ICTRS 2015

Fourth International Conference on

Telecommunications and Remote Sensing

Including a Special Session on

e-Health Services and Technologies (EHST)

Proceedings

Rhodes, Greece • 17-18 September 2015

Organized by: JICREST

Co-Organized by: ARSit, Aristotle University of Thessaloniki

Under the auspices of: U.R.S.I.

Cooperating Organizations: AMAKOTA, BULATSA, Bulgarian Academy of Sciences

ICTRS 2015

Proceedings of the

Fourth International Conference on Telecommunications and Remote Sensing

Including a Special Session on
eHealth Services and Technologies (EHST)

Rhodes, Greece

17-18 September 2015

Organized by
IICREST - Interdisciplinary Institute for Collaboration and Research on Enterprise Systems and Technology

Co-Organized by
ARSIT - Association for Radio Systems and Intelligent Telecommunications
AUTH - Aristotle University of Thessaloniki

In Collaboration with
AMAKOTA Ltd.
BULATSA - Bulgarian Air Traffic Services Authority
IMI - Institute of Mathematics and Informatics of Bulgarian Academy of Sciences

Under the Auspices of
URSI - International Union of Radio Science

Edited by:
Blagovest Shishkov, Andon Lazarov, Naoki Shinohara, Leszek Maciaszek, Boris Shishkov

Graphics Production by Canka Petrova

Printed in Bulgaria

ISBN: 978-989-758-152-6

Depósito Legal: 396778/15

http://www.ictrs.org

secretariat@iicrest.org

BRIEF CONTENTS

KEYNOTE SPEAKER

Dimitris Mitrakos

Aristotle University of Thessaloniki

Greece

CHAIRS AND PROGRAM COMMITTEE

Program Co-Chairs

Blagovest Shishkov, Bulgarian Academy of Sciences, Bulgaria
Andon Lazarov, Burgas Free University, Bulgaria
Naoki Shinohara, Kyoto University, Japan
Leszek Maciaszek, Wroclaw Univ. of Economics, Poland / Macquarie Univ., Australia

General Chair

Boris Shishkov, Bulgarian Academy of Sciences / IICREST, Bulgaria

PROGRAM COMMITTEE

Oleksiy Agapitov, Taras Shevchenko Kyiv National University, Ukraine

Catherine Algani, CNAM, France

Makoto Ando, Tokyo Institute of Technology, Japan

Mauro Assis, Brazilian Committee of URSI, Brazil

Krassimir Atanassov, Bulgarian Academy of Sciences, Bulgaria

Boncho Balabanov, NBU, Bulgaria

Vera Behar, Bulgarian Academy of Sciences, Bulgaria

Maurice Bellanger, CNAM, France

Jun Cheng, Doshisha University, Japan

Lyubka Doukovska, Bulgarian Academy of Sciences, Bulgaria

Kazimierz Fraczkowski, Wroclaw University of Technology, Poland

David Fuschi, Bridging Consulting Ltd., UK

Yoshiharu Fuse, USEF, Japan

Ivan Garvanov, Bulgarian Academy of Sciences, Bulgaria

Mihail Iliev, University of Ruse “Angel Kanchev”, Bulgaria

CHAIRS AND PROGRAM COMETEE (CONT.)

Hristo Kabakchiev, Sofia University "St. Kl. Ohridski", Bulgaria

Dorina Kabakchieva, UNWE, Bulgaria

Mesut Kartal, Istanbul Technical University, Turkey

Stefan Kirn, University of Hohenheim, Germany

Kazuya Kobayashi, Chuo University, Japan

Karol Kozak, Technical University Dresden, Germany / Wroclaw University of Economics, Poland

Mohamed Latrach, ESEO, France

Frank Little, Texas A & M University, USA

Marco Luise, University of Pisa, Italy

Francois Lefeuvre, LPC2E / CNRS, France

Olga Maltseva, Southern Federal University, Russia

Andrea Massa, University of Trento, Italy

Galia Marinova, Technical University of Sofia, Bulgaria

Wolfgang Mathis, Leibniz Universität Hannover, Germany

Lyudmila Mihaylova, Lancaster University, UK

Tomohiko Mitani, Kyoto University, Japan

Tadao Nagatsuma, Osaka University, Japan

Shoichi Narahashi, NTT DOCOMO, Inc., Japan

Marin Nenchev, Technical University of Sofia, Bulgaria

Martin O'Droma, University of Limerick, Ireland

Takashi Ohira, Toyohashi University of Technology, Japan

Jacques Palicot, SUPELEC / IETR, France

Chairs and Program Cometee (cont.)

Hermann Rohling, Hamburg University of Technology, Germany

Sana Salous, Durham University, UK

Hamit Serbest, Cukurova University, Turkey

Boris Shishkov, Bulgarian Academy of Sciences / IICREST, Bulgaria

Alexander Shmelev, Russian Academy of Sciences, Russia

Marten van Sinderen, University of Twente, The Netherlands

Angela Slavova, Bulgarian Academy of Sciences, Bulgaria

Jun-Ichi Takada, Tokyo Institute of Technology, Japan

Hiroyuki Tsuji, NIICT, Japan

Ing Widya, University of Twente, The Netherlands

Satoshi Yagitani, Kanazawa University, Japan

Tsuneki Yamasaki, Nihon University, Japan

Antoni Zwiefka, "A. Falkiewicz” Specialist Hospital, Poland

BEST PAPERS SELECTION

The authors of up to three selected ICTRS'15 best papers will be invited to submit revised and extended versions of their papers to a special issue of the international journal

ANNALS OF TELECOMMUNICATIONS

FOREWORD

Bringing together ***telecommunications*** (through areas, such as *radio communications* and *signal processing*) and ***remote sensing*** (through areas, such as *pattern recognition* and *context awareness*), and enriching this with *eHealth*-related topics, is characterizing ICTRS 2015 – the *Fourth International Conference on Telecommunications and Remote Sensing*.

This book contains the proceedings of ICTRS 2015, held in Rhodes, Greece, on 17-18 September 2015. The proceedings consists of 19 high-quality research and experience papers that have not been published previously. Those papers have undergone a detailed peer-review process and were selected based on rigorous quality standards.

Since 2012, we have enjoyed three successful ICTRS editions, namely: *Sofia 2012*, *Noordwijkerhout 2013*, and *Luxembourg 2014*, and this is mainly thanks to ***Prof. D.Sc. Blagovest Shishkov*** who inspired and led this conference, being its spiritual father. It is with great pain and sorrow that we mention the unexpected decease of Blagovest Shishkov at the age of 77. Our memory of him as a respected person as well as a great scientist will remain vivid. An inspiring person of a brilliant mind, he also always impressed with his honesty and modesty. Always devoted to his family, he was a loving husband and a caring, wonderful father of two sons. He was a great friend and an excellent team-mate.

Blagovest's work on telecommunications was remarkable, marked by impressive achievements and career. Graduating with honors secondary school in his home town – Topolovgrad, situated in South Bulgaria, he moved to Sofia to study electrical engineering at the Technical University. He got his MSc in 1961. This was followed by inspiring research activities related to noise detection, at the Bulgarian Academy of Sciences, Institute of Electronics. In 1974 Blagovest earned a Doctorate (PhD) in Physics (exploring the *noise processes in receiving visual information*) and in 1991 he got his D.Sc. degree in telecommunications (exploring the *optimal statistical synthesis of information systems within a-priori undetermination*). From 1994, Blagovest was Full Professor of Applied Mathematics at the Technical University of Sofia and after his retirement in 2003, he was Emeritus Full Professor of Statistical Signal Processing at the Bulgarian Academy of Sciences, Institute of Mathematics and Informatics. Blagovest published over 100 scientific and professional articles and several books, many of which touching upon *signals transmission* (signals, filtering, and detection) and especially *asymptotic methods in parameter estimation*, signal detection and identification, pattern recognition and data quantization. His last work was related to *signal processing of cyclostationary signals* and detecting and studying of nonlinear wave interactions by using *higher-order statistics*. Both areas have been successfully incorporated into *adaptive antenna beamforming* and analysis of time series associated with space data. Some of his last research was devoted to mathematical modeling and optimization of the sidelobe level of large antenna arrays toward SPS/MPT applications. Blagovest created a variety of algorithms applicable to telecommunications, radar, remote sensing systems, space research, biomedical systems, and so on. *Asymptotic algorithms* for detection and estimation of weak signals on a background of noise are to be especially mentioned as well, together with new algorithms in *adaptive antennas* and *space data analysis*.

He will be missed by everybody who had the privilege of knowing him, and we are honored to dedicate our work on furthering ICTRS to the memory of Blagovest!

Further, we are especially grateful to our EHST Colleagues who contributed interesting eHealth-related papers to the Special Session on *eHealth Services and Technologies* (EHST), reflecting possible applications of telecommunications, remote sensing, (context-aware) information systems, and IT in general, in resolving problems and addressing challenges relevant to a domain of high societal importance, such as Health-care.

ICTRS 2015 was organized and sponsored by the *Interdisciplinary Institute for Collaboration and Research on Enterprise Systems and Technology* (IICREST), being co-organized by the *Association for Radio Systems and Intelligent Telecommunications* (ARSIT) and *Aristotle University of Thessaloniki* (AUTH). Cooperating organizations were the BAS *Institute of Mathematics and Informatics* (IMI), the *Bulgarian Air Traffic Services Authority* (BULATSA), and *AMAKOTA* Ltd. Further, the conference is performed under the auspices of the *International Union of Radio Science* (URSI), covering topics relevant to the work in a number of *URSI Commissions*, namely commissions *A, B, C, D, E, F, G, H,* and *K.*

Publisher of the current proceedings is SCITEPRESS. Besides printed proceedings we also deliver an electronic version – all presented papers will be made available on the *SCITEPRESS Digital Library* by November 2015. Furthermore, the authors of up to three selected papers presented at ICTRS 2015 will be invited to submit revised and extended versions of their papers to a special issue of the international journal *Annals of Telecommunications*.

As it can be seen from the papers, ICTRS 2015 is addressing a large number of topics relevant to the areas of Telecommunications and Remote Sensing: Wireless Telecommunications and Radio Wave Propagation, Signal Detection and Signal Processing, Cognitive Radio and Context-Awareness, Radio Positioning and Localization, Data Transmission and Networking, Wireless Sensor Networks and Transmission Power, Surface Conductivity, Electromagnetic Radiation. This is complemented by the Special Session on e-Health Services and Technologies, addressing in turn topics, such as: Modeling of Hospital Processes and Data, Tele-Monitoring, Big Data and Machine Learning, Interoperability, Human-Computer Interaction, IT Products Usability, and also bringing forward practical experience on introducing e-Health services and technologies in Bulgaria, Burundi, and Turkey. This all allows for an inspiring interdisciplinary technical program that we hope will be interesting and stimulating for all participants.

FOREWORD (CONT.)

The high quality of the ICTRS 2015 program is enhanced by a keynote lecture, delivered by *Dimitris Mitrakos* (Aristotle University of Thessaloniki) who is a renowned expert in his field. The keynote speaker and some other ICTRS'15 / EHST'15 participants will take part in a panel discussion and also in other discussions stimulating community building and facilitating possible R&D project acquisition initiatives. Those special activities will definitely contribute to maintaining the event's high quality.

Organizing this interesting and successful conference required the dedicated efforts of many people. Firstly, we must thank the authors, whose research and development achievements are recorded here. Next, the program committee members each deserve credit for the diligent and rigorous peer-reviewing. Further, we would like to mention the excellent organization provided by the IICREST team, and especially *Canka Petrova* and *Aglika Dikova* plus the co-organizers, ARSIT and AUTH (supported by the logistics partner, AMAKOTA Ltd.) – the team did all the necessary work for delivering a stimulating and productive event, and we have to acknowledge also the inspiring support of our colleagues from Thessaloniki and our local partners in Rhodes, especially *Vassilis Pappas*. We are grateful to SCITEPRESS for their willingness to publish the current proceedings and we would like to especially mention *Vitor Pedrosa* with whom we collaborated excellently on the proceedings preparation. Last but not least, we thank the keynote speaker, *Dimitris Mitrakos*, for his invaluable contribution and for his taking the time to synthesize and deliver his talk.

We wish you all an inspiring conference and an enjoyable stay in the beautiful Rhodes island. We look forward to seeing you next year in Milan, Italy, for the Fifth International Conference on Telecommunications and Remote Sensing (ICTRS 2016), details of which will be made available on http://www.ictrs.org.

Boris Shishkov
Bulgarian Academy of Sciences / IICREST, Bulgaria

Andon Lazarov
Burgaz Free University, Bulgaria

Naoki Shinohara
Kyoto University, Japan

Leszek Maciaszek,
Wroclaw University of Economics, Poland / Macquarie University, Australia

CONTENTS

CONTENTS (CONT.)

KEYNOTE SPEAKER

Low Power Wireless Communications and the Internet of Things

Dimitris Mitrakos
Aristotle University of Thessaloniki, Thessaloniki, Greece

Abstract: The IEEE 802.15.4 standard offers the basis upon which a number of wireless data communications technologies are built. Characterised by key features such as high energy efficiency, low latency, high accuracy and low cost, these technologies play a significant role in the rapid development of new technological ecosystems that help shaping the emerging Internet of Things. The standard sets specifications for the RF, PHY and MAC layers. Augmented with higher layer ad hoc protocols and applications these specifications allow single chip complete radio controllers to create cost effective self-healing mesh networks of both sensors and actuators. Under the umbrella of Internet of Things, these devices allow people and objects in the physical world as well as data and virtual environments to interact with each other, thus enabling smart digital environments in important domains of human activity such as home, transport, health or energy. Mathematical modeling and performance assessment of the basic IEEE 802.15.4 technologies has been the focus of many studies in order to help understanding better the fundamental factors affecting their efficiency. Despite its complexity, the joint evaluation of communications mechanisms at both PHY and MAC layers is expected to lead to more realistic results and to help developing more efficient systems. At the MAC layer, efficiency is mainly determined by the foreseen multiple access mechanism CSMA/CA and its ability to resolve collisions of packets that originate from various network nodes. However, this efficiency is further affected by the path loss, multi path fading, shadowing, noise and interference exhibited in the underlying wireless channel. The presentation will review a number of approaches developed in the past for joint modeling of PHY and MAC layers in IEEE 802.15.4 wireless networks and present an analytical model based on the M/M/1 and M/G/1 queuing systems with service interruptions or vacations in an attempt to capture the stochastic behavior of signal propagation particularly in low power wireless networks. Simulation as well as experimental results will be discussed in the light of real life implementations for urban traffic management and public transport optimisation currently under development in a framework of smart cities applications.

Brief Bio: Dimitris Mitrakos is an Associate Professor at the Department of Electrical Engineering, School of Engineering, Aristotle University of Thessaloniki, Greece. His research interests include internet computing, multimedia communications, sensor and digital telemetry networks and distributed control and teleoperations systems. Dimitris has a Diploma in Electrical Engineering from Aristotle University of Thessaloniki, an MSc in Communications Engineering from University of Manchester Institute of Science and Technology, a DIC in Signal Processing and a PhD in Electrical Engineering from University of London Imperial College of Science and Technology. In the recent past, he has been Vice-Chairman of the Electrical Engineering Department and Director of the Electronics and Computer Section of the Electrical Engineering Department of

Aristotle University of Thessaloniki. Today he is member of the Managing Board and Director of Postgraduate Studies of the Aristotle University of Thessaloniki Interdisciplinary Postgraduate Program in “Language and Communications Sciences”. Since early 90s Dimitris has been member of Workprogramme Preparation Committees, Proposal Technical Evaluation Committees, Evaluation Process Assessment Committees and Project Technical Review Committees for the European Commission Advanced Informatics in Medicine, Telematics for Education, Trans-European Networks, COPERNICUS, ERDF Article 10, MEDIA Plus and Information Society Technologies Programs. He has also been member of the Greek Steering Committee for the European Commission Inter-Regional Information Society Initiative.

Full Papers

Usage of the Internet Resources for Research of the Ionosphere and the Determination of Radio Wave Propagation Conditions

Olga Maltseva
Institute of Physics Southern Federal University, Stachki, 194, Rostov-on-Don, Russia
mal@ip.rsu.ru

Keywords: GPS. Total electron content TEC. Ionospheric models. Radio wave propagation. Geomagnetic disturbances.

Abstract: In the field of ionospheric research is impossible to obtain the new results, new knowledge without use of Internet resources. The paper provides examples of the use of these resources on the basis of generalization and complement to reports made on the three previous ICTRS conferences. Results are presented in four areas of possible data: (1) vertical sounding, (2) total electron content TEC, measured by high orbit navigation satellites, (3) plasma frequencies measured by low orbit satellites, (4) empirical models of the ionosphere. The main achievement in the first direction is the creation of GAMBIT, designed to provide global maps of ionospheric parameters foF2 and hmF2 with delay of 15 minutes in relation to real time. The estimation of conformity of the IRI model to experimental data of foF2 for high-latitude station of southern hemisphere is made. Within the second direction the effectiveness coefficient of use of the observational median of the equivalent slab thickness in comparison with thickness of the IRI model is introduced. It is shown that this coefficient almost always exceeds 1, reaching values 1.5-2 globally. Behavior features of deviations of calculated foF2 from observational values during the strongest geomagnetic disturbances of April 2014 and March 2015 are given. In the framework of the third direction validation of a plasmaspheric part of a N(h)-profile according to satellite IMAGE data is performed. It was concluded that to disambiguate N(h)-profiles it is necessary to improve both values of TEC, and the shape of the topside part because the plasmaspheric part is close to existing model RPI. Within the fourth direction the statistics of comparisons of various models was increased including high-latitude region of the southern hemisphere.

1 INTRODUCTION

Internet affects the lives of every person, providing huge opportunities in the information sphere. This paper describes the opportunities associated with databases for ionospheric research. To solve many scientific and technological problems need to know the conditions in the ionosphere, within which there are the satellites that provide us, in turn, information about the state of the ionosphere. This article points out the most important databases and displays the results of their use. Possibilities of use of Internet data are illustrated with examples: (1) vertical sounding (VS), (2) global maps of the total electron content of the ionosphere TEC, (3) data satellites CHAMP and DMSP, and (4) the International Reference Ionosphere model IRI. Each of these areas corresponds to a separate section (2-5).

2 INTERNET POSSIBILITIES OF USAGE OF THE VERTICAL SOUNDING DATA

In the 20th century, the key parameter of the ionosphere was the peak concentration NmF2 (or critical frequency foF2), measured by special receiver-transmitters - ionosondes. By the end of the 20th century, digisondes, being ionosondes with automatic processing of samples, have been appeared when the need of the ionospheric data in real time and on a global scale has required to automate this process. A lot of programs have been created. System ARTIST for which now already the fifth version (Reinisch et al., 2009) was developed is most widely used. However, as shown by means of special expert program QualScan estimating quality of ionograms, usually 1/3 of digitized ionograms can be rejected (McNamara, 2006). The most promising is a new principle, which underlies dynasonde

(Zabotin et al., 2005) and provides the most accurate determination of the parameters of the ionosphere. The evaluation of improvements of ionospheric parameters determination by dynasonde in comparison with digisonde (ionosonde) was held in the paper (Maltseva et al., 2007). The most important are results for foF2 and hmF2. Figure 1 from this paper is an example of difference between automatic methods of determination of daily dependence of frequency foF2 by various methods of sounding: digisonde (method POLAN (P)) and dynasonde (method NeXtYZ (N) (Zabotin et al., 2005)) in comparison with the IRI model.

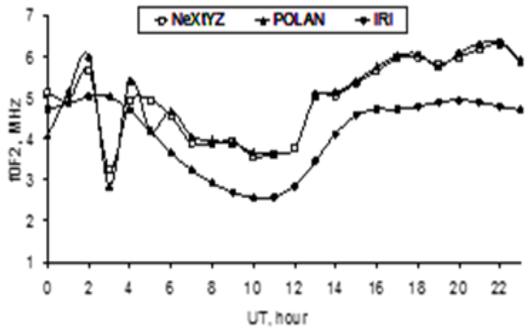

Figure 1: An example of differences between methods of foF2 determination: digisonde (method POLAN (P)) and dynasonde (method NeXtYZ (N)).

Table 1 shows the quantitative evaluation of absolute differences between critical frequencies obtained by these methods and the values of the IRI model. Results are shown for stations having dynasonde and for those days which had simultaneous measurements. The first column contains name of stations. The second column shows the month. The third column points out number of day when there were measurements. P-IRI column shows the difference between the method POLAN and the IRI model, N-IRI shows the difference between the method NeXtYZ and the IRI model, N-P column shows the difference between the methods NeXtYZ and POLAN. Values of P-IRI and N-IRI are an estimate of improving the determination of foF2 using experimental data. Values N-P are an estimate of improving the determination of foF2 by method NeXtYZ (dynasonde) compared with method POLAN (digisonde). If you take the average value of foF2 approximately 5 MHz, the improvement is 11-14%.

Table 1: Improvement of foF2 determination by new dynasonde.

station	month	Number of days	$\lvert\Delta$foF2$\rvert$, MHz		
			P-IRI	N-IRI	N-P
Lycksele	July	16	0.64	0.62	0.16
Tromso	July	16	0.57	0.59	0.19
BearLake	July	5	0.74	0.72	0.18
Tromso	Nove	3	0.6	0.6	0.10
Lycksele	April	5	0.83	0.76	0.14
Tromso	April	6	0.69	0.66	0.17

Figure 2 shows the difference between automatic methods of determining the values of hmF2, calculated by different methods of sounding: digisonde (method POLAN (P)) and dynasonde (method NeXtYZ (N)). Also values of hmF2 for the IRI model are given.

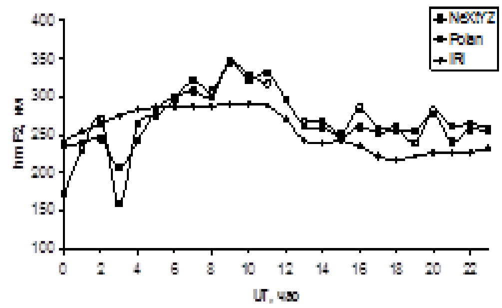

Figure 2: An example of differences between values of hmF2 obtained by various methods of sounding: digisonde (method POLAN (P)) and dynasonde (method NeXtYZ (N)).

Table 2 shows the quantitative evaluation of these differences (Maltseva et al., 2007). P-IRI column shows the difference between the method POLAN and the IRI model, N-IRI shows the difference between the method NeXtYZ and the IRI model, N-P column shows the difference between the methods NeXtYZ and POLAN.

As well as in a case with parameter foF2, values P-IRI and N-IRI characterize improvement in adapting the model to the experimental values of hmF2. N-P values have improved the definition of the maximum height by the new method. These values are significantly higher than the accuracy of measurement. Using as the mean value of hmF2 250 km we obtain approximately 15% improvement.

Table 2: Improvement of hmF2 determination by new dynasonde.

station	month	Numb-er of days	|ΔhmF2|, км		
			P-IRI	N-IRI	N-P
Lycksele	July	16	40.6	38.5	34.3
Tromso	July	16	48.5	44.7	29.2
Bea Lake	July	5	47.8	53.2	25.2
Lake	Novem	3	49.2	40.5	33.7
Tromso	April	5	20.3	20.5	24.5
Lycksele	April	6	29.5	19.2	29.7

Unfortunately, the stations equipped by dynasonde are not a lot. Data of other stations are collected in the form of multiple databases (GIRO, SPIDR, DIAS et al.) which are freely available, include data from several cycles of solar activity, and are permanently being updated and modified. A large contribution to this update provide reports from conference (URSI, 2015), in particular, the system GIRO was modified in GAMBIT (https://git.giro.uml.edu), which represents the global maps of foF2 and hmF2 delayed at least 15 minutes, compared with real time. The European System of DIAS was supplemented with the data of such stations as Moscow and Tromso.

