costs are added to a unit (1000 ALL), the client is 1.036 times more likely to use the card customer than not, or if the difference in domestic currency (ALL) is 20 thousand then a client is $2.03 = 1.036^{20}$ times more likely to use the client card. Customers who use credit cards are more likely to use client card compares to they that not have credit card because credit card customers are codified 1 while customers who have no credit card are codified with 0. Actually, the odds ratio for clients with a credit card (odds) who use the customers card is 6.015 ($=e^{1.794}$ Table 2) times greater than the odds ratio for customers who have no credit card. Also based on the 95% confidence interval for the expenditure we see that it contains the value 1 and in such cases it is said that change in expenditures is statistically insignificant, and for other intervals that do not contain the value 1 the opposite is true.

TABLE II.   RESULTS OF THE LOGISTIC REGRESSION ANALYSIS

Variables in the Equation									
		B	S.E	Wald	Df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1[a]	Total_Cost	.035	.048	.528	1	.467	1.036	.942	1.138
	Credit_Card	1.794	.538	11.104	1	.001	6.015	2.094	17.279
	Constant	-1.965	.583	11.381	1	.001	.140		

a. Variable(s) entered on step 1: Total_Cost, Credit_Cost.

We can now use equation (4) to estimate the probability of using the customer card for a particular type of customer. Thus, for this group of customers, the probabilities of using or not using the customer card when his total expenses are equality as some numerical indicators (minimum, maximum, mode, and the average costs) are shown in Table 3. From the results presented in Table 3, it appears that the probability of using the customer card (when his expenses are: 1235 ALL, 11366 ALL, 12640 ALL and 27957 ALL) is much higher for customers with credit card.

Supermarket is interested for those customers who have customer card and not a credit card, but the bank is interested more to those customers who have a credit card because for every purchase there is 10% off.

The analysis showed that the most reliable customers are those customers that have customer card and credit card.

TABLE III.　ASSESSING THE PROBABILITY ACCORDING TO SOME NUMERICAL CHARACTERISTICS

		Minimum =1.235	Average costs =11.366	Mode =12.640	Maximum =27.957
Credit card	1	46.81%	55.65%	56.74%	69.16%
	0	12.77%	17.26%	17.91%	27.16%

V.　BINARY LOGISTIC REGRESSION INTERPRETATION (BINARY LOGISTIC REGRESSION)

With logistic regression model, it is difficult to interpret the relation between the independent variables and the probability of occurrence of success, y=1, directly because the logistic regression equation is nonlinear. Instead of model the probability p directly to a linear model, we first consider the odds ratio as follows:

$$odds = \frac{p(A)}{1-p(A)} = \frac{prob \text{ of } 1\text{-s}}{prob \text{ of } 0\text{-s}} = \frac{P(Y=1/X_1=x_1, X_2=x_2)}{1-P(Y=1/X_1=x_1, X_2=x_2)} \quad (5)$$

The odds ratio measures the impact on the odds of a one unit increase in only one of the independent variables. The odds ratio is the odds that {Y=1}, a customer uses the customer card, given that one of the independent variables such as expenses has been increased by one unit ($odds_1$) divided by the odds that {Y =1}, a customer uses the customer card given no change in the values for the two independent variables ($odds_0$). So the odds ratio for customers with and without credit card is respectively as follows:

For credit card customers we have:

$$odds_1 = \frac{P(Y=1/X_1=x_1, X_2=1)}{1-P(Y=1/X_1=x_1, X_2=1)} \quad (6)$$

For customers without credit card we have:

$$odds_0 = \frac{P(Y=1/X_1=x_1, X_2=0)}{1-P(Y=1/X_1=x_1, X_2=0)} \quad (7)$$

The ratio of the two conditional probabilities is calculated as follows:

$$odds = \frac{odds_1}{odds_0} \quad (8)$$

In Table 4 are shown the ratios of two conditional probability (equation 8) for customers using the customer card when the total of her expenses as some numerical indicator (minimum, maximum, mode, and the average cost).

TABLE IV.　ASSESSMENT REPORT TWO CONDITIONAL PROBABILITY BY SOME NUMERICAL CHARACTERISTICS

		Min= 1.235	Average costs =11.366	Mode= 12.640	Max= 27.957
Credit card	$Odds_1$	0.88	1.25	1.31	2.24
	$Odds_0$	0.15	0.21	0.22	0.37
odds		6.01	6.01	6.01	6.01

VI.　CONCLUSION

The analysis showed that customers who use credit cards are more favorable to use customers card compared to those who don't have credit card and the chances increase when total cost rise. Also, it is seen that a credit card customer is 6 times more favorable to use the customer card compared to those who don't have customer card.