3 INTERNET POSSIBILITIES OF USAGE OF THE TOTAL ELECTRON CONTENT OF THE IONOSPHERE

Despite the fact that, apparently, the method of the VS can be considered the best method for measuring the ionospheric parameters, its main deficiencies are rare network and quite often data gaps on some stations. In the 21st century, with the development of network of navigation satellites has been a shift to the identification of the total electron content (TEC) of the ionosphere as the main parameter, and even proposed to replace the parameter NmF2. The advantages of this option is its continuous monitoring, the availability of online databases of global maps for a period longer than the cycle of solar activity, information about the N(h)-profile. In (Maltseva et al., 2012; 2013; Maltseva, 2014) the TEC parameter was used to determine foF2 with experimental median of ionospheric equivalent slab thickness τ(med). Much more conformity of the calculated values of foF2 with measurements in comparison with conventionally used value τ(IRI) (McNamara, 1985; Houminer, Soicher, 1996; Gulyaeva, 2011) was obtained. The improvement is achieved by taking into account a big difference between what was (τ(IRI) = TEC(IRI)/NmF2(IRI), NmF2(calc) = TEC(obs)/τ(IRI)), and what became (τ(med) = med(TEC(obs)/NmF2(obs)), NmF2(calc) = TEC(obs)/τ(med)). Illustration of this difference for the reference station Juliusruh is shown in Fig. 3. In this paper we introduce the efficiency factor Keff of using τ(med) compared with τ(IRI), and give the results of its determination for the individual stations in the various regions of the world and globally. Keff factor is defined as the quotient of the absolute deviations |ΔfoF2| of calculated foF2 from the experimental values using these two parameters τ(IRI) and τ(med). Fig. 4 shows the coefficients for the 6 stations: the European mid latitude Juliusruh station, the American mid latitude Goosebay station, the American high-latitude station Thule in the northern hemisphere, the equatorial Ascension Island station, the low-latitude and high-latitude Grahamstown and Mawson stations of the Southern hemisphere. For comparison, the value of K=1 is displayed. Integrated characteristics are given in Table 3.

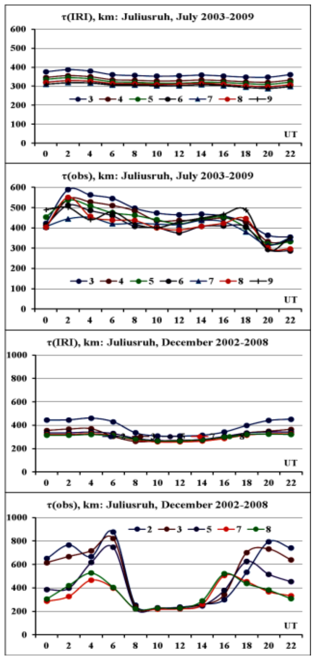

Figure 3: Differences between used values τ(IRI) and τ(med) on an example of the long-year data for July and December and the Juliusruh station.

.

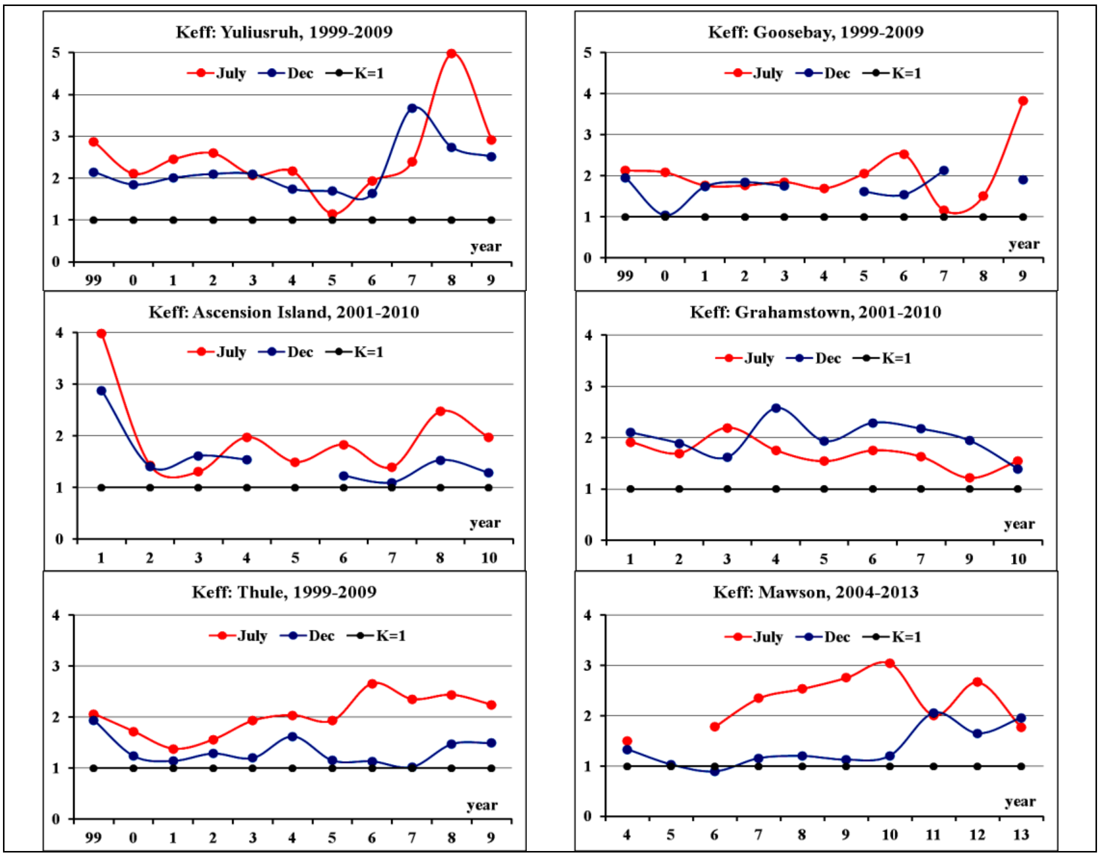

Figure 4: Coefficients Keff in various regions of the globe.

Table 3: Conformity averaging for all years between the calculated and observed values of foF2: coefficients Keff and the relative deviations σ (foF2), %

σ(foF2), %								
	July				**December**			
station	Keff	med	ins	τ	Keff	med	ins	τ
Juliusruh	2.52	9.82	15.03	6.91	2.20	13.30	17.68	8.45
Goosebay	2.03	9.41	15.46	9.36	1.72	13.54	17.61	11.36
Thule	2.03	11.46	15.43	8.46	1.34	10.41	18.21	13.91
Grahams	1.70	9.53	15.84	10.09	1.99	12.44	17.26	9.40
Ascension	1.98	23.40	27.21	14.89	1.57	13.04	17.52	12.45
Mawson	2.26	44.26	48.83	25.7	1.36	15.28	24.82	20.75

It can be seen that the efficiency factor is close to 2 for July, i.e. results are in 2 times better, and near 1.5 for December. The same can be said about the relative deviation σ(foF2),%. With the exception of the Antarctic station Mawson deviations for τ(med) lie within 10%. Keff has been calculated globally for months with disturbances. Figure 5 shows the average monthly variations depending on the latitude for τ(IRI) (green triangles). For τ(med) red dots show results for the global map JPL. Figure 5 also includes the efficiency factors. Blue circles in all Figures show results of comparison of model IRI values with experimental medians for an illustration of accuracy of the model in relation to measurement.

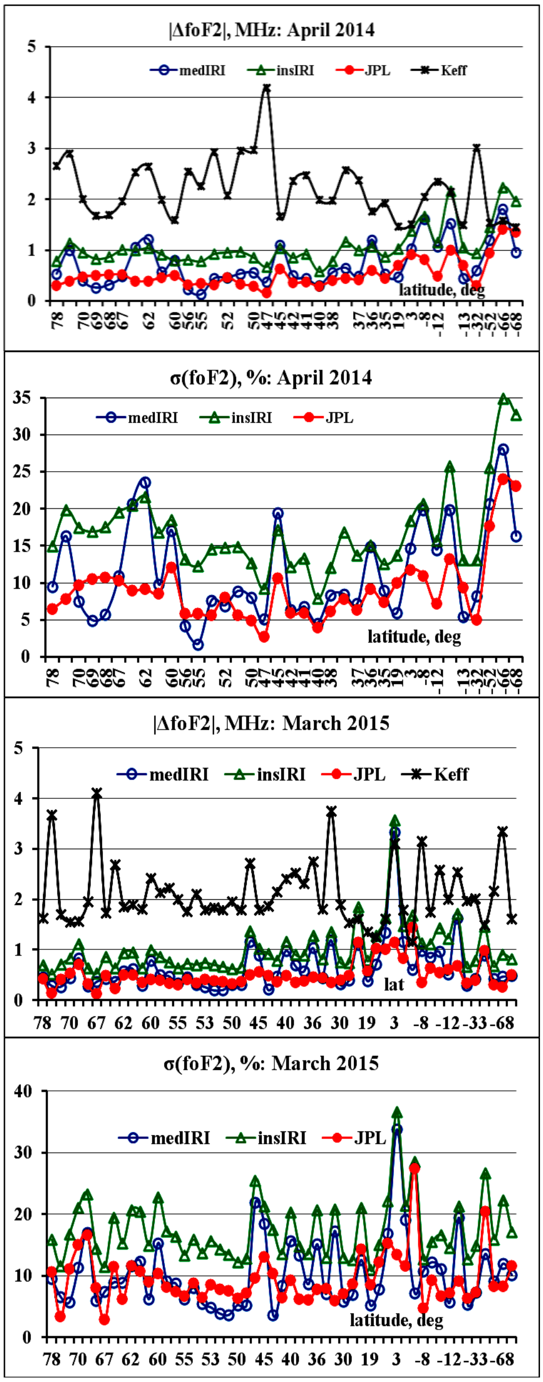

Figure 5: Behavior of Keff on a global scale.

Keff is close to 2 around the globe, and σ is less than 10% with the exception of the stations of the Southern hemisphere. This indicates that the simulation situation for the southern hemisphere has problems. For these months, the details of the comparison of deviations have been identified in disturbed periods. Fig. 6 shows the Dst-index for these months (http://wdc.kugi.kyoto-u.ac.jp/dstdir/).

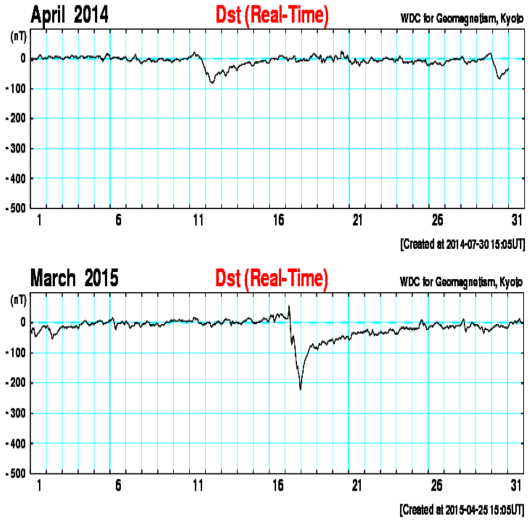

Figure 6: Behavior of Dst-indexes in months with disturbances

The behavior of Dst-index specifies which days had the greatest disturbances. In April 2014 this day is 12. In March 2015 these are 17-19 because disturbance was stronger. Fig. 7 presents the global results for specific days. In both cases values are also given for the quiet days preceding the disturbances. Two top figures show the results for τ(IRI), the lower - for τ(med), and these charts include in addition the values for τ(IRI), averaged over the 5 days considered.

We see that deviations for τ(IRI) are maximum in disturbed days. In the southern hemisphere still get big deviations also for τ (med). To conclude this Section it is necessary to say a few words about the alternatives. In recent years, assimilation techniques become widespread (e.g. Khattatov et al., 2005; Fuller-Rowell et al., 2006), in which N(h) -profile of the ionosphere is fitted by the very powerful methods so that it satisfies the measured TEC(obs), but emphasis on the determination of foF2 was not done. And when compared with measurements by ionosondes large discrepancies are obtained, which, for example, are shown in Fig. 8 from (Yao et al., 2014). Differences are seen not only in critical frequencies, but also in a profile of the topside part although in general density distribution throughout the profile provides conformity with the TEC(obs) for the account of the plasmaspheric part.

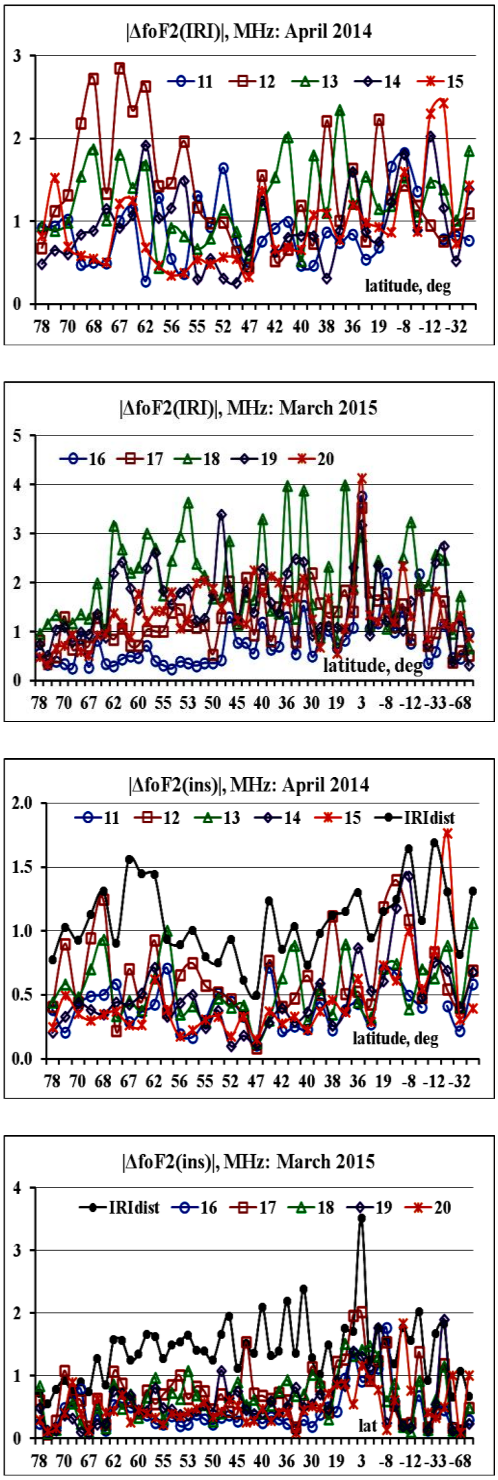

Figure 7: Details of deviation comparison in the disturbed periods.

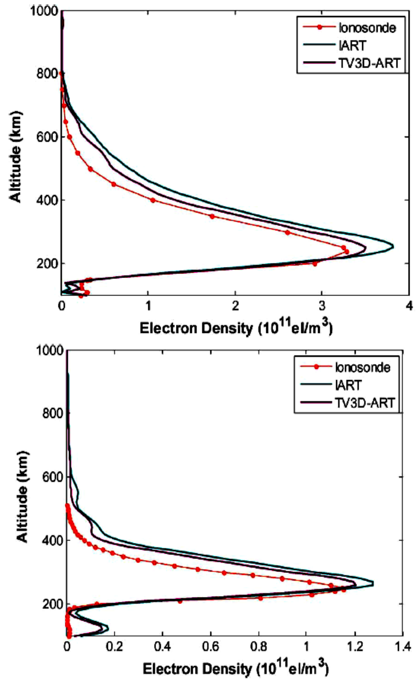

Figure 8: Conformity of assimilation profiles with the data of vertical sounding (from Yao et al., 2014).

4 INTERNET POSSIBILITIES OF USAGE OF THE LOW ORBIT SATELLITES DATA

By the end of the 20th century, data from different satellites began to play a larger role in the study of the ionosphere and radio propagation in it. One of the achievements of the 21st century is a creation of database of satellite CHAMP (http://isdc.gfz-potsdam.de/), which includes the value of the electron density at altitudes near the maximum height hmF2. Currently, this array also covers virtually one cycle of solar activity (2000-2011). Even earlier, similar data were provided by a series of satellites DMSP (http:// cindispace. utdallas.edu/ DMSP/ dmsp_data_at_utdallas.html). Since 2006, the huge amount of data was obtained by satellite COSMIC (http://tacc.cwb.gov.tw/en/). One of the goals of the simulation is to coordinate the model N(h)-profile and experimental TEC(obs). The IRI-Plas model (Gulyaeva, 2011) provides adaptation to TEC(obs), but the obtained profiles do not always ensure the coincidence with the plasma frequency, measured at the low orbit satellites. As shown in (Maltseva, 2014), adapting the model to the plasma frequency of one or two satellites gives values TEC(sat), not equal to TEC(obs). In this paper, the difference Δ(TEC(sat)) = TEC(obs)-TEC(sat) was attributed to a plasmaspheric part of the profile and introduced coefficient K(PL), modifying this part of the profile so that the TEC of the modified profile is equal to TEC(obs). This factor is a multiplier, which the density of the profile at the upper limit (20,000 km) is multiplied by. Examples of profiles of the station Juliusruh in April 2001 have been shown that if the difference Δ(TEC(sat)) is positive then the coefficient K(PL) may be greater or less than 1. When K(PL)> 1 the profile may have nonphysical shape. When K(PL)<0 it is necessary to suppose density on the upper limit equal to zero, since it cannot be negative. TEC for such a profile may differ from the TEC(obs). This indicates a need for testing a plasmaspheric part of the profile. To do this, in this paper the model of (Ozhogin et al., 2012), based on the experimental data RPI (Radio Plasma Imager) of satellite IMAGE, was used. This model was developed in a range of L-shells from 1.6 to 4, so this test cannot be carried out for the station Juliusruh, as its L-shell does not fall in this range. Example of test is given for the station Ascension Island.

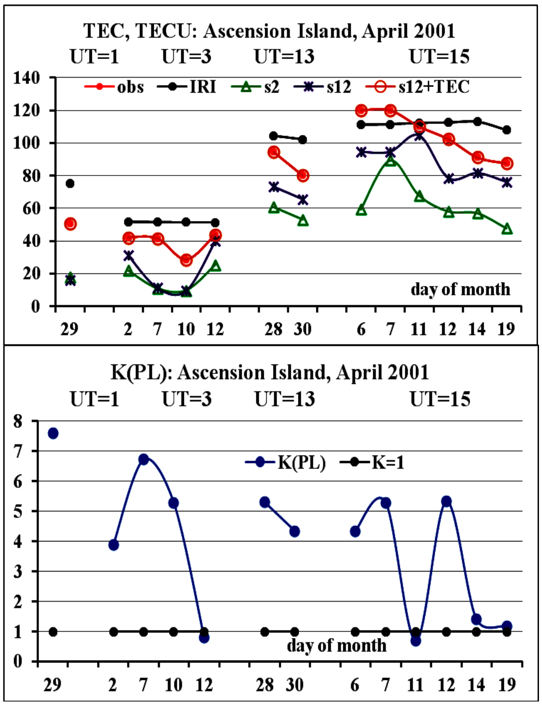

Figure 9: Comparison of values of TEC for various options of calculation and corresponding coefficients K(PL).

This example makes it possible to show the uncertainty of profiles associated with both the shape of the profile, and with the uncertainty of value of TEC(obs), given by various global maps. Comparison of TEC for that station for the calculation of various options is given in Fig. 9 during moments of flight of the satellite CHAMP over the station. The horizontal axis represents the days of the month. Time of observation is shown at the top of the graph. Red dots on the top panels show the observed values, the black dots are the values for the original (without adaptation) model IRI. Green triangles belong to profiles of models adapted to foF2(obs) and fne(sat2), violet asterisks indicate the TEC for the profile, which is the starting point for modifications. Orange circles show the TEC for profile modified by the coefficient K(PL). The lower panel shows the corresponding coefficients K(PL).

This situation is typical for all the stations, only values of the difference between the TEC may vary depending on the latitude. Interestingly, the N(h)-profiles passing through fne(sat), give more lower values TEC(sat) than TEC(obs). The modification results in full compliance with the TEC(obs) using K(PL) from the lower panel. Fig. 10 shows the N(h)-profiles for one of the cases (April 12, UT = 3). Profile is represented for clarity in two parts: from the beginning of the ionosphere to the height of 2,000 km and from an altitude 2,000 km to an altitude of 20,000 km.

Fig. 10 includes the profile of the original model (black dots), for which the TEC(IRI) exceeds the TEC(obs), the profile calculated for a model adapted to foF2(obs) and TEC(obs) and shown by purple diamonds (mark "All"), profiles passing through fne(CHAMP) and foF2(obs). These profiles cannot always build. They are shown with blue crosses, and there are only in two panels. Profiles going through foF2(obs) and fne(DMSP) are shown by green triangles. They are built for all cases. Large red dots show fne(sat) on the top panels. Small red dots show the values of the model RPI on the bottom panels. A plasmaspheric part of all profiles is close to the model of RPI. Profile s12+PL, providing TEC(JPL)=44 TECU, in the upper panel, does not pass through the plasma frequency. The second panel shows the profiles for UPC map with TEC(UPC) = 31.8 TECU. In the profile "All", the topside part was decreased, but the plasmaspheric part was increased. It can be seen that the profile s12 + PL passes through fne(DMSP), but again there is no passage through fne(CHAMP).

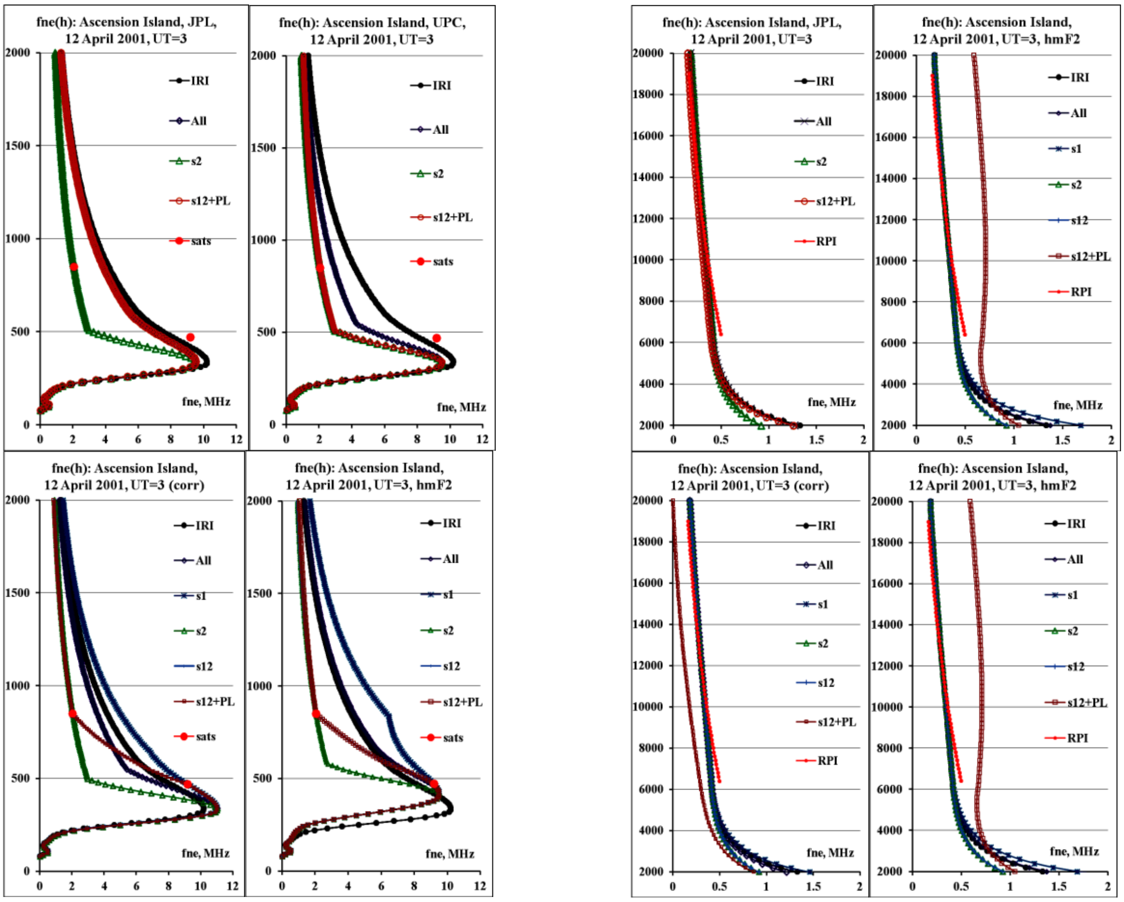

Figure 10: An illustration of conformity between model N(h)-profiles and observed TEC.