REFERENCES

[1] Bewick V, Cheek L., Ball J: Statistics review 8: Qualitative data–tests of association. Crit Care 2003, 8:46-53

[2] Chao-Ying Joanne Peng and Tak-Shing Harry So, (2002). Logistic regression Analysis and Reporting: A Primer. Understanding Statistics, l(1), 31-70. Lawrence ErlbaumAssociates, Inc

[3] D.w. Hosmer, S. Lemeshow (1989).Applied Logistic Regression.willey, NewYork, .

[4] George Antonogeorgos, Demosthenos B. Panagiotakos, Kostas N. Priftis and Anastasia Tzonou, (2009).Logistic Regression and Linear Discriminat Analysis in Evaluating Factors Associated with Asthma Prevalence among 10- to 12-Years –Old Children: Divergence and Similaritity of the Two Statistical Methods. Hindawi Publishing Corporation, International Journal of Pediatrics.

[5] Muça M., PUKA Ll., BANI K, SHAKAJ F. (2013) "Logistic Regression Analysis: A Model to Predict the Entrance Probability in Higher Education". "1st International western Balkans Conference of Mathematical Sciences – IWBCMS-2013", in Elbasan/ALBANIA PROCEEDINGS

[6] Sadri Alija, Lazim Kamberi and Llukan Puka, (2011). Logistic regressions and an application in teaching practice assessment. In AKTET, Journal of Institute Alb-Shkenca. Volume IV, 8.

Appendix 1

The multiple logistic regression model has the form (1)

$$Ln(\frac{p}{1-p}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \qquad (1)$$

Therefore
$$P = prob(Y = 1 / X_1 = x_1,..., X_k = x_k)$$
$$= \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)},$$

$\beta_0, \beta_1,...\beta_k$ they are the regression coefficients which are estimated by the maximum likelihood method, X-s are a set of predictors, p is the probability of occurrence of the eventA={Y=1/X=x}.

$Y = X\beta + \varepsilon$ it is a linear model with $Y \sim B(p)$. If $E(\varepsilon) = 0$ then $E(Y) = 1*p + 0*(1-p) = p$ (in general) and $E(y_i) = x_i'\beta = p_i$ (in particular). From the last equation shows that the mean of the response to the linear function $Y = X\beta$ is equal to the probability of the event A={Y=1/X=x}.

Some problems encountered in linear model:

1. If the response is binary, then the error terms can take on two values, namely,

$$\varepsilon_i = \begin{cases} 1 - x_i'\beta, & y_i = 1 \\ -x_i'\beta, & y_i = 0 \end{cases}$$

Because the errors is dicreate, normaly assuption is violated.

2. The error variance is not constant and it can be seen that it is a function of probabilities, because

$$D(y_i) = E(y_i - E(y_i))^2 = (1-p_i)^2 * P(y_i = 1) + (0-p_i)^2 P(y_i = 0)$$
$$D(y_i) = (1-p_i)^2 * p_i + p_i^2(1-p_i)$$
$$D(y_i) = (1-p_i)*p_i = E(y_i)(1-E(y_i))$$

where $E(y_i) = x_i'\beta = p_i$

Therefore the assuption of homoscadasticity does not hold.

From the last equation shows that the variance response (that It is equal to with the error variance because $\varepsilon_i = y_i - p_i$, where p_i an constant) is a function of waiting Mathematical and $0 \le E(y_i) = p_i \le 1$. Last inequality It causes problems if elected linear model.

The interpretation of βs is rendered using either the odds ratio (for categorical predictors) or the delta-*p* (for continuous predictors). The null hypothesis states that all βs equal zero. A rejection of this null hypothesis implies that at least one of the parameters does not equal zero in the popullation, which means that the logistic regression model predicts the probability of the answer better than the mean of the answer Y (see Item 2 to some problems encountered in linear model) that determined by draw (2).

$$P(Y = 1 / \underset{\sim}{X} = \underset{\sim}{x}) = p(y) = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)} \quad (2)$$

Active natural disasters policy – the case study of Lublin Voivodeship in Poland

Zbigniew Piepiora

Wrocław University of Environmental and Life Sciences
Wrocław, Poland
zbigniew.piepiora@up.wroc.pl

Paweł Piepiora

University School of Physical Education in Wrocław
Wrocław, Poland
pawelpiepiora@gmail.com

Abstract—**In the article, there is verified counteraction of natural disasters' effects in Lublin Voivodeship (NUTS 2). In the examined area, preventing the negative consequences of the extreme phenomena is financed from the budget of the voivodeship and from budgets of Lublin Voivodeship's self-governments. Expenses are assisted by the measures from e.g. the central budget, foreign states, ecological funds, credits and loans. It is worth noting that the low part of the expenditures for removal the effects of natural disasters in the expenses for public safety, fire and health protection in the years 2008-2014.**

Keywords - natural disaster, policy, Lublin Voivodeship, counteraction, funding, effects

I.   INTRODUCTION

In July 1931, the tornado ravaged area of Lublin, killing 6 people. In the spring of 2009, there was a landslide in Godziszów Drugi in Janowski County. During the flood in 2010 under water was, among others, Wilkowo municipality. In 2015, Lublin Voivodeship was affected by the drought.

Droughts, landslides, tornadoes and floods are examples of natural disasters. These phenomena can be defined, according to the Natural Disaster Act, as an events connected with impacting of nature's forces. Besides mentioned examples, catastrophes are also: atmospheric discharges, storms, seismic activity, intensive rainfalls, extreme temperatures, fires, ice's phenomena on rivers, lakes, water reservoirs and the sea, infestations of insects and diseases of plants, animals and people[18, art. 3].

The natural disaster can be started by natural factors: biological (e.g. the disease), geological (e.g. the seismic activity), hydrometheorological (e.g. the storm). It is often the sudden event with tragic effects which causes damages, suffering of people and also changes connected with the affected area[1, 2].

Active natural disasters' policy can be defined as the conscious activity of the government and council authorities, as well as the national and international institutions specialized to minimize the damages caused by natural disasters[7].

The aim of this article is to verify the active natural disaster's policy. It is verified at the regional level of Lublin Voivodeship (NUTS 2 – Nomenclature of Units for Territorial Statistics level 2). The article is part of the research project which concerns 16 Polish voivodeships (provinces).

II.   THE CASE STUDY IN EXAMINED AREA – INSTITUTIONAL AND FINANCIAL ANSPECTS

Lublin Voivodeship is located in the central-eastern Poland. From the East it borders with Belarus and Ukraine[8]. It neighbors with 4 voivodeships: Podlaskie, Mazowieckie, Świętokrzyskie and Podkarpackie. It counts 2.15 million inhabitants. It covers an area of 17 970 km^2. The population density is 86 persons per 1 km^2. The most important city of the examined area and the capital is Lublin[16]. Other important cities are Chełm, Zamość, Biała Podlaska, Puławy and Świdnik[11, 12].

TABLE I.   EXPENDITURES FOR PERMANENT ASSETS SERVING THE WATER MANAGEMENT IN LUBLIN VOIVODESHIP IN THE YEARS 1998-2013 – DIRECTIONS OF INVESTING (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

1.	2.	3.	4.	5.	6.	7.	8.
1998	37964	28776	465	3875	1174	3675	0
1999	25841	19005	2558	1724	632	1923	0
2000	24484	16325	1882	1800	844	3633	0
2001	19970	12458	1650	963	488	4407	4
2002	22585	10993	2117	1060	3239	5175	0
2003	37300	16154	2180	1314	7182	8555	1915
2004	34766	20515	2432	1731	1152	6774	2161
2005	19533	12039	1878	2368	146	3023	80
2006	25251	17159	1489	2967	1071	2396	170
2007	26180	13144	1535	3668	1532	5023	1279
2008	15301	11359	2165	813	778	185	0
2009	18369	16373	1441	47	342	167	0
2010	42693	25867	2747	1784	870	11424	0
2011	43225	23225	2830	2652	437	14082	0
2012	31616	20779	2319	1025	2569	4925	0
2013	32450	18412	3523	1010	3681	5825	0
total – all years	457527	282581	33210	28803	26135	81190	5609

1. years, 2. total, 3. intakes and deliveries of water, 4. building and modernization of water conditioning stations, 4. water reservoirs and stages, 5. regulation and development rivers and streams, 6. levees, 7. pump stations on breaking downs and depression areas
Source: Own study on the basis of: [4, 5].

The examined area includes both the lowlands (lakes, Polesie) and upland (Roztocze wish Puszcza Solska and Lasy Janowskie), cut by river valleys (Bug and Vistula), which, as a few have not been subjected to any regulatory activities[13]. Natural disasters that might occur within the examined province are: floods, droughts, landslides and storms[10]. Storms can affect the whole examined area. It is estimated that the areas at risk of flooding include about 60 thousand ha (2.4% of the area of the province). In 2010, areas of potential flood conditions (protected by dikes) were 264 km^2[9]. The occurrence of landslides are threatened edge zones within

Wyżyna Lubelska and Roztocze, as well as the slopes of river valleys at risk of undercutting by flood waters[14]. Drought is strongly threatened north-eastern part of the examined province. Drought hampers water supply for agriculture and industry, and part of the natural areas[15]. Irrigation system (a system Wieprz-Krzna Canal) built to counteract this phenomenon, in the present conditions does not meet expectations[6].