We attempted to modify the profile in two directions: to increase foF2(obs) (third panel) or hmF2 (fourth panel) in order to pass the profile through fne(CHAMP). In both cases, the profile has gone through fne(CHAMP), but the TEC of these profiles are greatly increased. In case of increase of foF2(obs), difference Δ(TEC(sat)) was negative, resulting in a zero value at the upper limit of profile. In the second case, the topside part was not changed much. This led to a positive difference and an increase of the plasmaspheric density. These results lead to the important conclusion that to disambiguate N(h)-profile it is necessary to specify more precisely the values of TEC and shape of the topside profile because the plasmaspheric part is close to the existing model.

.

5 INTERNET POSSIBILITIES OF USAGE OF EMPIRICAL MODELS OF THE IONOSPHERE

Presence of huge array of data and approximation methods has led to possibility of creating of empirical models of the ionosphere which with success are used for forecasting of its state. Among set of empirical models which are disposed on the Internet, the IRI model is one of the most demanded and most dynamically developing (http://irimodel.org). In this century it has undergone some updates: IRI2001, IRI2007, 2010, 2012 (Bilitza, 2001; Bilitza, Reinisch, 2008; Bilitza et al., 2014).

Main attention focuses on the correspondence of model parameters to experimental values. In this study, a comparison of the model with observed data was held including data of stations which have appeared on the Internet only recently (since 2013). It allowed us to obtain results in a global scale (Fig. 5, 7).

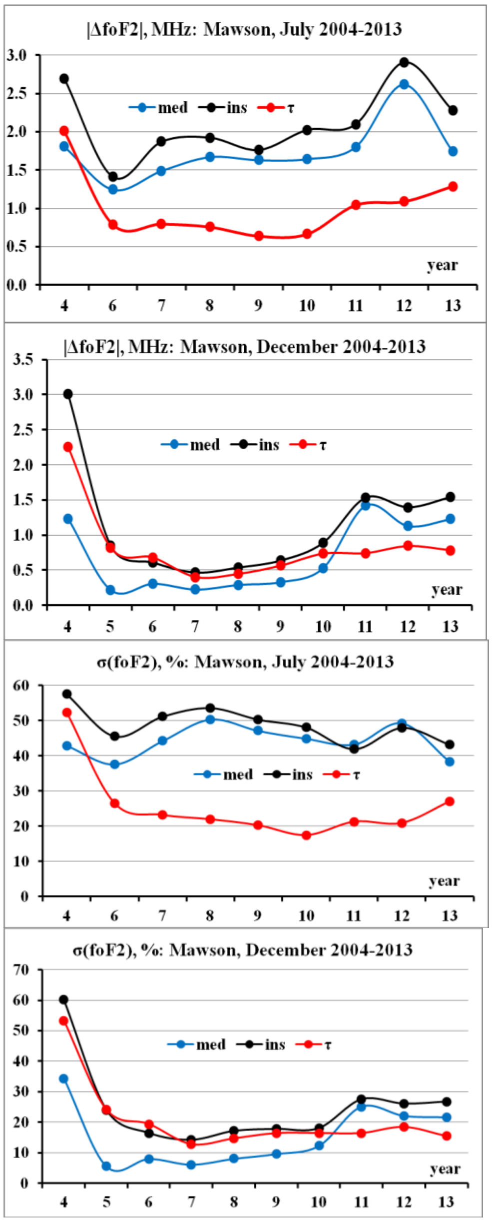

Figure 11: An estimate of accuracy of the model IRI according to data of the Antarctic station Mawson.

Of great interest is evaluation of the IRI model possibilities in the high latitudes of the southern hemisphere in connection with the project GRAPE (De Franceschi, Candidi, 2012). In this work, the first results were obtained according to data of the station Mawson. Fig. 11 shows the absolute and relative deviations of July and December for all the years with available data. Black dots show deviations from day to day, blue dots – deviations of medians and red dots – deviations of foF2 values, calculated with usage of the TEC and τ(med).

It is evident that in July (winter) deviations of curves with marks "med" and "ins" are significantly exceed the values typical for the northern hemisphere. Using the TEC and τ(med) improves the situation, except for 2004. Moreover, this Figure does not include the values for 2005 in connection with abnormal values of foF2, shown in Fig. 12 requiring additional study (the top panel). One reason may be an insufficient quantity of data: the lower panel shows the deviation |ΔfoF2| along with the number of days reduced by 10 times.

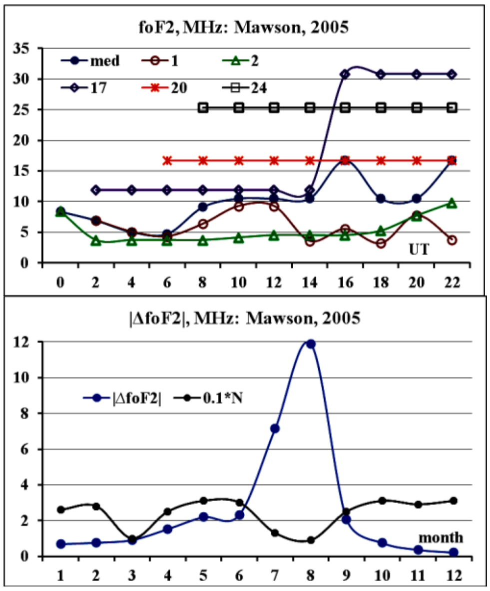

Figure 12: An illustration of the abnormal behavior of foF2 at the station Mawson in 2005.

In December results are comparable with the values typical for the northern hemisphere.

In (Maltseva et al., 2012; 2013; Maltseva, 2014) a lot of attention was paid to the comparison of different empirical models of foF2 and TEC (Gulyaeva, 2011; Hoque, Jakowski, 2011; Jakowski et al., 2011). In addition to these results, a huge volume of data has been used in (Maltseva et al., 2014; 2015) for comparison, covering more than 10-15 stations in different regions of the globe and several years. Despite this, an unambiguous answer was not obtained.

Great expectations are linked to a database of satellites COSMIC. The COSMIC EDP data are obtained from the COSMIC Data Analysis and Archive Center (CDAAC) at University Corporation for Atmospheric Research (UCAR). COSMIC RO measurements and products (such as electron density profiles) can be available from the Taiwan Analysis Center for COSMIC (TACC, http:// tacc.cwb.gov.tw/en/) and the COSMIC Data Analysis and Archive Center (CDAAC, http://www.cosmic.ucar.edu/cdaac/).

6 CONCLUSIONS

Only a few examples illustrate the possibility of obtaining new results using the resources of the Internet. For ionospheric studies, the most important are databases of the vertical sounding, measurements of TEC by high-orbit navigation satellites, measurements of plasma frequency by low-orbit satellites, data of solar and geomagnetic conditions. The obtained results show the large possibilities of use of resources of the Internet for study of the ionosphere and the representation of conditions of propagation of radio-waves. A series of new knowledge in this area was obtained.

But there is another side to such use: (1) it is not always possible to reconcile different data to each other because of their inaccuracies and ambiguities, (2) there is a danger in the access closing of certain countries, as it may be in an emergency situation.

ACKNOWLEDGEMENTS

The author thanks the scientists who provided the data of SPIDR, satellites DMSP, global maps of TEC, operation and modification of the IRI model, Dr A. Karpachev (IZMIRAN, karp@izmiran.ru) for CHAMP data, Southern Federal University for finantial support (project N 213.01-11/2014-22).

REFERENCES

Bilitza, D., 2001. International Reference Ionosphere, *Radio Sci*., *36(2), 261-275.*

Bilitza, D., Reinisch, B.W., 2008. International Reference Ionosphere 2007: Improvements and New Parameters. *Adv. Space Res., 42, 599-609*.

Bilitza, D., Altadill, D., Zhang, Y., Mertens, C., Truhlik, V., Richards, P., McKinnell, L.-A., Reinisch, B., 2014. The International Reference Ionosphere 2012 – a model of international collaboration, *J. Space Weather Space Clim*., *4, A07, 12p. DOI: 10.1051/swsc/2014004.*

De Franceschi, G., Candidi, M., 2013. GRAPE, GNSS Research and Application for Polar Environment, Expert Group of SCAR, *Annals of Geophysics*, *Special Issue 56. 2. P0215*; *Edited by Giorgiana, doi:10.4401/ag-6366.*

Fuller-Rowell, T., Araujo-Pradere, E., Minter, C., Codrescu, M., Spencer, P., Robertson, D., Jacobson, A.R., 2006. A new data assimilation product from the Space Environment Center characterizing the ionospheric total electron content using real-time GPS data US-TEC, *Radio Sci*. *41*, *RS6003, doi: 10.1029/2005RS003393.*

Gulyaeva, T.L., 2011. Storm time behavior of topside scale height inferred from the ionosphere-plasmosphere model driven by the F2 layer peak and GPS-TEC observations. *Adv. Space Res., 47, 913-920.*

Hoque, M.M., Jakowski, N., 2011. A new global empirical NmF2 model for operational use in radio systems. *Radio Sci., 46, RS6015, 1-13.*

Houminer, Z., Soicher H., 1996. Improved short –term predictions of foF2 using GPS time delay measurements. *Radio Sci., 31(5), 1099-1108.*

Jakowski, N., Hoque, M.M., Mayer, C., 2011. A new global TEC model for estimating transionospheric radio wave propagation errors. *Journal of Geodesy, 85(12), 965-974.*

Khattatov, B., Murphy, M., Gnedin, M., Sheffel, J., Adams, J., Cruickshank, B., Yudin, V., Fuller-Rowell, T., Retterer, J. 2005. Ionospheric nowcasting via assimilation of GPS measurements of ionospheric electron content in a global physics-based time-dependent model. *Q.J.R.Meteorol. Soc*., *131, 3543–3559*.

Maltseva, O., 2014. Choice of the definition method for the total electron content to describe the conditions in the ionosphere. *Proceedings of Third International Conference on Telecommunications and Remote Sensing, Luxemburg, Grand Duchy of Luxemburg, 26-27 June 2014, 51-61.*

Maltseva, O.A., Mozhaeva, N.S., Nikitenko, T.V., 2014. Validation of the Neustrelitz Global Model according to the low latitude ionosphere. *Adv. Space Res.,54, 463–472, http://dx.doi.org/10.1016/j.asr.2013.11.005.*

Maltseva, O.A., Mozhaeva, N.S., Nikitenko, T.V., 2015. Comparative analysis of two new empirical models IRI-Plas and NGM (the Neustrelitz Global Model). *Adv. Space Res.,55, 2086–2098*.

Maltseva, O., Mozhaeva N., Vinnik E., 2013. Validation of two new empirical ionospheric models IRI-Plas and NGM describing conditions of radio wave propagation in space. *Proceedings of Second International Conference on Telecommunications and Remote Sensing, Noordwijkerhout, The Netherlands, 11-12 July, 109-118*.

Maltseva, O., Mozhaeva, N., Zhbankov, G., 2012. A new model of the International Reference Ionosphere IRI for Telecommunication and Navigation Systems. *Proceedings of the First International Conference on Telecommunication and Remote Sensing, 129-138.*

Maltseva, O.A., Poltavsk, O.S., Schlyupkin, A.S., 2007. The IRI model residual difference and the new method of N(h)-profile determination. *Acta Geophysica, 55, 441-458*.

McNamara, L.F., 1985. The use of total electron density measurements to validate empirical models of the ionosphere. *Adv. Space Res., 5(7), 81-90.*

McNamara, L.F., 2006. Quality figures and error bars for autoscaled Digisonde vertical incidence ionograms. *Radio Sci., 41, RS4011, doi:10.1029/2005RS003444.*

Ozhogin, P., Tu, J., Song, P., Reinisch, B.W., 2012. Field-aligned distribution of the plasmaspheric electron density: An empirical model derived from the IMAGE RPI measurements, *J. Geophys. Res., 117, A06225*.

Reinisch, B.W., Galkin, I.A., Khmyrov, G.M., Kozlov, A.V., Bibl, K., Lisysyan, I.A., Cheney, G.P., Huang, X., Kitrosser, D.F., Paznukhov, V.V., Luo, Y., Jones, W., Stelmash, S., Hamel, R., Grochmal, J., 2009. The New Digisonde for Research and Monitoring Applications. *Radio Sci.,RS004115*.

URSI Atlantic Radio Science Conference, 18-22 May 2015, ExpoMeloneras, Grand Canaria, http://www.at-rasc.com

Yao, Y., Tang, J., Chen, P., Zhang, S., Chen, J., 2014. An Improved Iterative Algorithm for 3-D Ionospheric Tomography Reconstruction. *IEEE Transactions on Geoscience and Remote sensing, 52, 8*.

Zabotin, N.A., Wright, J.W., Bullett, T.W., Zabotina, L.Ye., 2005. Dynasonde 21: principles of data processing, transmission and web service. *11 International Ionospheric Effects Symposium, Alexandria, USA, 3-5 May 2005, 7B51-7B58.*

Cognitive Radio Management Benefiting from Flexible Reconfiguration

Xiguang Wu, Jacques Palicot and Pierre Leray
CentraleSupélec / IETR
Avenue de la Boulaie CS 47601, F-35576 Cesson-Sevigné CEDEX, France
{xiguang.wu, jacques.palicot, pierre.leray}@centralesupelec.fr

Keywords: Cognitive Radio, Dynamic Partial Reconfiguration, FPGA, Zynq, Management.

Abstract: Cognitive Radio (CR) has the ability to adapt its behavior to the changing environment. In order to achieve this goal in real practice, a CR equipment should be able to manage and reconfigure itself flexibly and efficiently. Therefore, it is necessary to integrate management in the CR equipment. Benefiting from the dynamic full and partial reconfiguration, Zynq based devices provide more flexible features and become a preferable hardware platform for CR. In this paper, we briefly introduce the cognitive radio management on a Zynq based platform, and study the benefit and cost of offloading computations from software to hardware. The results show that it is possible to win both performance and power consumption by flexible reconfiguration.

1 INTRODUCTION

As the digital communication systems evolve from GSM and now toward 5G, the supported standards are also growing. The desired communication equipments are required to support different standards in a single device at the same time. The software defined radio (SDR) (Mitola, 1995) has been considered a solution to this requirement, because typical way to change the function of a communication equipment needs to redesign the hardware. By using the SDR, the function of the communication equipment can be changed by software reprogramming without modifying the hardware.

Furthermore, Cognitive Radio (CR) (Mitola, 2000) has been proposed to integrate new capabilities, which can automatically adapt its behavior to the changing environment. But more is expected, in order to efficiently manage the CR features, we introduce a management architecture, namely Hierarchical and Distributed Cognitive Radio Architecture Management (HDCRAM) (Godard, Moy and Palicot, 2006), which has been proposed for CR management by our team.

In order to design CR equipments, flexible and efficiently reconfigurable hardware platforms are necessary. Zynq based platform becomes a favorable choice because it integrates ARM processor and FPGA in a single device (Xilinx, Inc., 2013, UG585), which provides both flexibility and performance. Especially, it enables dynamic full and Partial Reconfiguration (PR) (Xilinx, Inc., 2010, UG702). Benefiting from these features, Zynq based device is more suitable for developing CR equipment.

CR has also been considered as an enabling technology for green radio communications (Palicot, 2009). In this paper, we investigate the performance and power consumption taking the software and hardware FIR filter as a study case on Zynq based ZC702 evaluation board. The results provide the useful information for the management of the FIR filter, and with the HDCRAM, it is possible to benefit both performance and power consumption from flexible full and partial reconfiguration.

2 HDCRAM ARCHITECTURE

2.1 Introduction

A radio equipment consists of a set of functional components that are connected with each other, illustrated as processing elements (PEs) at the bottom of Figure 1. These PEs can either be software or hardware elements.

Our team has proposed a management architecture for Cognitive Radio in a previous work, namely HDCRAM (Palicot, 2013). A diagram of HDCRAM architecture featuring three levels is depicted in Figure 1.

HDCRAM consists of two aspects: Cognitive Radio Management (CRM) and Reconfiguration Management (ReM). The CRM part is responsible for gathering sensing information and making decisions according to the information obtained from PEs. The ReM part is in charge of taking actions to reconfigure the system. It takes a bottom-up approach in the CRM part. Sensing information is submitted from the lower level to the upper level. When a CRM unit has made a reconfiguration decision, it sends the reconfiguration parameters to its associated ReM unit at the same level. It uses a top-down method in the ReM part. The reconfiguration commands are sent from the upper level to the lower level.

HDCRAM is composed of three hierarchy levels. At level 1, only one cognitive radio manager and one reconfiguration manager can exist, because this is the top level. At level 2 and level 3, there are multiple couples of Cognitive Radio Management units (CRMu) and their associated Reconfiguration Management units (ReMu).

The architecture featuring three levels is sufficient. The level 1 manages the exchange of different standards; the level 2 manages the reconfiguration of the middle scale functions; and the level 3 manages the PEs.

According to this hierarchical management, a cognitive cycle can be on three different scales: 1) a local small cycle, in which the sensing, decision making, and reconfiguration action are processed, only includes the PE and its associated level 3 management; 2) a medium cycle that involves multiple PEs and a level 2 management, the reconfiguration of a PE needs the cooperation with other PEs; 3) or a large cycle that concerns all the three levels of management.

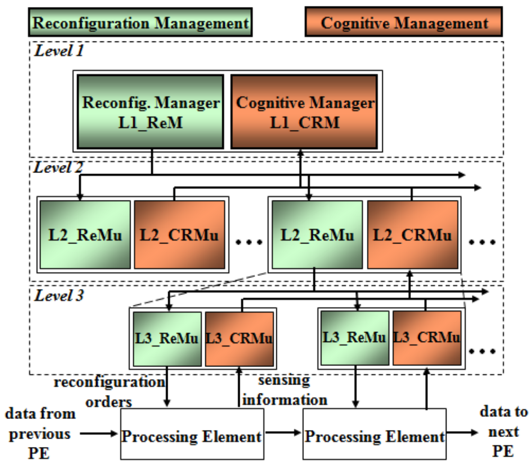

Figure 1: An example of HDCRAM.

2.2 Deployment Example

There may be many different choices to deploy HDCRAM. In this section, we only take one possible HDCRAM deployment method as an example, to introduce the deployment of HDCRAM, as shown in Figure 2.

It comprises a GPP, a DSP, a FPGA, and a Zynq based device. A straightforward way is placing the level 1 manager on the GPP, and maybe multiple level 2 and level 3 management units on it. A level 2 management unit and multiple level 3 management units are deployed on DSP, FPGA, as well as Zynq. An embedded processing core Microblaze is employed on the FPGA with the level 2 management unit on it. A PE may be hardware in logic or software on Microblaze. Therefore, a level 3 management unit that is in charge of managing a PE may also be hardware or software, or part of it is software executed on Microblaze and another part is hardware. On Zynq, similar to the PFGA, a level 2 management unit is on PS, and a PE may also either be hardware on programmable logic (PL) or software on processing system (PS). A level 3 management unit may also be hardware or software, or part of it is software executed on PS and another part is hardware on PL.

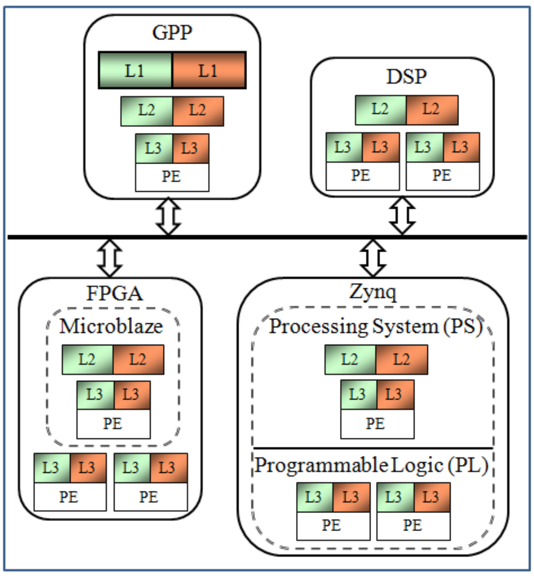

Figure 2: An example of HDCRAM deployment.

3 DYNAMIC PARTIAL RECONFIGURATION

3.1 Introduction

FPGA devices have provided the flexibility to do on-site device reprogramming, but a drawback of traditional FPGA is that it has to stop running and reprogram the entire logic even if a very small part of the logic needs to be updated. Recently, some FPGA families have provided a Dynamic Partial Reconfiguration (DPR) technique, which extends the inherent flexibility of the traditional FPGA. DPR allows designers to change the functionality of specific regions in an operating FPGA by dynamically downloading a partial configuration bitstream while the remaining logic continues to operate without interruption.

The logic in the FPGA design is divided into two different types, reconfigurable logic and static logic. Reconfigurable logic is any logical element that is part of a reconfigurable region. These logical elements are modified when a partial bitstream is loaded. Static logic is any logical element that is not part of a reconfigurable region. These logical elements are never partially reconfigured and always active when a partial bitstream is loaded (Xilinx, Inc., 2010, UG702).

As shown in Figure 3, the block portion labeled Reconfigurable Region represents reconfigurable logic and the light gray area of the FPGA block represents static logic. The function implemented in Reconfigurable Region is modified by downloading one of several available partial BIN files, PR1.bin, PR2.bin, PRn.bin, etc.

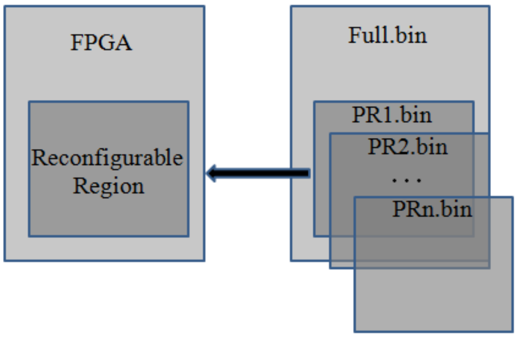

Figure 3: Reconfigurable logic and static logic.

There are many reasons why the DPR is advantageous over traditional full configuration.

- Flexibility. The functionality of part of the FPGA can be updated at any time by locally or remotely loading the partial bitstream that is needed on the fly, which makes the hardware software-like.
- Reduce reconfiguration time. Because a partial bitstream is smaller than the full bitstream, and the configuration time is proportional to the size of the bitstream, the reconfiguration time of DPR is shorter. Especially when the partial bitstream is quite small, compared with the reconfiguration of the entire device, DPR can significantly reduce the reconfiguration time, which is quite useful to applications requiring strict timing constraints.
- Improve performance. Only a portion of the device is reconfigured, the static logic remains functioning and is completely unaffected by the loading of a partial BIN file. There is no need to stop running and reprogram the entire device, therefore, it does not affect the performance of the rest of the device.
- Hardware sharing. DPR can realize the hardware reuse, which enables different functionalities to be implemented in the same portion of the device.
- Save space and resources. By taking advantage of the DPR, the same system can be implemented in smaller devices featuring less resource thus reducing the size of the FPGA.
- Consequently, reduce power consumption and overall cost.

3.2 Full and Partial Reconfiguration on Zynq

The ZC702 evaluation board utilizes a Xilinx Zynq-7000 All Programmable SoC (AP SoC), which

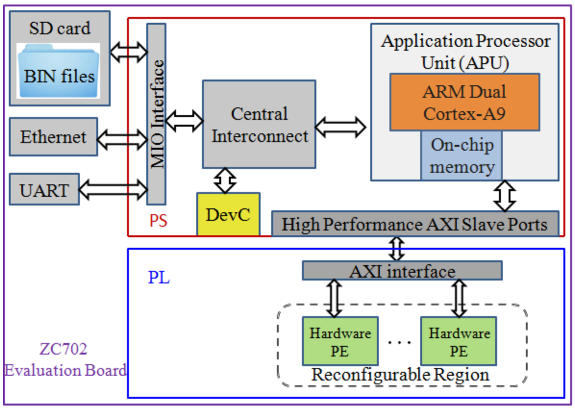

Figure 4: A simplified architecture of the ZC702 evaluation board.

integrates a dual-core ARM Cortex-A9 as the processing system (PS) and a Xilinx's 7 series FPGA Artix-7 as the programmable logic (PL) in a single device (Xilinx, Inc., 2013, UG850).

On Zynq, there are two ways for DPR to reconfigure the PL, i.e., either by the internal configuration access port (ICAP) primitive in the PL, or through the device configuration (DevC) / processor configuration access port (PCAP) interface in the PS (Christian Kohn, 2013).

ICAP can only perform partial reconfiguration in the PL, but PCAP supports both full and partial reconfiguration of the PL from the PS, which provides more flexibilities. Furthermore, the bitstreams are transferred to the PCAP interface by a Direct Memory Access (DMA) approach, which frees the processor to execute other tasks. Therefore, we utilize the PCAP method.

Different functions can be designed to share the hardware PL by dynamic full and partial reconfiguration in the field. The generated full and partial bitstreams can be stored in a database. Each function has a full bitstream and several partial bitstreams depending on the real needs. Figure 5 illustrates the storage organization of the BIN files database.

The reconfiguration bitstreams are stored in the database on the host computer. The full or partial bitstreams can be remotely downloaded through Ethernet to change the functionality of the complete or pre-defined regions of PL on the fly as needed. They can also be stored on the SD card on the ZC702 evaluation board if the level 2 management works standalone. It is also possible to dynamically download new full and partial bitstreams through Ethernet to update the database. Some partial bitstreams are able to be read into the on-chip memory in PS if they are frequently used.

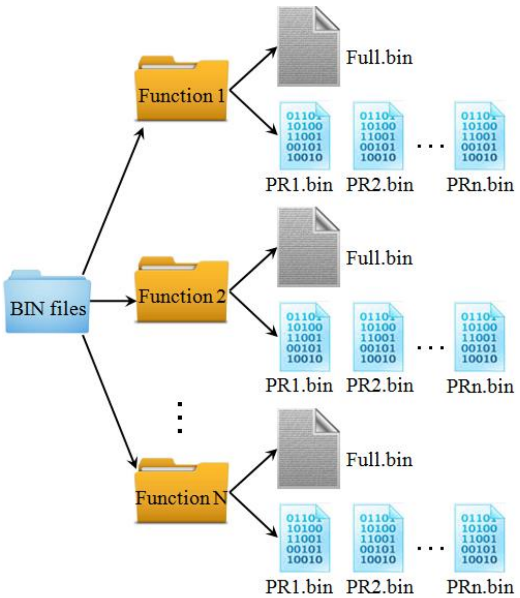

Figure 5: The storage organization of the reconfiguration bitstreams.

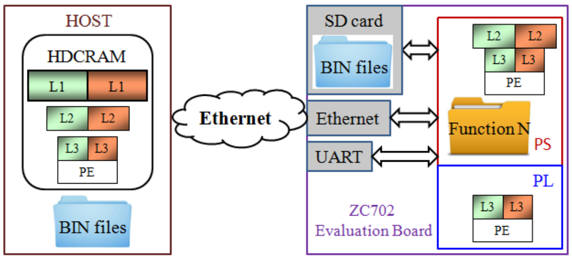

Figure 6: The HDCRAM implementation on the ZC702 evaluation board.

4 CASE STUDY

A finite impulse response (FIR) filter is a commonly used processing element in digital signal processing. It could be implemented either in software mapped onto the PS or in hardware mapped onto PL. Therefore, we would like to investigate the benefit and cost of the FIR filter implementation on PS and PL respectively, and then the results will provide helpful information for CRMu to make an appropriate decision.

4.1 Evaluation of performance and power consumption of FIR filter implementations

Take a 32-tap FIR filter as an example, which is implemented on PS and on PL respectively. The operations are executed in serial on PS, but on PL, the FIR filter could be implemented in serial or in parallel.

And the hardware serial and parallel implementations of the FIR filter reuse the PL logic by taking advantage of the PR. After generating the full and partial bitstreams for the PL following the PR design flow, we store them in the database on the host as shown in Figure 7. A blank full bitstream is also generated to clear the PL to save power if the PL part is not needed, which is stored in NOPL folder. Table 1 shows the resource available in the reconfigurable region and used by the FIR filter. The serial implementation consumes less resource, and it uses 2 DSP48E1s, which is 32 times less than the parallel implementation. But the serial way consumes more memory than the parallel approach.

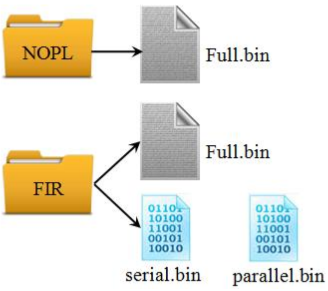

Figure 7: The full and partial bitstreams of the design.

Table 1: Resources available and used by the FIR filter.

Resource	Available	Serial	Parallel
LUT	10304	868	1096
FD_LD	20608	1516	3108
SLICEL	1564	141	288
SLICEM	1012	91	187
DSP48E1	72	2	64
RAMBFIFO36E1	36	8	4

The timing overhead of full and partial reconfiguration should also be considered. Because downloading a bitstream remotely from the host computer consumes longer time than that from the local memory, if we can benefit from remote reconfiguration, undoubtedly we can also benefit from local reconfiguration. The sizes of full and partial bitstreams, and the time consumed of remote full and partial configuration are listed in Table 2.

Table 2: Full and partial configuration time.

Type	Size (bytes)	Time (μs)
Full	4 045 564	215 736
Partial	707 712	51 865

We have also measured the power consumption of both PS and PL. The most convenient and simplest way to monitor the power consumption on ZC702 board is to use Texas Instruments' (TI) Fusion Digital Power Designer, which is a Graphical User Interface (GUI) used to monitor and display the real-time voltage and current of selected power rails of the board (Xilinx, Inc., 2013, UG850). Table 3 lists the power consumption of PL for blank design and the FIR filter.

In order to clearly and visibly observe the results, we have sent amount of data to the implemented software and hardware FIR filter. Each time we sent

Table 3: Power consumption of PL.

Function	NOPL	Serial	Parallel
Power(W)	0.06	0.095	0.101

4096 32-bit integers and then repeat 2000 times. When the hardware approach is chosen, the data are transferred between PS and PL by DMA approach. Table 4 gives the total time consumed by software and hardware implementations of the FIR filter.

Table 4: Execution time of the FIR filter.

Software (μs)	Hardware (μs)	
	Serial	Parallel
12 229 279	281 315	279 026

We can see that although the hardware approaches consume much less time than the software way, the hardware parallel implementation is not as fast as expected more than 32 times faster than the serial implementation, which is because the overhead of data transmission between PS and PL. It takes some time when the data and commands are transmitted from user space to Linux driver and then to the hardware. Therefore, if only offloading the FIR filter from the PS onto the PL, it is better to choose the serial implementation, which occupies less resource and consumes less power while not losing much performance.

The reason why we repeat 2000 times is that we cannot catch the power changes by TI Fusion Digital Power Designer when the execution time is too short. And even so, sometimes we still cannot catch PR and hardware FIR filter operations. For the sake of comparison and analysis, we put the operations of software FIR filter, PR, and hardware FIR filter together in Figure 8. At time 41:00, the software FIR filter are started execution, at around 41:25 PR is performed to reconfigure the PL, and at time 41:36, the hardware FIR filter operations are executed. The power risings at around 41:25 and at 41:36 are because the data transmission from PS to PL. We can see that the power increases from 0.33W to 0.44W during software FIR filter operations, which lasts about 12.23s. But the additional power increase

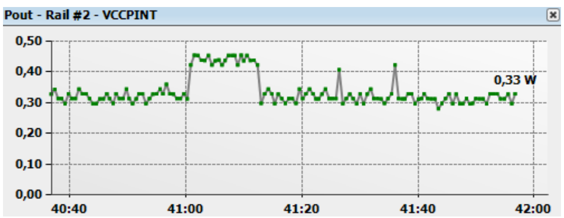

Figure 8: Power consumption of PS.

of the hardware serial and parallel implementations is around 0.04W on PL, which is less than 0.11W on PS.

4.2 Management of FIR filter by HDCRAM

Based on the above results, it is possible to benefit both performance and power consumption by offloading the FIR filter from the PS onto the PL. Another advantage is that it frees the PS to execute other tasks.

Therefore, we choose to implement the level 3 management of the FIR filter on the PS. The L2_CRMu makes the decision to implement the FIR filter on PS or on PL in serial or in parallel based on the information obtained from other L3_CRMus. And then the L2_ReMu sends the corresponding reconfiguration command to the L3_ReMu of the FIR filter, who then maps the FIR filter onto PS by calling the software FIR filter function or onto PL by dynamic full or partial reconfiguration.

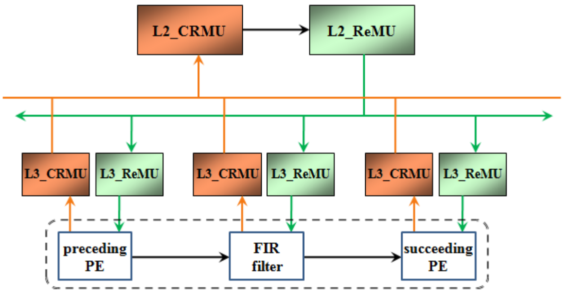

Figure 9: Management of FIR filter.

If the PL is occupied by other computation intensive PEs and has no more space for the FIR filter, there is no choice and the L2_CRMu decides to implement the FIR filter in software on PS, which consumes 0.11W more power and has a longer execution time.

Else if the preceding PE and the succeeding PE of the FIR filter is implemented on PS, the L2_CRMu decides to implement the FIR filter on PL in serial mode, because it uses less resource with additional 0.035W power consumption and the performance is close to the parallel way (see Table 4) due to the overhead of data transmission between PS and PL.

Else if the preceding PE or the succeeding PE of the FIR filter is implemented on PL, the L2_CRMu decides to implement the FIR filter on PL in parallel mode, because the speed is more than 32 times faster than the serial way and the data transmission is in hardware, which does not slow down the data processing. This way consumes 0.041W more power but has a higher performance.

5 CONCLUSIONS

In order to efficiently manage the CR features, it is necessary to integrate management into CR equipment. In this paper, we have briefly introduced the HDCRAM architecture as well as partial reconfiguration on Zynq. We have studied the commonly used FIR filter and the benefit and cost when it is implemented in PS and PL. The results show that we can win both performance and power consumption by flexible full and partial reconfiguration, which also provide useful information for the HDCRAM to make appropriate decisions to efficiently manage the FIR filter implementation. This also provides a reference to the implementation of potential green scenarios, which are our future works.

REFERENCES

Mitola, J. (1995). The software radio architecture. *Communications Magazine, IEEE, 33*(5), 26-38.

Mitola, J. (2000). Cognitive Radio---An Integrated Agent Architecture for Software Defined Radio.

Godard, L., Moy, C., & Palicot, J. (2006, June). From a configuration management to a cognitive radio management of SDR systems. In *Cognitive Radio Oriented Wireless Networks and Communications, 2006. 1st International Conference on* (pp. 1-5). IEEE.

Xilinx, Inc. (2013, September). Zynq-7000 all programmable SoC. Technical Reference Manual, UG585 (v1.6.1),.

Xilinx, Inc. (2010, April). Partial Reconfiguration User Guide. UG702 (v12.3).

Palicot, J. (2009, June). Cognitive radio: an enabling technology for the green radio communications concept. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly* (pp. 489-494). ACM.

Palicot, J. (Ed.). (2013). *Radio engineering: From software radio to cognitive radio*. John Wiley & Sons.

Xilinx, Inc., (2013, April). ZC702 Evaluation Board for the Zynq-7000 XC7Z020 All Programmable SoC User Guide, UG850 (v1.2).

Christian Kohn, (2013). Partial Reconfiguration of a Hardware Accelerator on Zynq-7000 All Programmable SoC Devices. XAPP1159 (v1.0).

Analysis and Design of Hybrid and Graphene-Based Plasmonic Waveguide Components

D. A. Ketzaki, V. Salonikios, I. Demirtzioglou and T. V. Yioultsis
Dept. of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece
traianos@auth.gr

Keywords: Graphene, plasmonics, plasmonic waveguides, hybrid waveguides.

Abstract: We present an efficient finite element formulation for the eigenmode analysis of graphene-based plasmonic waveguides with switching functionalities. The formulation is full-vectorial and addresses graphene as a surface conductivity, as opposed to bulky material considerations, thus eliminating the need for fine discretizations inside thin graphene models. Based on this technique, several graphene-enhanced plasmonic waveguides and components with promising characteristics are proposed.

1 INTRODUCTION

Graphene is a relatively new innovative material, with interesting new physics and several significant properties and effects, including the ability to support surface plasmon propagating modes and switching functionalities (Bludov et al., 2013). Optical conductivity of graphene has been shown to consist of a Drude intraband term and an interband contribution. These properties may result in either plasmonic modes in THz (Nikitin et al., 2011) or enhanced switching in photonic waveguides for the optical communications regime (Sun et al., 2014). In particular, for the case of the THz regime, where the Drude term is dominant, graphene surface plasmons offer the possibility of waveguiding with strong confinement, while in the optical communications spectrum where the interband contribution is substantial, the tunability of graphene's conductivity through electrostatic gating shows great potential for the design of switching components.

As for the analysis and design of graphene-based and enhanced waveguides and components, the finite element method (FEM) is a perfect candidate, due to its ability to deal with problems of considerable geometric complexity. However, the general trend is often to approach graphene as a bulky material, thus requiring very fine descritizations inside thin sheets and the surrounding space as well. We present here an efficient formulation for both the eigenmode and the 3D analysis of graphene-based plasmonic waveguides and components with switching functionalities which is full-vectorial and addresses graphene as a thin sheet with a surface conductivity. Based on this analysis, we propose a graphene-enhanced plasmonic CGS waveguide with an extinction ratio of 8.6 dB and a 2.15 dB insertion loss for a 10 um length, which can be considered highly satisfactory. A particular investigation of a three-dimensional microring /microdisk filter revelas the possibility of actual designs with extinction ratios exceeding 10 dB, which is also a firm basis for further study towards the development of switched plasmonic components in the photonics regime.

2 FINITE ELEMENT FORMULATION

The proposed finite element eigenmode formulation follows the general framework that has been proposed in (Selleri et al., 2001), where the electric field is used as a working variable. The formulation uses mixed finite elements for the discretization of the waveguide cross section, with tangentially continuous (H-curl) vector finite elements in the transverse plane and scalar (nodal) finite elements for the axial component. Using the Galerkin formulation for the Helmholtz equation, the form

$$\iint_S \mathbf{E}' \cdot \left(\nabla \times \overline{\overline{\mu_r}}^{-1} \nabla \times \mathbf{E} - k_0^2 \overline{\overline{\varepsilon_r}} \mathbf{E} \right) ds = 0 \qquad (1)$$

expresses the projected problem, reducing its solution to a finite-dimensional vector subspace. The electric field can be written in the form $\mathbf{E} = \left(\mathbf{E}_t + \hat{\mathbf{z}}E_z\right)e^{-\gamma z}$, where $\mathbf{E}_t = \mathbf{E}_t(x,y)$ represents the transverse component and $E_z = E_z(x,y)$ represents the axial component. The adjoint field $\mathbf{E}' = \left(\mathbf{E}_t' - \hat{\mathbf{z}}E_z'\right)e^{\gamma z}$ is selected as the test function in the Galerkin equation and the final eigenmode formulation expressed as a function of the effective refractive index $n_{eff} = -j\gamma / k_0$ is as follows:

$$\begin{aligned}
&\iint_S \nabla\times\mathbf{E}_t' \cdot \mu_{r,zz}{}^{-1}\nabla\times\mathbf{E}_t ds \\
&-\iint_S \left(\nabla E_z' \times\hat{\mathbf{z}}\right)\cdot \overline{\overline{\mu_{r,tt}}}^{-1}\left(\nabla E_z\times\hat{\mathbf{z}}\right)ds \\
&-k_0{}^2\iint_S \mathbf{E}_t' \cdot \overline{\overline{\varepsilon_{r,tt}}}\mathbf{E}_t ds + k_0{}^2\iint_S E_z'\cdot\varepsilon_{r,zz}E_z ds \\
&+n_{eff}\begin{bmatrix} -jk_0\iint_S \left(\mathbf{E}_t'\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}}}^{-1}\left(\nabla E_z\times\hat{\mathbf{z}}\right)ds \\ -jk_0\iint_S\left(\nabla E_z'\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}}}^{-1}\left(\mathbf{E}_t\times\hat{\mathbf{z}}\right)ds\end{bmatrix} \\
&+n_{eff}{}^2\left[k_0{}^2\iint_S\left(\mathbf{E}_t'\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}}}^{-1}\left(\mathbf{E}_t\times\hat{\mathbf{z}}\right)ds\right] \\
&-\oint_{\partial S}\mathbf{E}'\cdot\hat{\mathbf{n}}\times\nabla\times\mathbf{E}ds = 0
\end{aligned} \qquad (2)$$

The domain is terminated by perfectly matched layers. Following the discretization of the 2D-space, using basis functions and the degrees of freedom (nodal or edge-based, according to the field component) for the electric field quantities and assuming for the moment that the line integral vanishes, (2) leads to the quadratic eigenvalue problem expanded form of Galerkin formulation,

$$\left\{\begin{aligned}&\begin{bmatrix}\mathbf{S}^t - k_0{}^2\mathbf{T}^t & 0 \\ 0 & -\mathbf{S}^{z,m}+k_0{}^2\mathbf{T}^z\end{bmatrix} \\ &+n_{eff}\begin{bmatrix}0 & -jk_0\mathbf{P} \\ -jk_0\mathbf{Q} & 0\end{bmatrix} \\ &+n_{eff}{}^2\begin{bmatrix}k_0{}^2\mathbf{T}^{t,m} & 0 \\ 0 & 0\end{bmatrix}\end{aligned}\right\}\begin{bmatrix}\mathbf{E}_t \\ \mathbf{E}_z\end{bmatrix} = 0 \qquad (3)$$

Where

$$\begin{aligned}
&S_{ij}^t = \iint_{S_n}\nabla\times\mathbf{w}_i\cdot\mu_{r,zz}^{-1}\nabla\times\mathbf{w}_j ds, \\
&S_{ij}^{z,m} = \iint_{S_n}\left(\nabla N_i\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}^{-1}}}\left(\nabla N_j\times\hat{\mathbf{z}}\right)ds, \\
&T_{ij}^t = \iint_{S_n}\mathbf{w}_i\cdot\overline{\overline{\varepsilon_{r,tt}}}\mathbf{w}_j ds,\ T_{ij}^z = \iint_{S_n}N_i\cdot\varepsilon_{r,zz}N_j ds, \\
&T_{ij}^{z,m} = \iint_{S_n}\left(\mathbf{w}_i\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}^{-1}}}\left(\mathbf{w}_j\times\hat{\mathbf{z}}\right)ds, \\
&P_{ij} = \iint_{S_n}\left(\mathbf{w}_i\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}^{-1}}}\left(\nabla N_j\times\hat{\mathbf{z}}\right)ds, \\
&Q_{ij} = \iint_{S_n}\left(\nabla N_i\times\hat{\mathbf{z}}\right)\cdot\overline{\overline{\mu_{r,tt}^{-1}}}\left(\mathbf{w}_j\times\hat{\mathbf{z}}\right)ds
\end{aligned} \qquad (4)$$

To solve the quadratic eigenvalue problem we use first companion linearization to reduce it to

$$\begin{aligned}&\begin{bmatrix}0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{I} \\ -\mathbf{S}^t + k_0{}^2\mathbf{T}^t & 0 & 0 & jk_0\mathbf{P} \\ 0 & \mathbf{S}^{z,m} - k_0{}^2\mathbf{T}^z & jk_0\mathbf{Q} & 0\end{bmatrix}\begin{bmatrix}\mathbf{E}_t \\ \mathbf{E}_z \\ n_{eff}\mathbf{E}_t \\ n_{eff}\mathbf{E}_z\end{bmatrix} \\ &= n_{eff}\begin{bmatrix}\mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & k_0{}^2\mathbf{T}^{t,m} & 0 \\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{E}_t \\ \mathbf{E}_z \\ n_{eff}\mathbf{E}_t \\ n_{eff}\mathbf{E}_z\end{bmatrix}\end{aligned} \qquad (5)$$

which is a sparse form with a positive semidefinite matrix at the right hand side, suitable for sparse eigensolvers.

As far as the graphene implementation is concerned, its extremely small thickness (one-atom thick) dictates its consideration as an ideal two-dimensional surface with a corresponding surface conductivity σ_g (measured in S). Therefore, any graphene surfaces in the waveguide eigenmode analysis are basically represented by one-dimensional lines in the 2D cross-section of the structures (Figure 1).

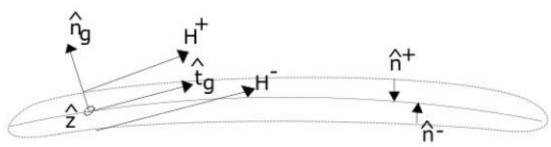

Figure 1: Representation of graphene as an infinitely thin sheet (2D cross section) and its surrounding surface.

A first route to incorporate graphene sheets in FEM simulations is to start from a bulky material approach and consider the limit of its thickness to zero. In this case, graphene's contribution would be apparent through its conductivity, thus affecting the 3rd and 4th term of (2) which include permittivity quantities. In these integrals, separating a finite surface of thickness δ corresponding to the bulky graphene area, we derive the additional terms

$$I_g = -k_0^2 \iint_{S_1} \mathbf{E}_t' \cdot \overline{\overline{\varepsilon_{r,tt}}} \mathbf{E}_t ds + k_0^2 \iint_{S_1} E_z' \cdot \varepsilon_{r,zz} E_z ds \quad (6)$$

where we replace the permittivity with the complex permittivity of graphene $\varepsilon_r^* = \varepsilon_r - j\sigma_b / \omega\varepsilon_0$, and σ_b describes an equivalent conductivity of bulk graphene (in S/m). Assuming that σ_b consists of non-zero real and imaginary parts, we can omit ε_r as being included in σ_b and have $\varepsilon_r^* = -j\sigma_b / \omega\varepsilon_0$. Therefore (6) becomes

$$I_g = jk_0\eta_0 \iint_{S_1} \mathbf{E}_t' \cdot \sigma_b \mathbf{E}_t ds - jk_0\eta_0 \iint_{S_1} E_z' \cdot \sigma_b E_z ds \quad (7)$$

Assuming an infinitesimal graphene thickness, variations are negligible in this dimension, giving

$$I_g = jk_0\eta_0 \int_{\partial S} \mathbf{E}_{t,p}' \cdot \sigma_g \mathbf{E}_{t,p} dl - jk_0\eta_0 \int_{\partial S} E_z' \cdot \sigma_g E_z dl \quad (8)$$

where $\delta\sigma_b$ equals the surface conductivity σ_g (in S). It has to be particularly noted that the electric field component $\mathbf{E}_{t,p}$ involved in the first integral is not the full transverse component but only its tangential projection on the one-dimensional graphene line.

An equivalent and more elegant way to arrive at the same expression is to consider graphene as a zero thickness sheet in the first place. In this case, the line integral term in (2) cannot be ignored, as the graphene current sheet introduces a discontinuity in the magnetic field, thus affecting the line integral term. In particular, the interface condition on the graphene sheet is written in the form

$$\begin{aligned} &-\frac{j}{k_0\eta_0}\left(\hat{\mathbf{n}}_g \times \nabla \times \mathbf{E}^+ - \hat{\mathbf{n}}_g \times \nabla \times \mathbf{E}^-\right) \\ &= \mathbf{J}_s = \sigma_g\left(\mathbf{E}_{t,p} + \hat{\mathbf{z}}E_z\right) \end{aligned} \quad (9)$$

where $\hat{\mathbf{n}}_g$ is the unit vector normal to the graphene sheet. However, to substitute (9) in the line integral term of (2) we need to consider a fictitious surface that surrounds the graphene sheet from both sides and being infinitely close to it. Therefore, the line integral is split into two parts, one for the upper surface, where the outward-pointing unit normal vector is $\hat{\mathbf{n}}^+ = \hat{\mathbf{n}}_g$ and one for the lower one, where $\hat{\mathbf{n}}^- = -\hat{\mathbf{n}}_g$ and the line integral takes the form

$$\begin{aligned} I_g &= \int_{\partial S^+} \mathbf{E}' \cdot \hat{\mathbf{n}}^+ \times \nabla \times \mathbf{E}^+ ds + \int_{\partial S^-} \mathbf{E}' \cdot \hat{\mathbf{n}}^- \times \nabla \times \mathbf{E}^- ds \\ &= \int_{\partial S^-} \mathbf{E}' \cdot \hat{\mathbf{n}}_g \times \left(\nabla \times \mathbf{E}^+ - \nabla \times \mathbf{E}^-\right) ds \\ &= jk_0\eta_0 \int_{\partial S^-} \mathbf{E}' \cdot \sigma_g \mathbf{E}_p ds \end{aligned} \quad (10)$$

which easily results in (9) as well.