TABLE II.   EFFECTS OF INVESTING IN THE WATER MANAGEMENT IN THE YEARS 2001-2013 IN LUBLIN VOIVODESHIP

1.	2.	3.	4.	5.	6.	7.	8.	9.
units of measure	'000 m3	'000 m3	km	unit	'000 m3	km	unit	km
2001	1960	1638	448	6	1137400	5	0	8
2002	9994	1834	544	3	62055	3	0	3
2003	4818	67	455	2	24275	7	0	37
2004	11017	5027	834	2	209600	17	0	3
2005	8090	505	218	3	4402	17	0	0
2006	14721	189	486	0	0	0	0	4
2007	4297	130	281	2	1041600	8	0	5
2008	6129	2148	291	1	19500000	3	0	6
2009	7608	1083	323	0	0	0	0	3
2010	6898	1866	590	1	29900	3	0	0
2011	7008	1166	491	2	176900	0	0	0
2012	772	4320	424	1	63230	6	0	0
2013	8196	1223	348	0	0	5	0	1
total – all years	91508	21196	5732	23	22249362	74	0	70

1. years, 2. intakes – efficiency per 24 hours, 3. water conditioning per 24 hours, 4. water supply system, 5. water reservoirs – total number, 6. water reservoirs – total capacity, 7. levees, 8. pump stations on breaking downs and depression areas, 9. regulation and development rivers and streams
Source: Own study on the basis of: [4, 5].

TABLE III.   EXPENDITURES FOR PERMANENT ASSETS SERVING THE WATER MANAGEMENT IN LUBLIN VOIVODESHIP – DIRECTIONS FINANCING SOURCES IN THE YEARS 2002-2013 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
2002	22585	8415	7345	1599	11	171	0	1023	914	3107
2003	37300	10050	2954	955	0	104	16711	2267	2044	2215
2004	34766	12222	2442	1223	11	72	13557	1735	2831	672
2005	19533	8471	1868	829	12	272	3810	2018	1042	1211
2006	25251	12495	1093	572	0	388	3911	2203	2482	2107
2007	26180	11297	2274	549	0	85	6631	1621	2642	1081
2008	15301	9449	595	478	36	70	665	863	2967	178
2009	18369	11785	19	309	0	90	604	1597	2818	1147
2010	42693	11025	9296	322	0	302	13386	303	7448	611
2011	43225	14273	8589	344	0	80	13464	1535	3889	1051
2012	31616	12513	3132	427	0	326	6796	4112	2471	1840
2013	32450	9958	4649	398	25	188	8211	4985	1592	2443
total – all years	349268	131954	44255	8004	96	2147	87747	24262	33140	17662

1. years, 2. total, 3. own measures, 4. measures from the central budget, 5. measures from the budget of voivodeship, 6. measures from budgets of counties, 7. measures from budgets of municipalities, 8. measures from foreign countries, 9. ecological funds, 10. country credits and loans, 11. other measures
Source: Own study on the basis of: [4, 5].

Expenses for permanent assets serving the water management in Lublin Voivodeship in the years 1998-2013 according to directions of investing are presented in the table I. Total expenditures expenses for permanent assets serving the water management in the Lublin Voivodeship in the years 1998-2013 amounted approximately 0.5 billion US dollars indexed to the year 2014, and in the years 2001-2013 – 0.37 billion US dollars. As we can see, total expenditures and expenses for levees after the year 1998-1999 (after flood in

basins of Vistula and Oder rivers significantly leveled down). The highest level of total expenses and expenditures for levees was noted a year after the great flood in 2010.

Effects of investing in the water management in the years 2003-2013 in Lublin Voivodeship are presented in the table II. As we can see, data from this table are corresponding with data from the table I. In the period 2001-2013, there were built 23 water reservoirs (0.02 km³ capacity) and approx. 74 km levees. Almost half the length of levees were built in 2004-2005.

Active natural disasters' policy in Lublin Voivodeship, on behalf of the government is led by Regional Water Managements in Warsaw and Krakow, and by the Governor of Lublin, who is the representative of the government in the examined area[20].

On behalf of the local government for active natural disasters' policy are responsible: Lublin Regional Assembly (the legislative and control body) and the Regional Board in Lublin with the Marshal at the head (the executive body) in Lublin and subjected to him, local governmental organizational units of the voivodeship - including Provincial Management of Land Reclamation and Water Facilities in Lublin. Preventing the negative impact of natural disasters is also supported by funds from other sources, including: measures from budgets of municipalities and counties, measures from foreign countries, 9. ecological funds, country credits and loans.

Expenditures for permanent assets serving the water management in the examined area in the years 2002-2013 according to directions of financing sources are presented in the table III. As we can see, the main source of expenses were own measures of investors. From these source, there were financed nearly 38% of total expenditures in the examined period which totally amounted nearly 350 million US dollars.

TABLE IV.   CHARGES AND INCOMES TO THE FUND FOR ENVIRONMENTAL PROTECTION AND WATER MANAGEMENT IN LUBLIN VOIVODESHIP IN THE YEARS 2002-2014 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
2002	0	0	0	0	0	0	0	0	0
2003	453	17628	6842	7526	2286	0	975	2516	205
2004	533	16162	6073	8236	1504	0	349	1179	126
2005	426	12141	4028	6576	1087	0	450	1988	81
2006	654	15297	5876	7243	1801	0	377	321	0
2007	642	16300	6277	7237	2677	0	109	238	42
2008	571	20249	6127	7682	5955	0	486	139	40
2009	697	19722	5705	6015	7615	0	388	112	44
2010	732	19765	5332	6384	7537	0	512	728	41
2011	508	19163	5841	6302	6891	0	129	514	57
2012	507	17929	5816	6150	5862	0	101	119	41
2013	380	8790	2824	2673	2926	0	367	101	36
2014	456	15103	5656	4638	4375	0	435	171	58
total – all years	6103	183146	60741	72023	46139	0	4243	7955	714

1. years, 2. level of measures at beginning of the year, 3. total incomes, 4. incomes from charges for wastewater management and water protection, 5. incomes from charges for air protection and climate, 6. incomes from charges for waste management, 7. incomes from charges for other fields, 8. other incomes, 9. administered penalties, 10. incomes from penalties
Source: Own study on the basis of: [4, 5].

Expenses for water management were co-financed from ecological funds. One of these funds is the Fund for Environmental Protection and Water Management. Charges

and incomes to this fund in Lublin Voivodeship in the years 2002-2014 amounted over 183 million US dollars. As we can see in the table IV, incomes from charges for air protection and climate were the main part of total incomes and exceeded 72 million US $. It is worth noting that just over 34% of total charges and incomes were incomes from charges for waste management.

Investing expenses for the low water retention in the area of examined region in the years 2003-2014 according to investing directions are presented in the table V. As we can see, investing expenditures for man-made water reservoirs exceeded almost 19 million US dollars. Total investing expenses for the low water retention amounted approximately 21.5 million US $.

TABLE V.   INVESTING EXPENSES FOR THE LOW WATER RETENTION IN LUBLIN VOIVODESHIP – INVESTING DIRECTIONS IN THE YEARS 2003-2014 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

years	total	man-made water reservoirs	independent staging structures and water intakes on primary water-races	independent staging structures and water intakes on secondary water-races	staging lakes	fish ponds	other
2003	1557	1089	468	0	0	0	0
2004	1926	1402	349	17	0	157	0
2005	945	945	0	0	0	0	0
2006	972	951	21	0	0	0	0
2007	9560	9173	0	3	0	0	385
2008	1109	1102	7	0	0	0	0
2009	1865	1865	0	0	0	0	0
2010	0	0	0	0	0	0	0
2011	708	666	42	0	0	0	0
2012	1484	1484	0	0	0	0	0
2013	63	29	0	0	0	35	0
2014	1345	296	1002	0	0	0	47
total – all years	21536	19002	1890	20	0	192	432

Source: Own study on the basis of: [4, 5].

TABLE VI.   INVESTING EXPENSES FOR THE LOW WATER RETENTION IN LUBLIN VOIVODESHIP IN THE YEARS 2003-2014 – FINANCING SOURCES (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

years	total	from budget of Voivod	from Fund for Environmental Protection and Water Management	from Fund for protection rural areas	from structural funds	from budgets of self-governments	other
2003	1557	434	373	0	0	248	502
2004	1926	317	362	0	0	249	997
2005	945	615	156	0	0	174	0
2006	972	648	279	0	0	40	4
2007	9560	494	3153	0	0	3943	178
2008	1109	7	311	0	394	340	57
2009	1865	0	0	0	1002	864	1615
2010	0	0	0	0	0	0	0
2011	708	0	0	0	543	166	0
2012	1484	0	0	0	12	646	838
2013	63	0	13	0	29	21	29
2014	1345	94	0	0	1213	38	1213
total – all years	21536	2610	4647	0	3191	6730	5434

Source: Own study on the basis of: [4, 5].

Investing expenses for the low water retention in Lublin Voivodeship in the years 2003-2014 according to financing sources are presented in the table VI. As we can see, expenses were financed mainly from budgets of self-governments (6.7 million US dollar) and from from Fund for Environmental Protection and Water Management (4.6 million US dollar).

Effects of investing in the low water retention are presented in the table VII. The range of objects in Lublin Voivodeship in the years 2003-2014 amounted 76. Total increasing capacity in the examined period exceeded 6.7 million m^3.

TABLE VII.   EFFECTS OF INVESTING IN THE LOW WATER RETENTION – THE RANGE OF OBJECTS IN LUBLIN VOIVODESHIP IN THE YEARS 2003-2014

years	objects	increasing capacity	staging of lakes		artificial water reservoirs		fish ponds	staging structures	other	surface of irrigation
	1.	2.	1	2.	1.	2	1.	2.	1.	(ha.)
2003	7	83	0	0	2	83	0	0	5	0	235
2004	11	387	0	0	3	350	1	38	7	0	718
2005	2	805	0	0	2	805	0	0	0	0	154
2006	2	34	0	0	1	34	0	0	1	0	30
2007	2	2875	0	0	2	2875	0	0	0	0	0
2008	7	956	0	0	3	956	0	0	4	0	15
2009	1	925	0	0	1	925	0	0	0	0	280
2010	25	30	0	0	25	30	0	0	0	0	2
2011	3	72	0	0	1	72	0	0	2	0	20
2012	5	567	0	0	5	567	0	0	0	0	118
2013	7	19	0	0	6	4	1	16	0	0	0
2014	4	18	0	0	1	18	0	0	2	1	160
total - all years	76	6772	0	0	52	6718	2	53	21	1	1733

1. objects (units), 2. capacity ('000 m3)
Source: Own study on the basis of: [4, 5].