Therefore, graphene's contribution can be implemented by adding two line integral terms in the initial formulation, expressed as

$$T_e^{t,g} = \int_C \mathbf{E}_{t,p}' \cdot \sigma_g \mathbf{E}_{t,p} dl, \, T_e^{z,g} = \int_C E_z' \cdot \sigma_g E_z dl \quad (11)$$

and by considering the corresponding matrices, it results in the linear eigenvalue problem similar to (5), where the term $-jk_0\eta_0\mathbf{T}^{t,g}$ is added to $-\mathbf{S}^t + k_0^2\mathbf{T}^t$ and $jk_0\eta_0\mathbf{T}^{z,g}$ is added to $\mathbf{S}^{z,m} - k_0^2\mathbf{T}^z$.

The three-dimensional FEM formulation is similarly based on a standard Galerkin formulation with vector finite elements in three dimensions, starting from the 3D version of (2). Since there is no split in transverse and axial components, the Galerkin formulation will include only 3D forms of the first, third and last terms in (2). Following similar principles, a graphene-raleted term of the form $jk_0\eta_0\mathbf{T}^g$ will be added to the standard FEM stiffness-mass matrix $\mathbf{S} - k_0^2\mathbf{T}$.

3 PLASMONIC AND SWITCHING COMPONENTS

The proposed formulations are is able to analyze both plasmon graphene ribbon waveguides in the THz regime and a switching-capable waveguide structure for telecom applications, based on the CGS waveguide (Dai and He, 2009), which is properly enhanced by graphene.

3.1 Graphene waveguide with high-index dielectric ridge

The first structure simulated was the graphene waveguide proposed in (Sun et al., 2014). Its concept takes advantage of a high-index dielectric ridge to achieve strong field confinement without using a finite width graphene ribbon which is harder to fabricate. Placing a wide graphene sheet over a dielectric ridge of appropriate size, the geometry of the structure facilitates waveguiding. The relation of the complex effective refractive index to the thickness of gap the graphene sheet and the dielectric ridge, as well as to the chemical potential are shown in Fig. 3, being in very good agreement with (Sun et al., 2014).

3.2 Graphene microribbon waveguide

To fully test the functionality of our formulation, we analyzed a plasmon graphene microribbon waveguide in the THz regime (Nikitin et al., 2011). This is a waveguiding structure for frequencies between 1 and 12 THz (as opposed to the telecommunications wavelength regime) taking advantage of the surface conductivity of a graphene microribbon. The analysis was conducted for a ribbon width of 5 um and the electric field intensity plots for the two transverse components are shown in Figure 2.

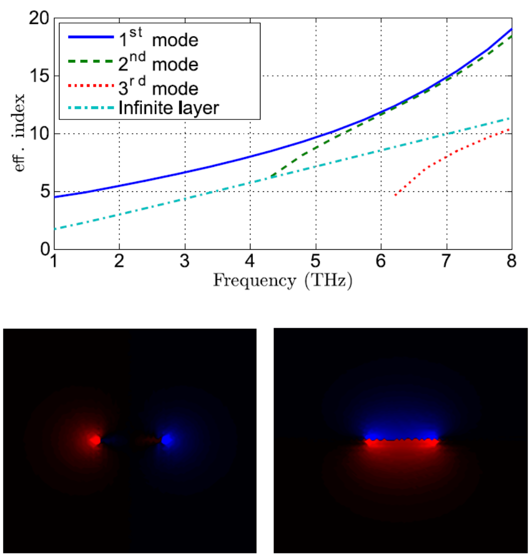

Figure 2: Effective refractive index (real part) and mode profiles for the first mode of a 5um graphene ribbon waveguide (vertical and horizontal E-field component, respectively).

3.3 Graphene waveguide with high-index dielectric ridge

The next structure simulated was the graphene waveguide proposed in (Sun et al., 2014). Its concept takes advantage of a high-index dielectric ridge to achieve strong field confinement without using a finite width graphene ribbon which is harder to fabricate. Placing a wide graphene sheet over a dielectric ridge of appropriate size, the geometry of the structure facilitates waveguiding. The relation of the complex effective refractive index to the thickness of gap of the graphene sheet and the dielectric ridge, as well as to the chemical potential are shown in Figure 3, being in very good agreement with (Sun et al., 2014).

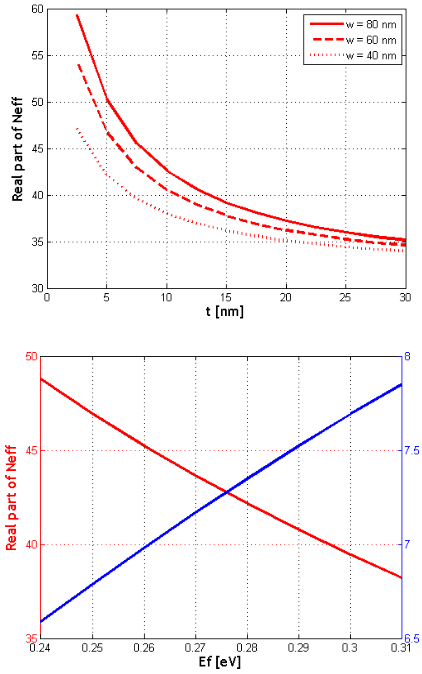

Figure 3: Real part of the effective refractive index with respect to gap thickness and chemical potential

3.4 Graphene switching component

Based on the analysis, we propose a switching capability for the classic plasmonic CGS waveguide (Dai and He, 2009), (Wu et al., 2010) by adding graphene layers on all interfaces between waveguide materials, including both sides of the oxide layer, and also the two vertical ridges of the waveguide. The ON and OFF states of the waveguide correspond to chemical potential values 1 eV and 0.1 eV. Selecting a structure length of 10 um, the insertion loss can be as low as 2.15 dB, almost entirely due to metal (not graphene) losses and the achieved extinction ratio is 8.6 dB, which is highly promising for further study.

3.5 Switched plasmonic components with frequency selective finctionalities

Following the proposed waveguide components we further proceed to the design of three dimensional components with frequency tuning functionalities, like the waveguide-coupled microring structure of Figure 4. For more enhanced performance, the microring concept is extended to cover the cases of a micordisk or a donut-shaped ring. The component's performance, in terms of the power transmission coefficient is shown in Figure 5, for both the microdosk and donut shapes, for the ON state, while the achieved extinction ratio between ON and OFF states for the graphene enhanced component exceeds 11.5 dB, which is a figure suitable for practical considerations.

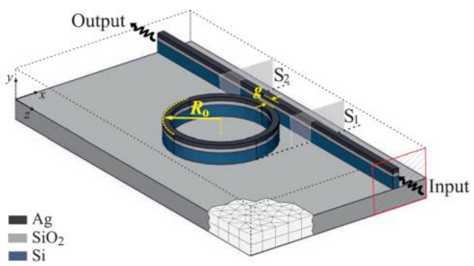

Figure 4: Microring resonator filter based on the CGS waveguide

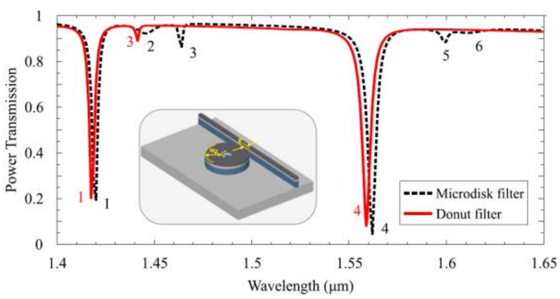

Figure 5: Frequency response of the microdisk and donut structures, in terms of the power transmission coefficient. Outer ring radius is 0.85 um, inner ring radius for the donut structure is 0.45 um and the gap is set to 150 nm.

4 CONCLUSIONS

We have presented a full-vectorial finite element formulation for the eigenmode analysis of graphene-based plasmonic waveguides and components with switching functionalities. The formulation addresses graphene as a surface conductivity thus eliminating the need for fine discretizations inside thin graphene sheets. Finally, several plasmonic or switched components with promising characteristics have been proposed.

REFERENCES

Bludov, Y.V., Ferreira, A., Peres N. M. R. and Vasilevskiy M. I., "A Primer on Surface Plasmon-Polaritons in Graphene," *Int. J. Mod. Phys*. B, 27, 1341001, 2013.

Nikitin, A. Y., Guinea, F. and García-Vidal, F. J, "Edge and waveguide terahertz surface plasmon modes in graphene microribbon," *Phys. Rev. B*, 84, 161407(R), 2011.

Sun, Y., Zheng, Z., Cheng, J. and Liu J., "Graphene surface plasmon waveguides incorporating high-index dielectric ridges for single mode transmission," *Optics Communications*, 328, pp. 124–128, 2014.

Selleri, S., Vincetti, L., Cucinotta, A. and Zoboli, M. "Complex FEM modal solver of optical waveguides with PML boundary conditions," *Optical and Quantum Eletronics*, vol. 33, no. 4-5, pp. 359-371, 2001.

Dai, D. and He, S., "A silicon-based hybrid plasmonic waveguide with a metal cap for a nano-scale light confinement", *Optics Express*, vol. 17, no. 19, 2009.

Wu, M., Han, Z. and Van, V., "Conductor-gap-silicon plasmonic waveguides and passive components at subwavelength scale," *Optics Express*, vol. 18, no. 11, pp. 11728-11736, 2010.

On the Homogeneous Transmission Power under the SINR Model

Evangelos Spyrou and Dimitris Mitrakos
School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Egnatia Street, Thessaloniki, Greece
evang_spyrou@eng.auth.gr, mitrakos@eng.auth.gr

Keywords: Transmission Power, Packet Reception Ratio, Signal-to-Interference-and-Noise Ratio, homogeneous, Clear Channel Assessment.

Abstract: Power control is quite important in the field of wireless sensor networks. Many works adjust transmission power in order to either achieve significant improvement on packet reception or to save energy. Even though the use of non-homogeneous transmission power utilisation benefits is evident in the literature, we study cases where the use of homogeneous transmission powers across parts of the network may accomplish high Packet Reception Ratio. We show examples of the above and provide experimental results that show that reception of packets may be high in appropriate topologies or parts of the topology, with the use of the same transmission power level. We evaluate two topologies with and without the use of Clear Channel assessment to present our point.

1 INTRODUCTION

Wireless sensor networks (WSN) are going to the second decade research, as shown by Breza, Martins, McCann, Spyrou, Yadav and Yang (2010). The main reason behind this trend is the plethora of civil and military applications that require the gathering of data and their successful wireless transmission within a large terrain. WSNs consist of small wireless devices that measure physical phenomena such as temperature, pressure, humidity, or the position of objects.

The distributed nature of WSNs, where devices exchange information with others within their transmission range, can be quite useful. However, this very advantage is the main drawback of WSNs. The nature of the radio employed by the devices is shared by all participants in the transmission range; hence, the issue of interference is addressed. It is intuitive that simultaneous transmission might result in drop of packets, since the communication medium suffers from interference. Such a phenomenon is evident with the use of graph-based models.

This has a direct impact on the network capacity and throughput. In order to face the problem of interference, we can address the issue of adjusting transmission powers. Successful adjustment of transmission powers result in a smaller set of interferers; hence, an increase of network throughput. It is imperative that we use the most appropriate interference model to attempt to tackle this problem. We utilize the physical SINR model by Gupta and Kumar (2000), where interference is continuous and decreasing polynomially with distance from the sending device. We will provide the reader with a formal description of the model at a later section of the paper.

Briefly outlining the model, the receivers successfully receives a message if the ratio of the signal strength of the sender and the sum of interference signals by devices, transmitting at the same time, is larger than the hardware-defined threshold. The denominator of the ratio also includes ambient noise. The speed that the signal fades depends on the variable called the path-loss exponent α, dictated by Rappaport (1996), which takes the value ranging from 2 - 6 according to environment of the transmission. The accumulative nature of interference provides a fruitful domain of research. Only recently have some theoretical guarantees been provided for SINR-based algorithms.

Power control is an important field in the field of wireless networks, since it can control the performance of the network. Furthermore, it may increase the number of receivers for a given sender, as well as tuning interference. However, power assignment has a significant impact on the complexity of the problems addressed by algorithms. In the literature, power assignment is distinguished between uniform and non-uniform settings. As implied by the two different approaches, uniform assignment sets the same transmission power to all nodes. On the other hand, in the non-uniform assignment, senders operate on different transmission powers offered by the communication medium.

Moscibroda, Wattenhofer, and Zollinger (2006) as well as Moscibroda, Wattenhofer, and Weber (2006), show that uniform power assignment exhibits performance disadvantages as opposed to a non-uniform one. However, cases where power control approaches outperform uniform power assignment schemes position the nodes in an area of exponential size in the number of nodes. These schemes require transmission power levels that differ by a factor exponential in the number of nodes, as shown by Avin, Lotker and Pignolet (2009). A uniform power control has a number of advantages due to it being simple.

Some of them include the lower cost of transmitting at the same transmission power. The simplicity of decision making implies lower cost, since devices do not need to decide their power level depending on factors, such as interference. In addition, Avin C., Emek Y., Kantor E., Lotker Z., Peleg D. and Roditty L., (2009), showed the convexity of reception zones of senders using a uniform scheme. This is not the case for the non-uniform scheme.

In this paper, we will build upon the results of Moscibroda, Wattenhoffer and Weber, (2006) that show that senders may transmit simultaneously and messages will be received without collision due to interference. We show that simultaneous transmission of messages is feasible even with the use of uniform transmission powers, which depends on the distance of the interfering nodes with the receivers.

2 RELATED WORK

There is a significant difference between the graph based and the SINR models. Early works investigate the SINR model based on the assumption of nodes being uniformly distributed in the plane, such as (Behzad and Rubin, 2003), (Grönkvist and Hansson, 2001). The complexity of these solutions, however, gave way to computationally efficient approaches, which provide guarantees that use SINR effects. These solutions include scheduling (Moscibroda and Wattenhofer, 2006) and topology control (Moscibroda, Wattenhofer, and Weber, 2006).

Since then, a plethora of research has been undertaken in scheduling (Calinescu and Tongngam, 2011), (T. Tonoyan, 2013), (Fan, Zhang, Feng, Zhang and Ren, 2012), (Halldórsson and Mitra, 2012), as well as topology control (Lou, Tan, Wang and Lau, 2012), (Bodlaender, Halldórsson and Mitra, 2013) under the SINR model.

In (Halldórsson, Holzer, Mitra, and Wattenhofer, 2013), the authors explicitly investigate the power of the non-uniform transmission power. On the other hand, in (Avin, Lotker, Pasquale, and Pignolet, 2009) valuable information is provided on the employment of uniform transmission power. This is close to our work, with the difference that we aim to show different cases of uniform transmission powers utilization under the SINR model. In (Whitehouse, Woo, Jiang, Polastre and Culler, 2005), the authors consider a scheme of collisions and not failures that make explicit the utilization of the capture.

Furthermore, we provide the reader with some early works regarding throughput increase in wireless networks. In (Biswas and Morris, 2005),, the authors propose a routing and MAC layer protocol, which aims to the maximization of throughput. Also, a scheme that surpasses graph-based models is suggested in (Katti, Rahul, Hu, Katabi, Medard, and Crowcroft, 2006).

3 PROMISSING EXAMPLES OF UNIFORM TRANSMISSION POWER

Initially we assume that the nodes are randomly distributed on a unit plane. Moscibroda, Wattenhoffer and Weber, (2006) showed that doubling the throughput is feasible when we employ non-uniform transmission powers in a 1-D setting. We consider a 2-D scenario where devices transmit with uniform transmission powers. We consider a network of devices, where a transmission from a device is successful if the receiver can decode the message. This occurs when $\frac{P}{I+N} \geq \beta$, where P is the signal strength, I is the sum of interferences from other devices and N is the ambient noise. Denote β as the hardware-dependent ratio.

Furthermore, under the physical model of propagation, the signal strength P is modeled as a polynomially decreasing function depending on distance between the sender-receiver pair of devices. Denote this as $d(x_s, x_r)$ and the aforementioned function as $\frac{1}{d(x_s,x_r)^\alpha}$ where α is the path-loss exponent ranging from 2 to 6 according to the setting of the network (e.g outdoor, indoor).

We assume that the path loss exponent α = 3, the SINR threshold β = 3 and the background noise N = 10nW. Note that the values above are reasonable for practical wireless sensor scenarios, as presented by Son, Krishnamachari, and Heidemann, (2006). We denote $\beta(x_i, x_j)$ as the SINR ratio of node x_i when node x_j is transmitting. Hence the power of node x_j is the signal and the powers of the other nodes transmitting simultaneously are considered as interference. Obviously, a transmission is successful if $\beta(x_i, x_j) \geq \beta$.

Consider the example that is given in figure 1. We observe that the distances between the transmitting and the interfering nodes are greater by approximately a factor of 2. This is that the interference distance is twice as great as the transmitting distance.

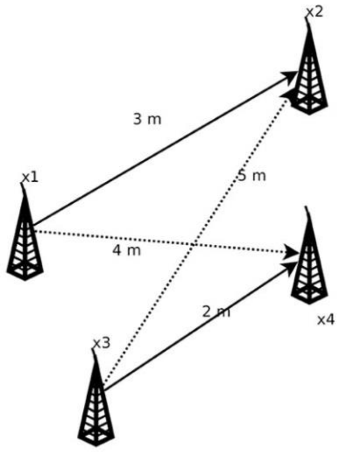

Figure 1: Nodes Transmitting Simultaneously with heterogeneous distances.

If we obtain the SINR values of the two transmission pairs we have the following:

$$\beta(x_3, x_4) = \frac{1000\mu W(2m)^3}{0.01\mu W + 1000\mu W/(4m)^3} \approx 8$$

and

$$\beta(x_1, x_2) = \frac{1000\mu W/(3m)^3}{0.01\mu W + 1000\mu W(5m)^3} \approx 4.5.$$

This shows that both messages transmitted go through in the case of simultaneous transmission. The values exceeding the SINR threshold hold even if nodes transmit with the minimum transmission power. Note that using any graph-based approach trying to send the two messages in parallel will fail because, intuitively, the medium between the two receivers can only be used once per time slot.

4 MULTIPLE NODE INTERFERENCE

The main issue with the utilization of the SINR model is the fact that it can get very complicated, constituting it intractable in terms of the protocol designer. In known network topologies, transmission power increase results in node degree increase, which implicitly means that the number of interferers increases as well. This may assist in the decrease of the packets decoded in the network; hence, a decrease in the PRR. Y. Gao, J. C. Hou, and H. Nguyen, (2008), p.3 introduce the term "interfering node", which is given by (1).

$$\frac{p_t(i)d_{i,j}^{-\alpha}}{N + p_t(k)d_{k,j}^{-\alpha}} < \beta \quad (1)$$

Where $d_{i,j}$ is the distance between the sender i and the receiver j. Also, $d_{k,j}$ is the distance between the interfering node k and the receiver j. This essentially provides the node, whose interference results in the packet to be dropped by the receiver. The authors also provide the term interference degree, which they show that it might not be minimized by using the minimal transmission power assignment.

Note that a node can be interfering with the transmission of packet and the packet may still be received. Hence, interference degree is the number of nodes that collide or interfere with a transmission that may result in a successful transmission or not.

Following the interference degree, it is useful to provide the reader with some notes on the potential number of interferers. Earlier in the paper, we assumed that the nodes are randomly distributed; hence, the number of interferers is a random variable. Nodes that are receiving in slot s-1 are transmitting in slot s, thus, interfering with nodes receiving in slot s. We refer the MAC layer slots as slots. We use the work of Vakil and Liang, (2006), p.4, to indicate that the number of interferers is given by (2)

$$N_{I_r}[s] = \sum_{i=1}^{N_p[s]} N_{I_r}^{i}[s] \quad (2)$$

where $N_{I_r}^{i}[s]$ is the number of nodes within the transmission range r of node i, which have been receiving in slot s and transmitting in the current slot. Also, $N_p[s]$ is the number of permissible sources – transmission ranges - in slot s. There is a difficulty in obtaining the number of interferers, since the calculation of inter-node distance is required and is quite difficult to obtain accurately. The parameter, which can be utilized in order to obtain the distance, is the Received Signal Strength Indicator (RSSI) value (Xu, Liu, Lang, Zhang and Wang, 2010), which may differ significantly from its actual value if the two nodes are not within Line-Of-Sight.

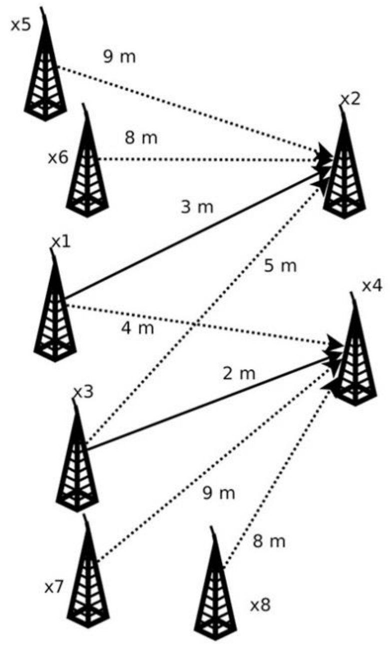

Figure 2: Multiple Nodes Transmitting Simultaneously with heterogeneous distances

We will continue our examples shown in the previous section of the paper, in order to show a case where the packets in the presence of multiple interfering nodes get decoded simultaneously. We have to mention that the intuitive action of the transmitting nodes in figure 2 is to reduce their transmission power in order to minimize the interfering nodes; hence, to increase the probability of decoding the packet successfully. However, the case we examine in the particular figure shows that even with a high transmission power, simultaneous reception of the packets is feasible, provided that the interfering nodes are at a quite larger distance than the transmitter-receiver pair. Note that all the nodes are transmitting with transmission power of 0dB. Specifically, the SINR ratio between nodes x1 and x2 with nodes x5, x6 interfering is

$$\beta(x_1, x_2) = \frac{1000\mu W/(3m)^3}{0.01\mu W + \frac{1000\mu W}{(4m)^3} + \frac{1000\mu W}{(8m)^3} + \frac{1000\mu W}{(9m)^3}} \approx 3.6,$$

which is higher than the SINR reception threshold. Similarly, for nodes x3, x4 when nodes x7 and x8 are interfering the SINR is

$$\beta(x_3, x_4) = \frac{1000\mu W/(3m)^3}{0.01\mu W + \frac{1000\mu W}{(4m)^3} + \frac{1000\mu W}{(8m)^3} + \frac{1000\mu W}{(9m)^3}} \approx 6.6.$$

Hence the packet is received successfully.

5 EXPERIMENTAL RESULTS

We decided to put some of our examples to a test reflecting some of the examples we carried out. We are considering a network of 15 nodes running for 30 minutes on the Indriya testbed (M. Doddavenkatappa, M. C. Chan, and A. L. Ananda, 2012). Note that there will be 108000 messages transmitted to all the nodes in the network, since every node is transmitting 4 packets per second. The devices in Indriya are employed with the Chipcon CC2420 radio, which uses the modulation and encoding specified by the IEEE802.15.4 standard.

We carefully selected two scenarios; one, where the nodes are connected but in a sparse manner and another that the nodes are in a dense area. We employed a form of synchronization, where the nodes transmit at the same time. Note that the transmission power, with which the nodes transmit is the maximum, 0 dB. The metric we utilize is the Packet Reception Ratio (PRR) and the number of successful receptions on the network. First, though, we provide the reader with the relationship between the SINR and the PRR.

SINR is the Signal-to-Interference-plus-Noise Ratio (SINR) of the transmission from node k to node j, which we denote as $\gamma_{k,j}$, which is given by

$$\gamma_{k,j} = \frac{h_{k,j} p_k}{\sum_{t \neq k, t \neq j} p_t h_{t,j} + N_0} \tag{3}$$

where $h_{t,j}$ is the channel gain between the interfering node t of node j. The Bit-Error-Rate (BER) for the CC2420 (Fu, Sha, Hackmann and Lu, 2012, p. 3), which we denote as ξ is

$$\xi_{k,j} = \frac{1}{2}(1 - \sqrt{\frac{\gamma_{k,j}}{1 + \gamma_{k,j}}}) \quad (4)$$

and finally, for any link $(k,j), k,j \in N$, $PRR_{k,j}$ denote as $P_{k,j}$ - can be expressed by

$$P_{k,j} = (1 - \xi_{k,j})^{l} \quad (5)$$

where l is the packet length in bits. As dictated in Zhao and Govindan (2003), at the physical layer, packet reception experiences variability by the existence of a grey area within the communication range of a node. Receiving nodes in this grey area are susceptible to unstable packet reception. Furthermore, the grey area is almost a third of the communication range in certain environments. The grey area also exhibits temporal packet reception variation.