The active natural disasters' policy is executed by the Lublin Regional Assembly in accordance with the Law on Regional Government [17]. The Lublin Regional Assembly enacts and coordinates the implementation of, among others, the following documents: the Strategy for Development of Lublin Voivodeship for the years 2006-2020, the Spatial Development Plan of Lublin Voivodeship, the Environment Programme of Lublin Voivodeship for 2012-2015 with a view to the 2015, the Waste Management Plan Lublin Voivodeship 2017, the Regional Innovation Strategy of Lublin Voivodeship until 2020, the Regional Operational Programme of the Lublin Voivodeship 2014-2020.

The active natural disasters' policy is also executed by the Governor with the Lublin Provincial Office according to the Law on palatine and government administration in the province [17] with the financial support of the Marshal with the Lublin Region Regional Council. The Marshal supervises the Department of Safety and Crisis Management (WBiZK) in th Lublin Provincial Office. It operates on the basis of, among others, the law on crisis management [19]. The WBiZK prepares plans, including crisis management plan or plans, to counteract the effects of drought. The Lublin Provincial Office pays benefits targeted for individuals or families affected by the events exhibiting characteristics of a natural disaster. It runs the state budget funds to local government units to tasks related to the removal of the effects of events exhibiting characteristics of a natural disaster – losses in agriculture and public infrastructure and to activities related to the execution of tasks of the Provincial program for decommissioning of landslides

and protection against erosion loess ravines. The Lublin Office carries out a program of assistance to farm families in which farms and special sectors of agricultural production created in 2014, occurred damages caused by floods, torrential rain and hail. It recovers the energy supplements for vulnerable customers. In 2015, the Minister of Administration and Digitization, in the form of promissory notes, admitted nearly 1.5 million US dollar funding for the protection of loess ravines of Lublin Voivodeship and Świętokrzyskie Voivodeship [3].

Expenditures for public safety, fire and health protection in Lublin Voivodeship according to divisions of the Budget Classification in the years 2008-2011 are presented in the table VIII. As we can see, similar amounts were spent for the public safety and fire protection – approx. 380 million US $. Total expenses in the examined period amounted nearly 763 million US $. The largest share in expenditures for the public safety and fire protection had budgets of municipalities. The largest share in expenditures for the health protection had budgets of counties.

TABLE VIII.   EXPENDITURES FOR PUBLIC SAFETY, FIRE AND HEALTH PROTECTION IN LUBLIN VOIVODESHIP ACCORDING TO DIVISIONS OF THE BUDGET CLASSIFICATION IN THE YEARS 2008-2014 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

	public safety (ps) and fire protection (fp)	share of measures of individual budgets in total expenditures for ps and fp	health protection (hp)	share of measures of individual budgets in total expenditures for hp
measures from budget of Lublin Voivodeship	4172	1%	129649	34%
measures from budgets of counties	128629	34%	150175	40%
measures from budgets of municipalities	250144	65%	99942	26%
total	382945	100%	379765	100%
total ps, fp and hp	762711			

Source: Own study on the basis of: [4, 5].

Expenses for removing natural disasters' effects according to the Budget Classification in Lublin Voivodeship in the years 2008-2014 are presented in the table IX. As we can see, total expenditures amounted nearly 124 million US $. The largest share in expenses for removing effects of these kind of phenomena had measures from budgets of municipalities (69.2%) and the smallest – from the budgets of counties (10.8%). The largest part of measures from budgets of counties were spent for transport and communication – almost 12 million US dollar. From the budget of Lublin Voivodeship, there were mainly financed expenses in agriculture and hunting (23.7 million US dollar).

Considering data from tables VIII and IX, it is worth noting that the minimal share of expenses for removing natural disaster's effects in expenditures for the public safety, fire and health protection. As we can see in the table X, only 5 million US dollars expenses were spent for counteracting natural disaster's negative consequences from 383 million US dollar expenditures for the public safe and fire protection. Only 0.6 million US dollar expenses were spent for counteracting

natural disaster's negative consequences from 380 million US dollar expenditures for the health protection.

TABLE IX.   EXPENSES FOR REMOVING NAUTRAL DISASTERS' EFFECTS ACCORDING TO THE BUDGET CLASSIFICATION IN LUBLIN VOIVODESHIP IN THE YEARS 2008-2014 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

1.	2.	3.	4.	5.	6.	7.	8.
agriculture and hunting	23742	25.3%	0	0.0%	482	0.1%	24225
foerestry	0	0.0%	0	0.0%	0	0.0%	0
fishing	0	0.0%	0	0.0%	0	0.0%	0
mining	0	0.0%	0	0.0%	0	0.0%	0
manufacturing	0	0.0%	0	0.0%	0	0.0%	0
production and supply in electricity, gas and water	0	0.0%	0	0.0%	0	0.0%	0
hotels and restaurants	0	0.0%	0	0.0%	0	0.0%	0
transport and communication	495	0.5%	11571	22.7%	48155	14.8%	60222
tourism	0	0.0%	0	0.0%	0	0.0%	0
houising	291	0.3%	70	0.1%	889	0.3%	1250
service activities	0	0.0%	16	0.0%	2	0.0%	18
public administration	0	0.0%	27	0.1%	264	0.1%	291
public safety and fire protection	122	0.1%	1388	2.7%	3490	1.1%	5001
justice	0	0.0%	0	0.0%	0	0.0%	0
education	0	0.0%	52	0.1%	1723	0.5%	1775
higher education	0	0.0%	0	0.0%	0	0.0%	0
health protection	0	0.0%	98	0.2%	578	0.2%	676
social security	0	0.0%	29	0.1%	28190	8.7%	28220
other tasks in the field of social security	0	0.0%	0	0.0%	0	0.0%	0
educational care	0	0.0%	92	0.2%	0	0.0%	92
municipal engineering and environmental protection	0	0.0%	43	0.1%	617	0.2%	660
culture and national heritage protection	50	0.1%	0	0.0%	18	0.0%	68
botanical and zoological gardens, natural areas and objects of protected nature	0	0.0%	0	0.0%	0	0.0%	0
physical education and sport	0	0.0%	0	0.0%	1041	0.3%	1041
total expenditures	24701	20.0%	13385	10.8%	85450	69.2%	123536

1. division of budget classification, 2. measures from budget of Lublin Voivodeship, 3. share of measures from budget of Lublin Voivodeship in total expenditures, 4. measures from budgets of counties, 5. share of measures from budgets of counties in total expenditures, 6. measures from budgets of municipalities, 7. share of measures from municipalities in total expenditures, 8. total
Source: Own study on the basis of: [4, 5].

TABLE X.   THE SHARE OF EXPENSES FOR REMOVING NATURAL DISASTERS' EFFECTS IN EXPENDITURES FOR THE PUBLIC SAFETY, FIRE AND HEALTH PROTECTION IN LUBLIN VOIVODESHIP ACCORDING TO DIVISIONS OF THE BUDGET CLASSIFICATION IN THE YEARS 2008-2014 (IN '000 US DOLLARS INDEXED TO THE YEAR 2014)

	1.	2.	3.	4.	5.	6.
measures from budget of the Lublin Voivodeship	4172	122	2,9%	129649	0	0,0%
measures from budgets of counties	128629	1388	1,1%	150175	98	0,1%
measures from budgets of municipalities	250144	3490	1,4%	99942	578	0,6%
total	382945	5001	1,3%	379765	676	0,2%

1. public safety and fire protection, 2. including: removing natural disasters' effects, 3. share of measures for removing natural disasters' effects in expenditures for public safety and fire protection
4. health protection, 5. including: removing natural disasters' effects, 6. share of measures for removing natural disasters' effects in expenditures for health protection
Source: Own study on the basis of: [1].

III. SUMMARY

After the analysis of active natural disasters' policy in Lublin Voivodeship, the following conclusions can be drawn.

In the examined area, the active natural disasters' policy is financed from the budget of Lublin Voivodeship and budgets of local (self) governments. Expenditures are supported by means, among others, from the central budget, foreign, environmental funds, loans and advances.

Total expenditures expenses for permanent assets serving the water management in the Lublin Voivodeship in the years 1998-2013 amounted approximately 0.5 billion US dollars indexed to the year 2014. In the period 2001-2013, there were built 23 water reservoirs (0.02 km^3 capacity) and approx. 74 km levees. Almost half the length of levees were built in 2004-2005.

Total investing expenses for the low water retention amounted approximately 21.5 million US $. Investing expenditures for man-made water reservoirs exceeded almost 19 million US dollars. The range of objects in Lublin Voivodeship in the years 2003-2014 amounted 76. Total increasing capacity in the examined period exceeded 6.7 million m^3.

In Lublin Voivodeship there are executed programs for counteracting natural disasters' effects, e.g. Strategy for Development of Lublin Voivodeship for the years 2006-2020, the Spatial Development Plan of Lublin Voivodeship, the Environment Programme of Lublin Voivodeship for 2012-2015 with a view to the 2015, the Waste Management Plan Lublin Voivodeship 2017, the Regional Innovation Strategy of Lublin Voivodeship until 2020, the Regional Operational Programme of the Lublin Voivodeship 2014-2020.