Physical layer coding schemes exist capable of masking some of the variability of packet reception. The 802.15.4 standard uses a 32:4 DSSS chip-to-bit encoding. Because the CC2420 uses soft chip decision, there's no real concept of a "bit error." Instead, it effectively calculates the closeness of each chip to 0 and 1 and then chooses the symbol sequence which is closest to the soft chip decisions.

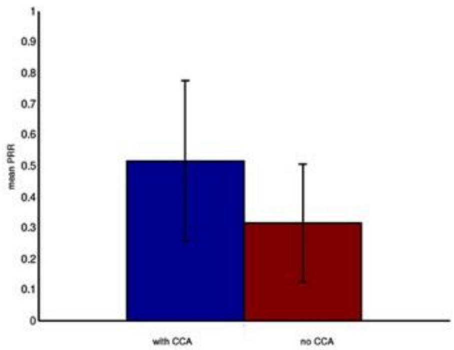

Figure 3: Mean PRR with and without CCA of dense topology

In most scenarios a high SINR means high PRR. We performed two experiments; in the first setting the nodes use the Clear Channel Assessment (CCA), in order to sense the channel before they proceed to a transmission of a packet. In the case that the channel is busy, the node performs an exponential back-off and attempts to transmit again. In the second scenario, using the same configuration of nodes, CCA is being disabled. Our intuition is that we will find a difference in the performance of the two settings in the sparse case. As for the dense configuration, we believe that the CCA enabled setting will outperform the CCA-disabled one. This is due to the examples in the previous section, that the nearest interferer will block the transmission.

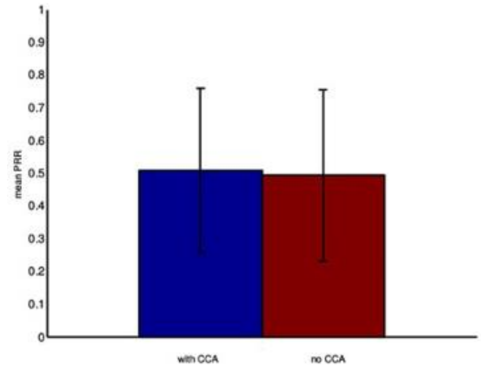

Figure 4: Mean PRR with and without CCA of sparse topology

In figure 3, we observe that for the dense scenario, disabling the CCA significantly affects performance. In fact, the difference between the CCA- enabled and CCA-disabled is approximately 20%. Furthermore, we investigated the number of the received messages received by all nodes in the network configuration. Our findings show that the CCA-enabled setting achieved 15,3% more messages received than the CCA-disabled setting. This is natural since, the messages transmitted by the CCA-disabled network are being dropped, since the channel is not sensed first, due to the density of the configuration.

Thereafter, we studied the performance of the same two settings for a sparser configuration, where the interferers have a greater distance from the receivers. In figure 4 we can see the mean PRR obtained for both settings. We note that the mean PRR is similar, which does not give us enough information on which settings accomplishes the best performance.

Table 1: Received messages of sparse topology with/without CCA

Configuration	Received Messages
With CCA	54864
No CCA	53244

Table 1 provides information of the received messages across the network. We observe that the number of the received messages of the configuration with the CCA disabled reaches the number of messages of the CCA-enabled one. This is due to the fact that when CCA is disabled, provided the sparsity of the configuration, interferers and transmitters pass messages at the same time, without performing a backoff, which may result in packet drop. On the other hand, even in the sparse configuration, nodes sense the channel's state first; hence they delay in the transmission of their packets.

6 CONCLUSIONS

In this paper we showed that utilizing uniform transmission powers may result in increase of PRR. This is dependent on the distance between the receiver and the interferer. We studied two settings, one with CCA enabled and the other with CCA disabled. We have seen that in a sparse configuration, using the CCA-disabled setting results in the network reaching the quality of messages reception of the CCA-enabled setting; thus, exhibiting a similar PRR. On the other hand, in a dense configuration, the CCA-enabled setting outperforms the one where CCA is disabled.

The use of the aforementioned results implies the necessity of spatiotemporal optimization and stability of wireless sensor networks. That is, WSN power control optimization methods may employ the careful selection of receivers to indicate whether a network should use uniform or non-uniform transmission power settings in specific regions. Furthermore, depending on the network density as well as the network neighbor and interference degrees, the network protocol designers, may find that CCA is a holding back factor of the network throughput increase. This may be valid in outdoor topologies where the signal is not affected by factors, such as Wi-Fi devices (Wu, Stankovic, He and Lin, 2008).

At this point, we have to mention it would be interesting to experiment with nodes when distances are fixed, according to the examples discussed previously. Furthermore, since the Indriya testbed is spread across different rooms, another interesting experiment would be to test the topological configurations under Line-Of-Sight, where the path loss exponent does not fluctuate. These experiments may provide us with useful insight regarding the PRR and rate of transmission.

Finally, this approach may indicate the fact that interference may not be high enough to require lowering the transmission power level of a node, even if the transmission power used is high. This may give a helpful insight on the behavior of the network PRR in a two-hop neighborhood.

REFERENCES

Avin C., Lotker Z., Pasquale F. and Pignolet Y.-A., 2009. A note on uniform power connectivity in the sinr model. *In Algorithmic Aspects of Wireless Sensor Networks, pages 116-127. Springer.*

Avin C., Lotker Z., and Pignolet Y.-A., 2009. On the power of uniform power: Capacity of wireless networks with bounded resources. *In Algorithms-ESA 2009, pages 373-384, Springer.*

Avin C., Emek Y., Kantor E., Lotker Z., Peleg D. and Roditty L., 2009, Sinr diagrams: towards algorithmically usable sinr models of wireless networks. *In Proceedings of the 28th ACM symposium on Principles of distributed computing, pages 200-209. ACM.*

Behzad A. and Rubin I., 2003. On the performance of graph-based scheduling algorithms for packet radio networks. *In Global Telecommunications Conference, GLOBECOM'03, IEEE, volume 6, pages 3432-3436. IEEE.*

Biswas S. and Morris R., 2005. Exor: Opportunist6ic multihop routing for wireless networks. *In ACM SIGCOMM Computer Communication Review, volume 35, pages 133-144. ACM.*

Bodlaender M. H., Halldórsson M.M. and Mitra P., 2013. Connectivity and aggregation in multihop wireless networks. *In Proceedings of the 2013 ACM Symposium on Principles of distributed computing, pages 355-364. ACM.*

Breza M., Martins P., McCann J. A., Spyrou E., Yadav P., and Yang S., 2010. Simple solutions for the second decade of wireless sensor networking. *In Proceedings of the 2010 ACM-BCS Visions of Computer Science Conference, page 7. British Computer Society.*

Calinescu G. and Tongngam S., 2011. Interference-aware broadcast scheduling in wireless networks. *Ad Hoc Networks, 9(7):1069-1082.*

Doddavenkatappa M., Chan M. C., Ananda A. L., 2012. Indriya: A low-cost, 3-d wireless sensor network testbed. *In Testbeds and Research Infrastructure. Development of Networks and Communities, pages 302-316, Springer.*

Fan S., Zhang L., Feng W., Zhang W. and Ren Y, 2012. Optimisation-based design of wireless link scheduling with physical interference model. *Vehicular Technologies, IEEE Transactions on, 61(8):3705-3717.*

Fu Y., Sha M., Hackman G. and Lu C., 2012. Practical control of transmission power for wireless sensor networks. *In Network Protocols (ICNP), 2012. 20th IEEE International Conference on, pages 1-10. IEEE*

Gao Y., Hou J. C., Nguyen H., 2008. Topology control for maintaining network connectivity and maximizing network capacity under the physical model. *In INFOCOM 2008. The 27th Conference on Computer Communications, IEEE.*

Grönkvist, J., & Hansson, A. (2001, October). Comparison between graph-based and interference-based STDMA scheduling. In *Proceedings of the 2nd ACM* *international symposium on Mobile ad hoc networking & computing* (pp. 255-258). ACM.

Gupta P. and Kumar R.P., 2000, The capacity of wireless networks. *Information Theory, IEEE Transactions on, 46(2):388-404.*

Halldórsson, M. M., Holzer, S., Mitra, P., & Wattenhofer, R. (2013, January). The power of non-uniform wireless power. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms* (pp. 1595-1606). SIAM.

Halldórsson M. M. and Mitra P., 2012. Wireless capacity and admission control in cognitive radio. *In INFOCOM, Proceedings, IEEE, pages 855-863.*

Katti S., Rahul H., Hu W., Katabi D., Medard M. and Crowcroft J., 2006. Xors in the air: practical wireless network coding. *In ACM SIGCOMM Computer Communication Review, volume 36, pages 354-254. ACM.*

Lou T., Tan H., Wang Y. and Lau F. C., 2012.Minimizing average interference through topology control. *In Algorithms for Sensor Systems, pages 115-129, Springer.*

Moscibroda T., Wattenhoffer R. and Weber Y., 2006. Design Beyonf Graph-Based Models. *In 5th Workshop on Hot Topics in Networks (HotNets), Irvine California, USA.*

Moscibroda T. and Wattenhoffer R., 2006. The complexity of connectivity in wireless networks. *In INFOCOM.*

Moscibroda T., Wattenhoffer R. and Zollinger A., 2006. Topology control meets sinr: the scheduling complexity of arbitrary topologies. *In Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing, pages 310-321. ACM*

Rappaport., 1996, *Wireless Communications: principles and practice,* volume2., Prentice hall PTR New Jersey.

Son D., Krishnamachari B. and Heidemann J., 2006. Experimental analysis of concurrent packet transmissions in wireless sensor networks. *ACM SenSys, Boulder, USA*

Tonoyan T., 2013. Comparing Schedules in the sinr and conflict-graph models with different power schemes. *In Ad-hoc, Mobile and Wireless Network, pages 317-328, Springer*

Vakil S. and Liang B., 2006. Balancing cooperation and interference in wireless sensor networks. *In Sensor and Ad Hoc Communications and Networks, 2006, SECON'06. 3rd Annual IEEE Communications Society on, volume 1, pages 198-206, IEEE.*

Whitehouse K., Woo A., Jiang F., Polastre J. and Culler D., 2005. Exploiting the capture effect for collision detection and recovery*. In Proceedings of the 2nd IEEE workshop on Embedded Networked Sensors, pages 45-52.*

Wu Y., Stankovic J. A., He T. and Lin S., 2008. Realistic and efficient multi-channel communications in dense sensor networks. *In Proceedings of the 27th IEEE International Conference on Computer Communications (InfoCom '08).*

Xu, J., Liu, W., Lang, F., Zhang, Y., & Wang, C. (2010). Distance measurement model based on RSSI in WSN. *Wireless Sensor Network*, *2*(08), 606.

Zhao, J., & Govindan, R. (2003, November). Understanding packet delivery performance in dense wireless sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems* (pp. 1-13). ACM.

SHORT PAPERS

Wide-Range Tuner for Generators in THz Bands

Alexander Denisov[1,2], Jinghui Qiu[2], Shengchang Lan[2], Ming He[3], Yang Xiong[3], Fan Zhang[3], Han Zhang[3] and Francesco Soldovieri[4]

[1]*State Research Center "Iceberg", Kiev, Ukraine*

[2]*Department of Microwave Engineering, Harbin Institute of Technology, P.R.China*

[3]*College of Electronic Information and Optical Engineering, Nankai University, Tianjin300071, P.R.China*

[4]*Istituto per il Rilevamento Elettromagnetico dell'AmbienteConsiglio Nazionale delle Recerche ,Napoli, Italy*

alex_denisov2009@ukr.net, qiujh@hit.edu.cn, heming@nankai.edu.cn, soldovieri.f@irea.cnr

Keywords: Josephson junction, unknown power, frequency measuring, sensitivity, electromagnetic radiation, terahertz band.

Abstract: Before experiments can be conducted on new generation structures in the THz band when the level of power is small and especially the frequency are practically unknown correctly (or it has place some accidental character) it is necessary to have instruments capable of measuring weak signals in a wide-range frequency band. Receiving instruments for the above-described evaluations must be based better by one measuring block and not only in the limits of the intermediate frequency (IF) amplifier band. Best natural "candidate" for this task is Josephson junction (JJ) – it has super wide voltage electronic regulation on frequency by basic ratio $2eV= h\omega$**.** JJ can be used as the criterion for a single-block super-wide band frequency-meter and as a sensitive element for the tuning the generation structures. Short description presents the realized innovative idea about simple tuner for generators in THz bands.

1 INTRODUCTION

The idea for this work came after visiting several labs which had interest to realize electromagnetic radiation (EMR) using Josephson media (JM) (H. B. Wang, S. Guénon, etc., 2010) and practical requests for possible good but cheap receivers covering the short part of millimeter band with enough wide bandwidth of the intermediate frequency (IF) amplifier. These labs worked only with liquid nitrogen and oriented on the structures based on high temperature superconductor (HTSC) such as $Bi_2Sr_2CaCu_2O_{8+\delta}$ (A. Koshelev and L. Bulaevskii., 2008, K. Kadowaki, L. Ozyuzer etc, 2008, D. Crankshaw, Enrique Trias., 2001).

Finally, it is worth to mention that purchasing a reliable enough cryocooler based on closed Giffard-McMagon (GM) cycle working till the temperature of a liquid helium is not a big issue today (www.fullsharecryogenics.com).

At least the price of GM machine is not principally more than famous network analyzer "Agilent". If to talk about the perspective using for the above task the Josephson junctions in spite that it is necessary to use cooling from the point of possible alternative to "Agilent" or some the same - from the "the rules of contraries"- there absent other good alternative.

The primary trend in superconductive microwave radio-electronics is the using JJ through the development of the devices employing Josephson oscillation, based on non-stationary Josephson-effect (K. Likharev, B. Ul'rich -1978, I. Kulik, I. Yanson, 1972). This can be used for the development principally new devices which are difficult or impossible to realize by the other methods.

On the basis of the previous theoretical and experimental investigation (**A.**Denisov, V. Obolonsky, 1990, A.Denisov, V.Gorishnyak, 1981**,** A. Denisov, V.Gaevsky., 1985) there shortly described the innovation decision about wide-range tuner for generators in THz bands.

Why tuner? What is necessary for the tuning of generator? It is need some visual control about frequency and the power of EMR. In reality the calculation of new generation structure can be done according to some current theoretical model and naturally in the simplest approximation. According to this and the real possible mistakes the frequency of EMR has some accident character, so the receiver with standard mixer even with big band of the intermediate frequency is not enough for this task. In the practical case at the beginning of the experiments the generation power is small enough too. Real situation. Besides of this the microwave losses in the transmission line can be very big too according to practically full absence the devices for such control and the matching in THz band.

There raised the question - where is the best suitable decision?

For such task it is need the tuning device with super wide band and with best sensitivity. In this case it will be real guarantee we would not "sleep past cash-box".

This is proved by two devices based on the JJ. First one is the sensitive receiver based on self-pump mode regime of JJ and second one is the using JJ as the criterion according to 2eV = hω **(**for the frequency measurement by the super wide frequency-meter. Combination of this two devices in one single block was published before**)** (A.Denisov, Qiu Jing Hui, 2014, A. Denisov, A.Gudkov, 2015)

Perhaps, it's interesting to say the fact (A. Denisov , A.Gudkov., 2015) that in 1978 it was the first S&R works to create the practical devices of the generators based on Impatt structures in Scientific Research Institute "Saturn" in Kiev, was oriented on R&D of low and super low noise microwave receivers in former USSR. It was practically lack of measuring devices in frequency bands more than 100 GHz. Spectrum analyzer C4-28 works till~38,5 GHz, and the existing certificated frequency-meter and power-meter worked practically till the same limiting frequency. *It was* realized the first variant of the panoramic sweep generator with monitor for measuring Standing Wave Ratio (VSWR) had been working in 3 mm. It was a real task to construct the generator of 70 GHz. It was came up with the idea to use 'Shapiro steps" (K.Likharev, B.Ul'rich -1978) – their level and position along voltage-frequency scale for the tuning and matching the "room-temperature" Impatt generator to the biggest power. Experiments were successful, but not only for 70 GHz. we elaborated the second harmonic and tuned the Impatt generator on 140 GHz. However, since the certificated measuring equipments were lucking, it was really difficult to confirm the obtained results and to convince the corresponding technical standard control personnel. We observed the harmonics of 140 GHz till more than 1,5 THz with the help of our first analogue variant of the Josephson frequency-meter (A.Denisov, V.Obolonsky, 1990) built on Nb point contact had been working according to the "Shapiro steps" on Current-Voltage Characteristic (IVCh) of the JJ (patented in USSR in 1978).

First of all, this device had extremely super wide momentary wave band. It was not necessary to change the measuring block to modify the measuring frequency band.

2 TECHNICAL EXPLANATION

On fig.1 and fig, 2 presented typical Volt-Amper Characteristic (VACh) of the JJ. Result on fig.1 was made from Tl based HTSC film by one of the author of this work. EMR of the experimental generator falls on JJ which produces microwave radiation according to 2eV= hω and later visual control of the Shapiro steps on VACh can be used for the tuning and matching of the experimental generator. To obtain a quantitative measure of the microwave power sensed by the Josephson junction, the well known dependence **of** the Shapiro step width on incident microwave power can be employed.

In term of Josephson-effect theory, the reduction of zeroth Shapiro step width is equivalent in quantity to diminish the critical current I_c of JJ which depends on EMR power as zeroth Bessel function. This current which is the result of interference of two signals on JJ has direct influence on the width of the steps. So, the width of the steps or ΔI on Fig 1 is the measure of EMR falling on JJ**.**

And naturally the position of Shapiro steps on the voltage scale of VACh corresponds to the frequency which falls on JJ.

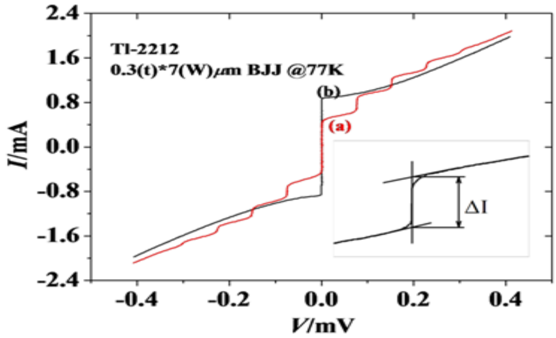

Figure 1: Typical VACh of Josephson junction.

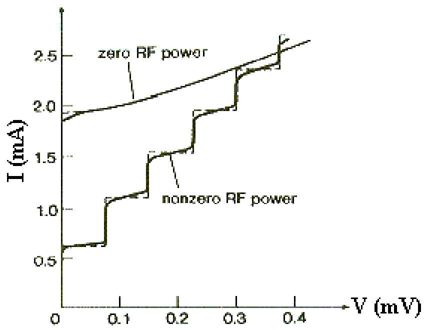

Figure 2: Typical picture of JJ with Shapiro steps

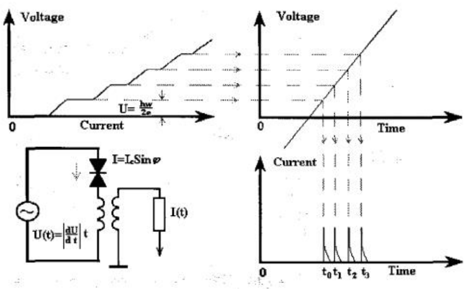

Figure 3: Explanation of how Shapiro steps measurements's JJ frequency meter is operating.

If to transform typical picture of the Shapiro steps on Fig.2 from co-ordinate I-V to co-ordinate V-I and to use transformer in bias network (there is the linear voltage shifting law, where dV/dt ~ fixed constant) it will be situation as on fig.3.

According to Shapiro steps there arise impulses from the second winding. These impulses are given directly by:

$$\omega = (2e|dV/dt| \cdot \Delta t)\, h \qquad (1)$$

The interval between impulses from the second (step-up) winding of the transformer bring needed information about frequency of EMR and after calibration can be re-calculated into the automatic scheme for the frequency measuring.

This time it is reason to remember the result with Tl -based HTSC on fig.1. Based on the Josephson - effect the intrinsic abilities the gap frequency of HTSC is about 30 THz (P. Wang, Zh. Wang, B. Fan, W. Xie, X. J. Zhao, M. He, X. Zhang, L. Ji, and S. L. Yan., 2012), which is higher than that of low temperature superconductor**.**

It denotes that the wide-range THz tuner which is combination of simplest visual power meter and frequency meter based on the Shapiro steps properties in the single block can respond at all the THz band.

3 CONCLUSION

In this paper it was done simplest explanation how to use Shapiro steps which have place in case EMR of JJ for the wide-range tuning of experimental generator in THz band.

It will be especially popular for the tuning of cooled generator.

As the real prognosis - possible later such device will take name *Shapiro-steps tuner of generators in THz band*.

ACKNOWLEDGEMENTS

This work was prepared according to the finance support by national program of the People's Republic of China "The Recruitment Program of Foreign Experts" under the State Administration of Foreign Experts Affairs.

REFERENCES

Wang H.B., Guénon S., Gross B., Yuan J., Jiang Z. G., Zhong Y.Y., Grünzweig M., Iishi A., Wu H., Hatano T., Koelle D., Kleiner R.., 2010. "Coherent Terahertz Emission of Intrinsic Josephson Junction Stacks in the Hot Spot Regime",*Phys. Rev. Lett. 105, 2010.*

Koshelev A.E., Bulaevskii L.N. 2008. "Resonant Electromagnetic emission from intrinsic Josephson - junction stacks with laterally modulated Josephson critical current" *Phys. Rev. B 77, 014530,2008.*

Kadowaki K., Ozyuzer L., Koshelev A., Kurter C., Gray K.., 2008 “Direct observation of Terahetz Electromagnetic waves emitted from intrinsic Josephson junctions in single crystalline Bi-ceramics*”, Physica C, vol.468, p. 634, 2008.*

Crankshaw D. **S.**, Trias E., and Orlando T.P. 2001, Magnetic Flux Controlled Josephson Array Oscillators. *IEEE Trans. On Appled Superconductivity, Vol. I I , No. I, March 2001, I223* ww w.fullsharecryogenics.com

Likharev K., Ul'rich B.., 1978. *“Systems with Josephson contacts”,* MGU Press, Moscow, 1978. I.Kulik, I. Yanson., 1972. *Josephson-effect in Superconducting Tunneling Structures,* Halsted press, 1972.

Denisov A. Obolonsky V., 1990. “Frequency measurement in millimeter in submillimeter electromagnetic wave bands using Josephson junctions*” Second World Congress on Superconductivity, Houston, Texas,USA, 10-15, 1990.*

Denisov A., Gorishnyak V., 1981 “Digital frequency meter based on Josephson Junctions*”, Soviet “Electronika of microwaves, TsNII “Electronika”, vol.4, 1981, p. 45-55, 1983.* [

Denisov A., Gaevsky V., Gassanov L., Kuzenkov S., Nazarenko L., 1985. “Study of the operation of Josephson Down-Converter in the Self Pump Mode”, *Soviet “Radiotechnika i elektronika”, №.5, p.1006, 1985.*

Denisov A., Qi J., Hui Q. J., Gudkov A..,2015 Possibilities to make the panoramic receiver- frequencymeter in terahertz band based on Josephson junctions, *IEEE Trans. On Appl. Superconductivity, Vol. 25, No. 2, April20151800909.*

A.Denisov, Hui Q. J., Qi J., Chen L.. 2014 Towards to make universal device for doing experiments with new generation structures in terahertz band. *20 International conference on microwaves, Radar and Wireless Communication, MIKON –2014, Gdansk, June 16-18, 2014*

Wang P., Wang Zh., Fan B., Xie W., Zhao X. J., He M., Zhang X., Ji L., and Yan S. L.., 2012. Millimeter-Wave Radiation from Misaligned Tl2Ba2 CaCu2O8Thin-Film Intrinsic Josephson Junctions. *IEEE Trans. on Applied Superconductivity, Vol. 22, No. 5, October 2012: 1500503*

Improving Signal of Opportunity Localisation Estimates in Multipath Environments

Thomas O. Mansfield, Bogdan V. Ghita and Adrian Ambroze
School of Computing and Mathematics, Plymouth University, Plymouth, UK
{thomas.mansfield, bogdan.ghita, m.ambroze}@plymouth.ac.uk

Keywords: Leading Edge Detection, Time of Arrival Localisation, Indoor Navigation, Signals of Opportunity.