The share of expenses for removing natural disaster's effects in expenditures for the public safety, fire and health protection is minimal. It is still too low level of financing.

REFERENCES

[1] Abbott P. L., *Natural disasters*, San Diego State University, San Diego 2009.

[2] Alexander D., *Natural disasters*, Wydawnictwo Springer Science & Business, Berlin 1993.

[3] S. Bukowski, Ponad 1 mld zł na usuwanie skutków klęsk żywiołowych, http://www.wspolnota.org.pl/aktualnosci/aktualnosc/ponad-1-mld-zl-na-usuwanie-skutkow-klesk-zywiolowych, 19.02.2015

[4] GUS, Bank Danych Lokalnych, http://www.stat.gov.pl/bdl/app/strona.html?p_name=indeks, 20.09.2015.

[5] MyBank.pl, http://kursy-walut-kupno-sprzedaz.mybank.pl/, 09.10.2015.

[6] A. Pichla, Rola nawodnień Kanału Wieprz-Krzna w gospodarce wodnej regionu, w: Wiadomości Melioracyjne i Łąkarskie, nr 2/2011 tom 54, wyd. Stowarzyszenie Inżynierów i Techników Wodnych i Melioracyjnych, Warszawa 2011.

[7] Z. Piepiora, Ekonomiczne aspekty lokalnej polityki przeciwdziałania skutkom katastrof naturalnych, wyd. Zbigniew Piepiora, Kowary 2012.

[8] Plan gospodarki odpadami dla województwa lubelskiego 2017, Zarząd Województwa Lubelskiego, Lublin 2012.

[9] Plan zagospodarowania przestrzennego województwa lubelskiego. projekt, Zarząd Województwa Lubelskiego, Lublin sierpień 2015.

[10] Program ochrony środowiska województwa lubelskiego na lata 2012-2015 z perspektywą do roku 2019, Zarząd Województwa Lubelskiego, Lublin 2012.

[11] Regionalna strategia innowacji województwa lubelskiego do roku 2020, Zespół Zadaniowy ds. Aktualizacji Regionalnej strategii innowacji województwa lubelskiego do roku 2020, Lublin 2014.

[12] Regionalny Program Operacyjny Województwa Lubelskiego na lata 2014-2020, przyjęty decyzją Komisji Europejskiej C(2015)887 z dnia 12 lutego 2015 r.

[13] Rewitalizacja terenów zalewowych i starorzeczy, http://www.bagna.pl/index.php?page=metody/rewitalizacja_zalewowych, 05.06.2009.

[14] J. Rodzik, Wpływ deszczów ulewnych i roztopów na rozwój wąwozu lessowego, w: Landform Analysis, Vol. 8, wyd. Stowarzyszenie Geomorfologów Polskich, Poznań 2008.

[15] U. Somorowska, Wzrost zagrożenia suszą hydrologiczną w różnych regionach geograficznych Polski w XX wieku, w: Prace i Studia Geograficzne, tom 43, Uniwersytet Warszawski, Warszawa, 2009, s. 97-114.

[16] Strategia rozwoju województwa lubelskiego na lata 2006-2020, Zarząd Województwa Lubelskiego, Lublin lipiec 2005.

[17] Ustawa z dnia 5 czerwca 1998 r. o samorządzie województwa, Dz. U. 1998 nr 91 poz. 576 z późniejszymi zmianami.

[18] Ustawa z dnia 18 kwietnia 2002 r. o stanie klęski żywiołowej Dz. U. 2002 nr 62 poz. 558 z późniejszymi zmianami.

[19] Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym, Dz.U. z 2007 r. nr 89, poz. 589 i 590 z późniejszymi zmianami, art. 16.

[20] Ustawa z dnia 23 stycznia 2009 r. o wojewodzie i administracji rządowej w województwie, Dz. U. z 2009 r. nr 31, poz. 206.

- All published papers undergone single blind peer review.
- All published papers are in English language only. Each paper had assigned 2 reviewers and each paper went through two-tier approval process.
- Papers are published as delivered by authors without content modification. All accepted papers have been formally checked by the conference Technical Committee and regularly reviewed in single blind peer reviewing process by the conference Scientific Committee (Reviewers Committee).
- The publication contents are the sole responsibility of the publication author and it does not reflect the Publishing Society, Zilina, Slovakia.

Open Access Online archive is available at: **http://www.rcitd.com/archive**
(proceedings will be available online one month after the publication release).

In case of any questions, notes or complaints, please contact us at: **info(at)rcitd.com**.

Paper Citation Example: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, Proceedings of The 3rd International Virtual Research Conference In Technical Disciplines (RCITD-2015), Vol. 3, pg #-#. Zilina, Slovakia.

⚠ **Warning**

RCITD 2015