Abstract: Network based geographic localisation has been widely researched in recent years due to the need to locate mobile data communication nodes to a level of accuracy equivalent to that provided by global navigation satellite systems (GNSS) in multipath urban and indoor environments. This paper investigates whether direct sequence spread spectrum (DSSS) signal processing can be applied to narrow-band radio channels to improve the ranging estimates. The DSSS signal processing application is then developed further to provide a method of deriving a measurement confidence indicator, allowing the optimisation of time separated measurements in a dynamic signals of opportunity radio environment. A set of validation tests demonstrates that the proposed method provides a significant improvement in the accuracy and robustness of the ranging estimate compared to simple threshold analysis in multipath environments.

1 INTRODUCTION

Radio positioning systems have achieved common use in a diverse range of systems. The most commonly used radio positioning systems are global navigation satellite systems (GNSS). These systems use signals received from satellite to calculate the position of the user to within 4m during 95% of the time (Norman Bonnor, 2012). GNSS systems rely on a line of sight (LoS) view of at least 4 satellites. This requirement cannot however, be guaranteed in urban or indoor environments where 'urban canyons' and roof cover, block sight to much of the surrounding sky. Research has been carried out into using signals of opportunity for localisation in such environments, particular success has been achieved by using time of arrival (ToA) systems to derive a user's location (Norman Bonnor, 2012), even in urban or indoor environments where multipath propagation is one of the main sources of system error 0. Constructive and deconstructive interference between the non-line of sight (NLoS) propagating signals can destroy or obscure the LoS signal that is required to derive an accurate ToA estimate.

Ultra wide band (UWB) signal analysis techniques, originally developed for low emission radar 0, have achieved promising results when applied to localisation in wide bandwidth direct sequence spread spectrum (DSSS) networks (00). These techniques rely on the differing multipath properties of the wide spread of frequencies to provide an improved leading edge time of arrival (ToA) estimate and to achieve GNSS levels of accuracy in wide bandwidth multipath environments.

This paper builds on the use of prior art wide bandwidth signal processing techniques and investigates their use in signals of opportunity networks that commonly collect time separated narrow bandwidth measurements such as frequency hopping spread spectrum (FHSS) networks. FHSS networks are typical to military 0 and civilian (IEEE, 2014. *IEEE 802.22)* systems and challenges remain to use them to achieve GNSS levels of location accuracy in multipath environments 0 due to the time separated nature of the received signals.

This paper proposes a method that allows the system to use time separated ToA estimates and, without prior training or additional data collection, generate a low latency and high bandwidth filtered ranging estimate. The benefits of the proposed method are verified through simulation. The accuracy and responsiveness of the ranging estimate shall be analysed in both static and mobile receiver environments.

This paper is organised as follows; Section 2 discusses the prior art. Section 3 proposes a method to use the leading edge detection algorithm to extract the data required to weight the values in a recursive filter. Section 4 provides details of the simulation environment and evaluates the ranging estimate performance. Section 5 concludes and discusses further work.

2 PRIOR ART

2.1 Leading Edge Detection

Basic ToA detection systems commonly use simple threshold based leading edge detection 0), which relies on the assumption that the LoS message will arrive first via the shortest direct path. In many situations however, the LoS component may be heavily attenuated by deconstructive multipath interference providing a leading error driver for indoor or urban ranging system accuracy.

Search-back algorithms improve on the ToA accuracy by analysing the received packet and performing a search-back to determine physical layer properties of the message to determine the time of arrival more robustly (Haneda K., 2009). These algorithms require prior knowledge of the multipath environment which cannot be provided in many applications.

The Multiple Signal Classification (MUSIC) algorithm (Schmidt R. O. 1986) extends the analysis to allow multipath signals to be used as a further information source and has become widely used in research. This algorithm requires a substantial training period to determine the number of multipath signals present to achieve better performance than relying on leading edge detection alone. Again, a training period is not practical in many applications where the device is to be used to navigate around an unknown area.

UWB signal processing techniques utilise the wide frequency range of the received signals to provide an improved ToA estimate. The analysis of the full frequency range available allows the user to determine frequency specific multipath variations and make an improved estimation of the true ToA reading. A widely implemented example of an existing UWB signal processing technique, described in 0, has been selected for further development in this paper. This technique was developed to detect the leading edge of a signal obtained from a wide bandwidth transmission. It has been selected for further development due to the fact that the running filters applied to the raw data may provide additional data to the user following further analysis.

The UWB signal processing technique is applied to any wide band received data as follows: if $h(t)$ represents the received signal in the time domain, it is first passed through a rectified moving average filter as shown in (1).

$$y(t) = \frac{1}{n} \sum_{i=t-n+1}^{t} abs(h([t])) \tag{1}$$

The averaged signal $y[t]$ is then passed through two filters of sizes n_1 and n_2 which return the maximum value from a sliding window, as shown in (2) and (3).

$$\max\_n_1[t] = \max(y_{t-n_1} \ldots y_t) \tag{2}$$

$$\max\_n_2[t] = \max(y_{t-n_2} \ldots y_t) \tag{3}$$

A binary indicator of whether a leading edge has been detected can be obtained from (4).

$$\begin{aligned} r[t] &= (\max\_n_1[t] * 2 > \max\_n_2[t]) \\ &\quad \& (\max\_n_2[t] > thresh) \end{aligned} \tag{4}$$

The threshold detection level, *thresh*, is typically set to 3σ of inter message in-channel received signal noise.

2.2 Application Considerations for Navigation Filters

Recursive averages are commonly used in navigation systems to produce a low noise and low latency location estimate from a noisy measurement input. In order to provide an efficiently filtered output, the measurement system that populated the recursive filter must provide not only a measurement value, but also a dynamic confidence indicator.

When using a simple threshold detection algorithm to detect the leading edge of a received signal, the only information that can be provided to the navigation filter is the time when a received value is greater than the selected threshold. If this information is available for each FHSS channel, a simple un-weighted recursive filter shown in (5) can be constructed to update the users filtered location based on the its previous position and the latest sensor data where, as commonly used in filter notation, $\hat{x}$ represents the filter output, $\bar{x}$ represents the previous state and $\tilde{x}$ represents the latest sensor value. The measurement confidence is represented by α.

$$\hat{x} = \alpha\bar{x} + (1-\alpha)\tilde{x} \quad (5)$$

The filter represented in (5) may be tuned by adjusting the value of α by a predetermined value.

A value of $\alpha < 0.5$ reduces the noise of the filter output at the expense of a higher latency if the receivers true location changes. A value of $\alpha > 0.5$ generates a more responsive, lower latency filter output but the filter output noise will be adversely affected. Both of these options are unsuitable for many system applications.

3 PROPOSED METHOD

The leading edge detection algorithm described in section 2.1 has been developed for wide band signal processing and analyses all of the data from the wide frequency range with each measurement.

The receiver system to be developed by this paper makes a ranging estimate upon detection of the leading edge of a received signal using the signal processing technique described in section 2.1. The process of running the n_2 filter (3) to return the maximum value in the longer sliding window continues for the duration of the first message in the current FHSS channel. The data obtained from the maximum value sliding windows is placed into a column vector and a standard deviation taken to determine the presence and magnitude of multipath present throughout the message. This is then correlated to provide a numerical confidence value.

The process is represented in equations (6) and (7). The standard deviation, σ, is first calculated in (6) with n_2 as the filter length, x_i is each iterative filter value and x_a is the current filter average. This standard deviation is then normalised in (7) to produce a dynamic measurement confidence, α.

$$\sigma_{n_2} = \sqrt{\frac{1}{n_2}\sum_{i=1}^{n_2}(x_i - x_a)^2} \quad (6)$$

$$\alpha = (1 - \frac{\sigma_{n_2}}{\mu_{n_2}}) \quad (7)$$

α represents a confidence factor with a weighted value between 0 and 1 for low to high confidence measurements respectively. This confidence measure can then be used to dynamically tune the filter shown in (5) to generate a recursive filter input that benefits from both low noise and low latency. This has been achieved by providing a high weighting value to ranging estimates received with good confidence and a low weighting to estimates with a low confidence, even if there has been true movement by either the transmitter or receiver.

The ability to achieve this from a multipath data source dynamically and without prior knowledge is of a key benefit in higher level navigation systems, as discussed in section 2.2. This confidence weighting has been achieved without the use of any additional information or averages over the ones implemented to allow the improved leading edge detection.

4 SYSTEM VALIDATION

4.1 Simulator Validation

A simulated radio frequency (RF) environment was modelled in Matlab® and Simulink® to evaluate the effectiveness and performance of the techniques discussed in section 3. The simulation uses the standard multipath simulation model (Alsindi N.A, 2004) shown in (8) where L_p is the number of multipath components, α is the complex attenuation and τ is the propagation delay.

$$h(t) = \sum_{k=0}^{L_p-1}\alpha_k\delta(t-\tau_k) \quad (8)$$

The simulation assumes that an idealised transmitter generates a single frequency modulated pulse; for validation, the FHSS network parameters included 100 20 kHz channels evenly spaced from 3 to 5 GHz. The transmitted pulse is then subjected to empirically derived propagation and receiver distortions to produce a received signal for analysis. The resulting signal includes simulated effects of multipath with the use of separate propagation channels, the number of which can be set by the user. The simulations evaluated throughout this paper will consider a LoS propagation path of 10 m with several multipath reflection paths with an apparent time path from the transmitter to the receiver consistent with 10.1 m to 11.2 m propagation distances.

This simulated environment has been used to ascertain the performance of a simple threshold detection algorithm in a Monte Carlo based simulation of a wide range of FHSS channels in a fixed geometry. A typical single transmitted message and the received signal patterns in a high multipath environment can be seen in Fig 1.

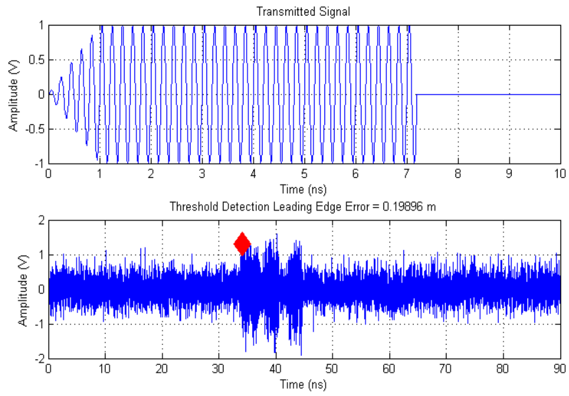

Figure 1: Transmitted (top) and received (bottom) pulse with the location of the detected leading edge of the pulse marked by the red symbol.

The threshold detection algorithm has been simulated assuming a static receiver and transmitter across a range of FHSS channels to benchmark the simulation. The results can be seen in Fig. 2 and shows properties that are expected in multipath environments, as seen in (Norman Bonnor, 2012)] and 0). The similarity to data collected by practical test in previous research provides confidence that the simulation is representative.

4.2 Technique Validation

A comparison of edge detection seen by employing UWB signal processing techniques to each narrow bandwidth channel as opposed to simple threshold detection can be seen in Fig. 2.

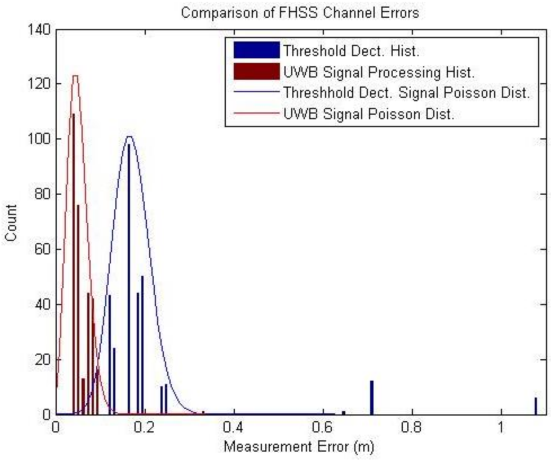

Figure 2: Comparison of threshold based and UWB signal processing leading edge detection methods.

Analysis shows that the Poisson distribution variance has a λ value of 17 for the threshold detection algorithm and an improved λ value of 5 for the UWB threshold detection. The received estimates across the range of networks not only have less average error but also a greater distribution density than can be obtained from simple threshold detection alone. As well as a significant improvement in the Poisson distribution, the UWB based edge detection algorithm removes the erroneous outliers seen at ≈ 0.7 m and ≈ 1.1 m error in the threshold detection algorithm. This behaviour may account for the high multipath uncertainty seen in (Faragher R. M., 2007) where a simple threshold detection algorithm was used to detect the ToA to estimate range.

Detail of the detected trigger timing at the leading edge of a signal with light multipath is shown in Fig. 4.

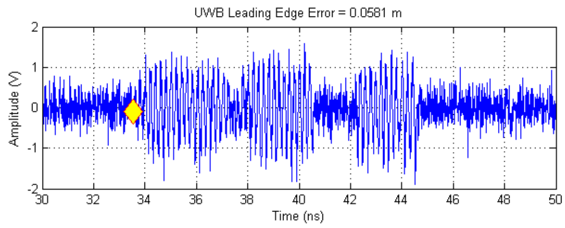

Figure 3: UWB leading edge detection of pulse in a noisy multipath environment.

Figure 3 is a magnification of the area of interest, related to the transmission pulse as shown in Fig. 2. Areas of constructive and deconstructive multipath effects can be seen throughout the 34 ns to 42 ns region where a non-multipath signal would be expected to produce a stable series of 1 V peaks.

The simulation has shown that the evaluation tests for the UWB algorithms discussed in section 2.1 produce a significant improvement over threshold detection when providing ToA estimation in high multipath FHSS networks when only a single narrow bandwidth channel can analysed at a time.

Further to the improvement shown in ToA estimates in a high multipath environment, the application of the additional data available, described in section 3, to a recursive navigation filter is analysed in the remainder of this section.

The application of threshold analysis data, where no weighting data is available for the new samples, into the simplified recursive filter leads to a noisy and poorly filtered position estimate. Fig 4. compares a plot of the raw measured and filtered ranging estimate obtained from a simulation of a static system that sweeps through 100 FHSS channels over a 5 second period.

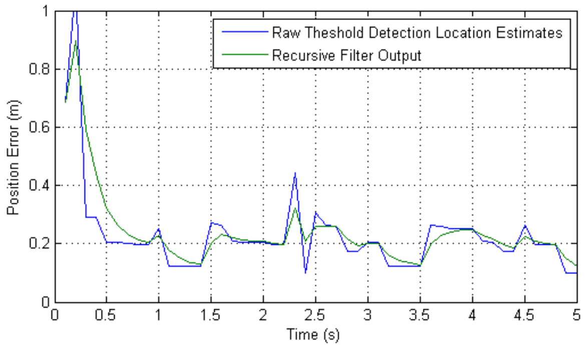

Figure 4: The raw and filtered output from the threshold detection algorithm with a pre-selected static confidence interval.

The results displayed in Figure 4 verify that the filtered position estimate from an un-weighted recursive filter is comparatively noisy and produces a large filtered error in the event of a multipath $\tilde{x}$ leading edge detection received from the sensor, as seen approximately 0.2 seconds into the simulation. The application of the position estimates and the relative variance derived using the method described in section 3 has been applied to a weighted navigation filter. The application of this navigation filter in the simulation leads to improved stability to the position estimate which, combined with the improvement in leading edge detection reliability and the absence of outliers, leads to a greatly improved position estimate over the threshold detection algorithm, as shown in Fig. 5.

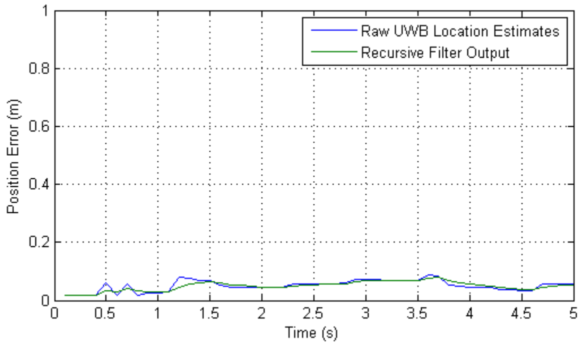

Figure 5: The raw and filtered output from the navigation filter with UWB leading edge detection and dynamically obtained confidence interval. This should be compared with Fig 6 to see the improvement achieved.

In a physically static system, as simulated in Fig 4 and Fig. 5, where the relative position of the transmitter and receiver does not change, the sensitivity to erroneous data could be mitigated by weighting the raw sensor data by a pre-determined factor of < 1 depending on sensor noise. While this will limit the filter error in the event of erroneous multipath readings and produce a more accurate location estimate, it also introduces high latency if the receiver or transmitter truly moves location. The application of a dynamically weighted recursive filter prevents an erroneous multipath ToA reading from causing filter noise. If however, the system truly moves, a new filter input with a new position estimate with a high weighting will be received and the filter output will respond with little latency.

A further simulation was run to evaluate the effect of a true receiver motion on the filter output. To simplify the simulation, a single narrow bandwidth channel with no frequency hopping was used throughout the experiment. After approximately 1.2s into the simulation, the receiver node instantaneously moves 1m within a multipath environment and remains static for the remainder of the simulation.

Fig. 6 shows a comparison of the filter response to the applied motion with both threshold detection and UWB detection inputs.

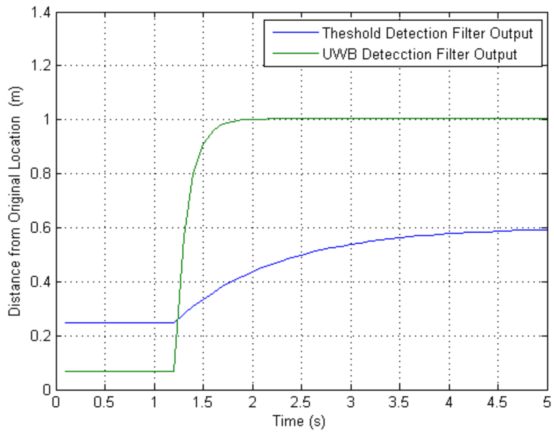

Figure 6: The response of the filters to an instantaneous 1 m movement of the receiving node.

The threshold detection filter still has a greater error before and after the 1 m move of the receiver node than the UWB filter, as expected. The area of interest highlighted by this simulation is the difference in time taken for the filter output to identify the change in location. The dynamic weighting to $\tilde{x}$ allows the UWB filter to respond with minimal latency in the event of true receiver or transmitter movement. The improvement seen in Fig 6 is due to both the improved UWB ranging estimate, shown in Fig 2 and the ability to weight the measurements. These contributing factors have not been analysed separately due to the fact that the weighted recursive filter may be implemented without any additional data collection and should always be used to provide an optimised solution.

5 CONCLUSIONS AND FURTHER WORK

This paper proposed a set of algorithms and application techniques that improve narrow bandwidth channel ranging estimates in signals of opportunity environments. The novel application and further development of DSSS signal processing techniques to provide not just an improved ranging estimate but, by re-analysing existing data, an additional confidence weighting.

By re-analysing the available data, a filter confidence factor can be obtained that can be calculated dynamically without the need for a training period and without any prior knowledge of the radio system and environment. More specifically, the use of UWB signal processing techniques provided an approximately 4 times improvement in ranging estimation over simple threshold detection even in narrow bandwidth channels, including a better Poisson distribution and higher resilience to false detections.

The main benefit of applying this technique is that a filtered ranging estimate can be obtained that is more accurate, lower noise and lower latency than can be obtained by using simple threshold detection techniques to detect the leading edge of a message.

The analysis of the proposed technique performance throughout this paper has been carried out only in multipath environments. It is anticipated that the benefits of the technique will be significantly less apparent in less hostile environments.

Future work should include the physical test of this system to verify the model. The integration of the algorithm into higher level systems is also required to verify the higher level benefits shown during simulation. The close coupling of this system with higher level navigation systems, in particular Kalman filtering schemes may also allow the development of a significantly improved signal of opportunity based localisation system.

REFERENCES

Air Commodore Norman Bonnor, 2012. *A Brief History of Global Navigation Satellite Systems*, Royal Institute of Navigation Journal of Navigation vol. 65.

Faragher R. M, Duffett-Smith P.J, 2010. *Measurements of the effects of multipath interference on timing accuracy in a radio positioning system*, IET Radar, Sonar and Navigation, vol. 4, iss. 6, pp. 818-824.

Fontana R. J., 2004. *Recent System Applications of Short-Pulse Ultra-Wideband (UWB) Technology*. Published in IEEE Microwave Theory and Techniques, Vol. 52, Issue 9.

Mahfouz M. R., et. al, 2008. *Investigation of high accuracy indoor 3-D positioning using UWB technology*, IEEE Trans. Microwave Theory and Technology, pp. 1316-1330.

Merkl B., 2008. *The future of the operating room: surgical preplanning and navigation using high accuracy ultra wide-band positioning and advanced bone measurement*, Ph.D Dissertation, The University of Tenessee, USA.

Grumman Northrop, 2013. *Understanding Voice and Data Link Networking*, Guidebook, Northrop Grumman, San Diego, USA.

IEEE, 2014. *IEEE 802.22 Working Group – Enabling Broadband Wireless Access Using Cognitive Radio Technology* http://www.ieee802.org/22/ Accessed 14th November 2014.

Xu P., Palmer R.J and Jiang Y, 2005. *An Analysis of Multipath for Frequency Hopping Spread Spectrum Ranging*. IEEE Conference on Electrical and Computer Engineering.

Haneda K., 2009. *Performance Evaluation of Theshold Based UWB Ranging Methods – Leading Edge vs, Search Back,* 2rd European Conference on Antennas and Propagation.

Schmidt R. O. 1986, *Multiple Emitter Location and Signal Parameter Estimation*, IEEE Trans. Antennas Propagation, Vol. AP-34 (March 1986), pp.276-280

Alsindi N.A, 2004. *Performance of TOA Estimation Algorithms in Different Indoor Multipath Conditions*. PhD Thesis, Worcester Polytechnic Institute.

Faragher R. M., 2007. *Effects of Multipath Interference on Radio Positioning Systems*, Ph.D. dissertation, Dept. Physics., University of Cambridge, Cambridge

Non-Stationary Random Wiener Signal Detection with Multistatic Acoustic System

Volodymyr Kudriashov
Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
25A Acad. G. Bonchev Str., Sofia, Bulgaria
KudriashovVladimir@gmail.com

Keywords: Acoustic signal processing, ultra wideband antenna, localization, imaging.

Abstract: The paper presents detection rule for multistatic reception of the non-stationary acoustic signal. The detection rule is obtained using maximum likelihood approach. Usually angular beam forming is applied to microphone array to localize spatially distributed emitters. In the paper, the time difference of arrival estimates of incoming acoustic emissions are used to localize their sources. The paper proposes wide frequency band acoustic noise source detection and localization enhancement using multistatic reception system. The paper shows experimental result on localization of source of wide frequency band emission by sound pressure imaging. All passband bandwidth of incoming signal is processed simultaneously. The localization is provided in range-cross range-elevation coordinates. The proposed technique may be suitable for 4D imaging in non-destructive testing and in ultra-wideband acoustic emitters' detection and localization. One of particular applications concerns testing of aircrafts landing regime and health monitoring of their engines at landing/take off.

1 INTRODUCTION

Detection rule is required to localize source of non-stationary random Wiener signal in range – cross range – elevation coordinates (Wentzell 1996, Levin 1969). The rule enables to define threshold level and the detector block diagram (Rozov 1987, Gusev 1988, Shirman 2007). Existing systems for acoustic noise source localization use pre-defined range to generate sound pressure images in cross range – elevation coordinates. As well, the systems generate conformal sound pressure images. The generation of those images uses beamforming techniques based-on estimates of phase difference of arrival of incoming signals in predefined frequency passband. The cited works do not contain detection of incoming signals. The paper presents rules for detection the signal against non-stationary random Wiener interference via bistatic and multistatic acoustic systems as well as corresponding threshold levels and block diagrams.

2 PROBLEM STATEMENT AND SOLUTION

Emission of an object is considered as a realization of non-stationary random Wiener signal. The signal frequency bandwidth is wide (Brüel & Kjær Sound and Vibration Measurement A/S 2009, Christensen and Hald 2004, Hald et al. 2004). Receivers and microphones limit it by their bandwidth B. The microphones are significantly spaced. Estimated parameter is time difference of arrival of incoming signal to the microphones. The pair of receivers' output signals are denoted as $y_I(t)$ and $y_{II}(t)$, correspondingly. The signals may contain the incoming signal (condition $A=1$) or not contain it (condition $A=0$) (Rozov 1987, Gusev 1988 and Shirman 2007).

The detection rule is derived for the incoming signal $x(t)$ against mix of interfering signals $c_I(t)$, $c_{II}(t)$ and intrinsic noise of microphones and receivers $n_I(t)$, $n_{II}(t)$. The intrinsic noises' power spectral density is N_0, for B of the equipment. The signal model is denoted as (Rozov 1987, Gusev 1988 and Shirman 2007):

$$\begin{aligned} y_I(t) &= A\,x(t-t_x)+n_I(t)+c_I(t-t_c), \\ y_{II}(t) &= A\,x(t-\tau)+n_{II}(t)+c_{II}(t-\tau), \\ & 0<t<T \end{aligned} \tag{1}$$

where $x(t)$, $n_I(t)$, $n_{II}(t)$, $c_I(t)$ and $c_{II}(t)$ are not correlated in pairs; t_x and t_c are TDOA for the incoming signal and interference (industrial noise, multipath propagation on the scene etc.); τ is time delay that introduced to compensate the t_x; and T is acquisition time.

According to the Wiener process property, the considered $x(t)$, $n_I(t)$, $n_{II}(t)$, $c_I(t)$ and $c_{II}(t)$ have independent increments those obey normal distribution (Wentzell 1996, Levin 1969). The exact time interval, which enables to obtain the normal distribution of the increments, may be obtained by further experimental investigations.

The digital signal processing assumption enables to present the signals (1) as Kotelnikov series with constant interval $1/2B$ of time sampling. Elements of the $\vec{Y}$ are the noted above increments $\Delta y_{I,i}$ and $\Delta y_{II,i}$.

Probability densities of the $\vec{Y}$ are obtained for the two conditions: $A=1$ and $A=0$, in order to obtain likelihood ratio $L(\vec{Y})$ and the detection rule. At condition $A=1$, the incoming signal is correlated, as well as the interference. Joint probability density of corresponding samples $\Delta y_{I,i}$ and $\Delta y_{II,i}$ obeys two-dimensional distribution function of two normally distributed random variables (Levin 1969). The corresponding probability density function is obtained based-on following equality:

$$p(\vec{Y}/A=1)=\prod_{i=1}^{k} p(\Delta y_{I,i}, \Delta y_{II,i}/A_i=1),$$

where $k=2BT$. At the condition $A=0$, the probability density function is obtained similarly. At the latter condition, no elements of the $\vec{Y}$ are correlated, except the interference. Relation of $p(\vec{Y}/A=1)$ to $p(W/A=0)$ is the likelihood ratio. For the technical implementation, natural logarithm of the obtained $L(\vec{Y})$ is more appropriate. One assumes that variances of increments of the noise and the interference are larger than variance of increments of the signal. Thereat, one of addends of the obtained expression do not depends on the incoming signal. The addend defines the threshold level. Assumption that variances of increments of the noise are larger than variances of the interference enables to obtain weight of the integration in the expression. The obtained detection rule estimates autocorrelation functions of increments of signals (1) and their cross-correlation function. Only the latter depends on time difference of arrival of incoming signal. Thus, the detection rule envisages calculating the expression:

$$Z_1 \approx \int_0^T k_I \Delta y_I(t)\; k_{II} \Delta y_{II}(t)\, dt \tag{2}$$

where $k_{I,II}$ define gain values for receivers 1 and 2, correspondingly; $\Delta y_{I,II}(t)$ define increments of the signals (1).

The rule for non-stationary random Wiener signal detection in bistatic reception system is obtained (Fig. 1, baseline 1).

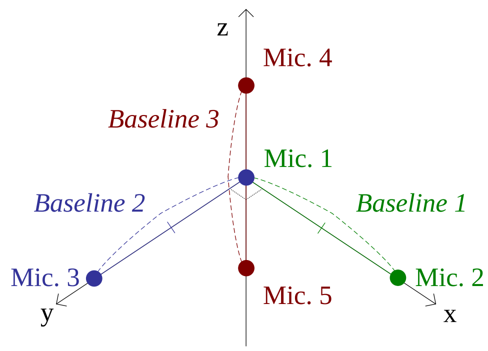

Figure 1: Basic geometry of the multistatic acoustic system.

The detection rules for other baselines (Fig. 1) may be expressed similarly to (2). Output signals of the bistatic reception systems are denoted as $u_1(t)$, $u_2(t)$ and $u_3(t)$, correspondingly.

The further detection rule obtainment is similar to the above one. But, the new $\vec{Y}$ consists of $u_1(t)$, $u_2(t)$ and $u_3(t)$ samples. The samples are denoted as $u_{1,i}$, $u_{2,i}$ and $u_{3,i}$. At the condition $A = 1$, the signal and interference components of the $\vec{Y}$ are correlated in pairs. Joint probability density of corresponding samples $u_{1,i}$, $u_{2,i}$ and $u_{3,i}$ obeys distribution function of normally distributed random variables (Levin 1969). At the condition, the corresponding probability density function is obtained based-on following equality:

$$p\left(\vec{Y} / A = 1\right) = \prod_{i=1}^{k} p\left(u_{1,i}, u_{2,i}, u_{3,i} / A_i = 1\right).$$

At the condition $A = 0$, the samples of $\vec{Y}$ are independent. Variations of these samples are same. Relation of the latter probability density functions is the new likelihood ratio $L\left(\vec{Y}\right)$. One of addends of the obtained ratio defines the threshold level, as the signal power in the multistatic system is low. Other one addends sum of power estimates from the three considered bistatic systems (Fig. 1). All possible cross-baseline cross correlation functions are subtracted from the latter addend. The last addend provides multiplication of power values of output signals of the bistatic systems. The latter is agreed to detection quality at limited number of samples (Shirman 2007). The input signals squaring is valuable for small signal-to-noise-plus-interference ratio at outputs of the bistatic systems. Spatial localization of the emission source is utilized by the considered multistatic system (Fig. 1) by the latter addend:

$$Z_{123} \approx \int_0^T k_1 u_1^2(t)\, k_2 u_2^2(t)\, k_3 u_3^2(t)\, dt \qquad (3)$$

where k_i are gain values of corresponding bistatic systems 1-3. All intermediate results and threshold level expression were dropped down.

The obtained requires to estimate TDOA of the signal by each bistatic system and to provide further calculation according to (3), for each node of spatial grid.

The non-stationary random Wiener signal detection rule for three bistatic systems is obtained according to the maximum likelihood method with respect to the threshold level.

3 EQUIPMENT OF THE ACOUSTIC CAMERA

Acoustic camera, manufactured by Brüel & Kjaer (Sound and Vibration Measurement A/S) is used. The camera uses 18 microphones type 4958, 12-channel and 6-channel input modules type 3053-B-120 and 3050-B-060, correspondingly. The acoustic camera includes Pulse LabShop software. The latter was used to transfer the multichannel equipment output signals for further post-processing.

The microphones dimensions are: 34 mm long, 7 mm diameter. Sensitivity of the microphones is 11.2 mV/Pa. Operating temperature range of the microphones is from 10˚C to +55˚C. The microphones dynamic range is from 28 dB to 140 dB. The microphones have CCLD preamplifier with transducer electronic datasheet (TEDS - IEEE 1451.4 V.1.0).

Both input modules support TEDS transducers and deliver REq-X technology, which flattens the transducers frequency responses by “mirroring” them. These input modules are mounted in 5-slot Mainframe LAN-XI type 3660-C-000 with battery module type 2831. The 3050-B-060 input module delivers Dyn-X technology that expands its dynamic range depending on exact signal quantization and bandwidth.

The acoustic camera upper frequency is 25.6 kHz and its quantization rate is about 65 kHz. The signals are synchronized using IEEE 1588 Precision Time Protocol.

The camera calibration may be provided in advance to assure precision of sound pressure estimates. The acoustic camera incudes hardware and software for the calibration. The portable calibrator is battery operated. The calibration frequency is 251.2 Hz. Pistonphone calibrator type 4228 with external barometer satisfies ANSI S1.40-1984 and IEC 942 (1988) Class 0L. The calibrator has following adaptors: DP-0775 for sequential calibration of the microphones and adaptor WA-0728-W-003 for calibration of groups of 6-microphones. The calibrator can be used over a wide range of temperature, humidity and pressure while still maintaining high accuracy.

Optic camera with resolution 640×480 pixels and microphones in 0.33 m slice wheel array of the acoustic camera are mounted on 3D tripod head

Manfrotto 229 and tripod Manfrotto 058B. The load capacity of both head and tripod is 12 kg (safety payload).

The acoustic camera data transfer cables limit spatial separation between the equipment blocks. The acoustic camera consists of three main blocks: a microphone array with optic camera, input modules in a frame and a laptop with software. Cable harness type WL1297-W-004 2013W21 with length ~4.5 m limits separation between the microphone array and the mainframe. Both optic camera USB cable with length ~6.3 m and LAN cable type AO1450-D-020 2013W13 with length ~2 m limit separation between the mainframe and laptop with the acoustic camera software. The main features of the laptop type 7201-E-GB2 (Dell Latitude E6430) are listed below: E6430 CPU, 6 GB RAM, 1 TB HDD, Wi-Fi, Ethernet 1 Gb, DVD-RW.

The acoustic camera hardware and software modules are supplied with full documentation (instruction manual, specifications). Acoustic images may be generated using Array Acoustics Post-processing (Version 17.1.2.308). Measurement data may be collected using Pulse LabShop (Customized Solution Version 17.1.2). Other existing software and drivers are not listed in this work.

4 DETAILS AND RESULT OF EXPERIMENT

Experiment is focused on localization of acoustic noise emission source with the multistatic system (Fig. 1). Microphone 1 is placed in the origin of Cartesian coordinate system. Baselines equal to 1 m. Emitter coordinates in the field of view are as following: range 1 m, cross range -0.8 m and elevation 0.15 m. Center frequency of the incoming signal is about 5 kHz and bandwidth of the signal is about 10 kHz. The latter corresponds to TDOA resolution of about 3.5 cm, at a baseline.

The signal processing contains multiplication of the bistatic systems output signals normalized squared. The amplitude calibration of the Acoustic Camera was carried out in advance. Obtained results are displayed in logarithmic scale (Fig. 2). Multipath propagation inside a typical office room affects equality of responses of the bistatic systems. Thus,–6 dB threshold is applied to the results. The normalization is not present in (3). Slight irregularity of responses of three bistatic systems (Fig. 2) is affected by insufficiently small pixel size.

The result shows opportunity to localize the emission source with the presented approach (3).

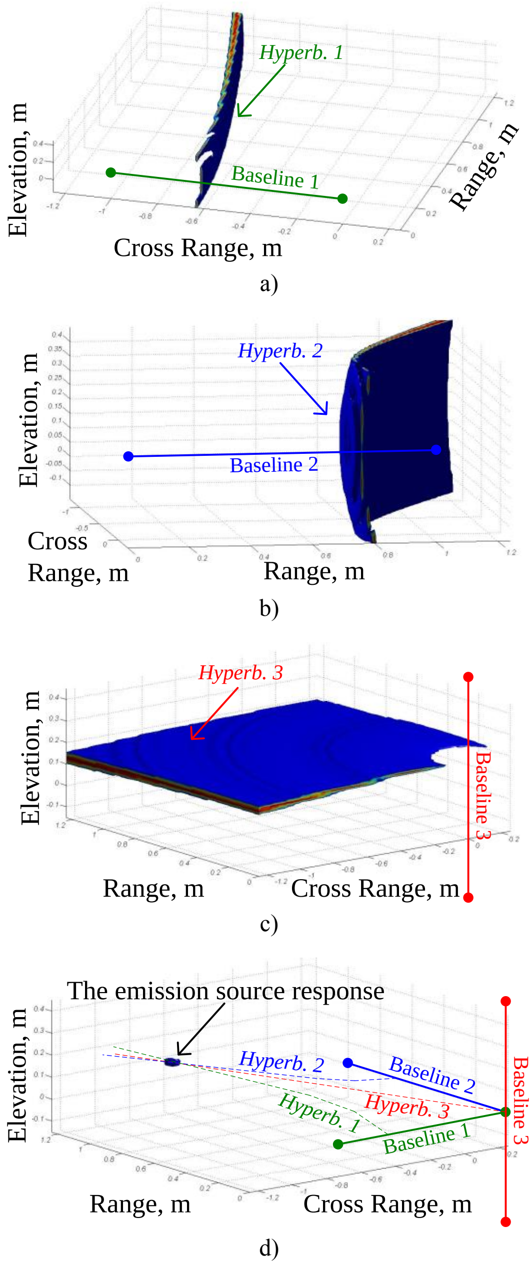

Figure 2: Acoustic images generated using the shown bistatic systems and according to the proposed approach using the multistatic system.

5 CONCLUSIONS

Newly developed detection rules of non-stationary random Wiener signal against such interference are proposed for bistatic and multistatic acoustic systems. The rules enable to define corresponding threshold levels and technically feasible block diagrams. The four-site system is considered for spatial localization of acoustic emission source. The proposed approach uses time difference of arrival estimates of incoming signal instead of its phase difference of arrival estimates to localize the source. Test source passband bandwidth about 10 kHz is processed simultaneously, in the experiment. Further implementation of the approach is promising for wideband acoustic noise source localization.

ACKNOWLEDGEMENTS

The research work reported in the paper was partly supported by the Project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).

REFERENCES

Wentzell, A.D., 1996. *Course of the theory of random processes*. Science. Moscow. In Russian.

Levin, B.R., 1969. *Theoretical bases of statistical radio engineering, Vol. 1*. Soviet Radio. Moscow. In Russian.

Rozov, A.K., 1987. *Detection of signals in non-stationary hydroacoustic conditions*. Shipbuilding. Leningrad. In Russian.

Gusev, V.G., 1988. *Systems for space-time processing of hydroacoustic information*. Shipbuilding. Leningrad. In Russian.

Shirman, Ya.D. (ed.), 2007. *Radio Electronic Systems: Fundamentals of Theory and Design*. Soviet Radio. Moscow, 2nd edition. In Russian.

Brüel & Kjær Sound and Vibration Measurement A/S., 2009. *Brüel & Kjaer Product data*. Brüel & Kjær. Nærum.

Christensen, J.J., Hald J, 2004. Beamforming. In *Brüel & Kjaer Technical Review No. 1.* Brüel & Kjær. Nærum.

Hald J., Ishii Y., Ishii T., Oinuma H., Nagai K., Yokoyama Y. and Yamamoto K., 2012. *High-resolution Fly-over Beamforming Using a Small Practical Array*. Brüel & Kjaer Technical Review No. 1. Brüel & Kjær. Nærum.

Fast Direction-of-Arrival Estimation for Single Source
Near- and Far-Field Approaches for 1D Source Localization

Iurii Chyrka
Institute of Information and Communication Technologies, Bulgarian Academy of Sciences,
25A, Acad. G. Bonchev str., 1113 Sofia, Bulgaria
Yurasyk88@google.com

Keywords: Direction-of-Arrival Estimation, Far-field, Near-field, Source Localization, Autoregressive Moving Average Model, Spatial Frequency Estimation.

Abstract: The new approaches for a single narrowband source direction-of-arrival estimation in a far-field scenario and both direction-of-arrival and range estimation a near-field scenario are proposed. The main idea is to estimate the spatial frequency directly along the uniform linear array aperture from the single-shot measurement. The algorithm based on the autoregressive moving average model of the sinewave is applied for the frequency estimation. The effectiveness of proposed methods is analysed via computer simulations.

1 INTRODUCTION

The problem of direction-of-arrival (DOA) estimation of multiple plane waves generated by narrowband signal sources have attracted considerable interest in the literature due to a variety of applications in communication, seismology, oceanography, radar, acoustics, and so on. This problem is considered in the framework of the array signal processing and signal parameter estimation in particular. Usually the objective is to estimate parameters, such as azimuth, elevation, range, center frequency etc. associated with each signal.

Localization problem can be generally divided into two types, based on the distance between the source and the antenna array: far-field (when $r \geq 2D^2/\lambda$, r is the range between the source and the array reference point, D is the array aperture, λ is the wavelength of the source signal), and near-field localization. In far-field case, the wavefront of the signal impinging on the array is assumed to be planar (Johnson, 2006). When the source is located in the Fresnel region ($0.62\sqrt{D^3/\lambda} < r < 2D^2/\lambda$) or even closer in the near-field $r \leq 0.62\sqrt{D^3/\lambda}$ the wave front gets some curvature. It is reasonable to split processing algorithms onto ones based on the planar wave assumption and ones for the circular wavefront.

For the far-field estimation there are a lot of methods that can be separated onto three categories. The first one is beamforming algorithms like delay-and-sum or minimum variance distortionless response (Bai, Ih, Benesty, 2013), which obtain a nonparametric spatial spectrum by application of a data-adaptive spatial filtering. The subspace algorithms like MUSIC (Stoica, Nehorai, 1989), ESPRIT (Gao, Gershman 2005) use the low-rank structure of the noise-free signal. The maximum likelihood methods (Wax, 1982), (Stoica, Besson, 2000), (Chen, Lorenzelli, Hudson, Yao, 2008) work with statistical properties, but require precise initialization to ensure convergence to a global minimum. Due statistical nature, they need sufficiently big data amount for accurate estimation.

In the case of single source localization, direction finding of the narrowband singal can be interpreted as a problem of a sinewave signal parameter estimation, particularly estimation of the spatial frequency. Besides, reduction of the problem allows using of simplified algorithms. (Wu, Liu, So, 2009).

In the near-field scenario it is necessary to estimate simultaneously two position parameters: a pair of coordinates or DOA and range. Therefore, traditional approaches like MUSIC must be extended to a two-dimensional field. Swindlehurst and Kailath (1988) suggest a quadratic (Fresnel) approximation of the wavefront in the near-field.

Using this approximation, the rotational invariance property can be used with the symmetric subarrays to estimate the DOA by ESPRIT (Zhi, Chia, 2007). In the paper of (Grosocki, Abed-Meraim, Hua 2005) position is obtained through estimation of two angles by weighted linear prediction. Another approach of transformation near-field localization problem to far-field one via interpolation is considered in (Yang, Shi, Liu, 2009).

In this work, we focus on the problem of estimation the DOA of a single source in both far- and near-field situations and the alternative approach of single-shot direct estimation of the spatial frequency from the one source is considered.

2 DOA ESTIMATION ALGORITHMS

2.1 Far-field Scenario

Let us consider a single narrowband signal $s(t)$ that comes from far-field and its source is located far enough to assume a wavefront as a linear one. The signal is received from direction θ by a uniform linear array (ULA) of M sensors. In order to avoid spatial aliasing distance d between them must be lesser than a signal wavelength λ_c.

The narrowband signal can be simply written as the next time-harmonic dependence

$$s(t) = A(t)\exp(\omega_c t)\,, \tag{1}$$

where $A(t)$ is the baseband signal, ω_c is the signal center angular frequency.

The signal from far-field received by the microphone array can be written as the next vector

$$\mathbf{x}(t,\theta) = \begin{bmatrix} x_1(t,\theta) \\ \vdots \\ x_M(t,\theta) \end{bmatrix} = \begin{bmatrix} e^{j2\pi d\cdot 0\cdot\sin(\theta)/\lambda} \\ \vdots \\ e^{j2\pi d\cdot(M-1)\cdot\sin(\theta)/\lambda} \end{bmatrix} \cdot s(t) + \begin{bmatrix} \eta_{c1}(t) \\ \vdots \\ \eta_{cM}(t) \end{bmatrix} = \tag{2}$$
$$= \boldsymbol{\alpha}(\theta)\cdot s(t) + \boldsymbol{\eta}_c(t).$$

$x_m(t,\theta)$ is a signal captured by mth sensor, $\boldsymbol{\eta}_c(t)$ is a corresponding complex additive noise assumed as a white Gaussian noise with zero mean, $\boldsymbol{\alpha}(\theta)$ is an array manifold vector or the steering vector (Bai, et al., 2013) that depends on the DOA θ. One can see that captured signals $x_m(t)$ have constant phase shift between each other $\omega = kd\sin(\theta)$, $k = 2\pi/\lambda_c$. This shift is a spatial frequency that has to be estimated.

In many real situations the received signal is not complex or sensors record only a real part of it. If we assume that the received signal is a single-tone one with an angular frequency ω_c, amplitude A and is sampled onto N samples with sampling interval τ than in any discrete moment of time $t_n = \tau(n-1)$, $n = \overline{1, N}$ can be described as

$$\mathbf{x}(t_n,\theta) = \begin{bmatrix} A\sin(\omega_c t_n + 0\cdot\omega) \\ \vdots \\ A\sin(\omega_c t_n + (M-1)\cdot\omega) \end{bmatrix} + \begin{bmatrix} \eta_1(t_n) \\ \vdots \\ \eta_M(t_n) \end{bmatrix}. \tag{3}$$

where $\eta_m(t_n)$ are real parts of the noise vector $\boldsymbol{\eta}_c(t)$ in the equation.

The minimal sufficient information for frequency estimation is contained in only the one vector $\mathbf{x}(t_n)$ taken in any arbitrary moment of the discrete time. Hence, for simplicity we can chose the n=1 and write the corresponding measurement vector

$$\mathbf{x}(\theta) = A\sin(\omega(m-1)) + \eta_m\,,\ m = \overline{1, M}\,. \tag{4}$$

For estimation of the spatial frequency in the signal (4) the algorithm considered in the paper of Prokopenko, Omelchuk, Chyrka (2012) is used. It is based on the representation of the noised sinewave signal as an autoregressive moving average model of the second order. The DOA estimation procedure with frequency estimation steps can be summarized as follows:

a) calculation of the signal statistics

$$B(\bar{x}) = \sum_{m=2}^{M-1}\left[(x_{m+1} + x_{m-1})^2 - 2x_m^2\right] \Big/ \left[2\sum_{m=2}^{M-1}(x_{m+1}x_m + x_m x_{m-1})\right]$$

b) obtaining of two values of the autoregressive model parameter

$$\hat{\alpha}_{1(2)} = B(\bar{x}) \pm \sqrt{B(\bar{x})^2 + 2}\,;$$

c) spatial frequency estimation

$$\hat{\omega}_{1(2)} = \arccos\left(\hat{\alpha}_{1(2)}/2\right);$$

d) choice of the value $\hat{\omega}$ that is located in the zone of the method uniqueness $(0, \pi/2)$ that is a working range;

e) final calculation of the direction angle as

$$\theta = \arcsin(\omega\lambda_c / 2\pi d)\,.$$

2.2 Near-field Scenario

In the near-field case the received signal is not a plane wave anymore and can be described as

$$\mathbf{x}(t,\mathbf{r}) = \begin{bmatrix} x_1(t,r_1) \\ \vdots \\ x_M(t,r_M) \end{bmatrix} = \begin{bmatrix} \frac{e^{j2\pi r_1/\lambda}}{r_1} \\ \vdots \\ \frac{e^{j2\pi r_M/\lambda}}{r_M} \end{bmatrix} \cdot s(t) + \begin{bmatrix} \eta_{c1}(t) \\ \vdots \\ \eta_{cM}(t) \end{bmatrix} = \tag{5}$$

$$= \boldsymbol{\alpha}(\theta) \cdot s(t) + \boldsymbol{\eta}_c(t).$$

Here $r_m = |\mathbf{s}_0 - \mathbf{s}_m|$ is a distance between the source point $\mathbf{s}_0$ and the *m*th sensor position $\mathbf{s}_m$.

In the case of a real signal it can be written in the form similar to (4) as

$$\mathbf{x}(\mathbf{r}) = A_m \sin(kr_m) + \eta_m,\; m = \overline{1, M}\,, \tag{6}$$

where $A_m = A/r_m$. In the near-field scenario the signal comes to each sensor from some direction that can be roughly defined as $\sin\theta_m \approx (r_m - r_{m-1})/d$, $m = \overline{2, M}$. The corresponding spatial frequency at the sensor's position is $\omega_m = kd\sin(\theta_m)$. We can estimate these local frequencies and use these angles to find the source position.

A single frequency can be calculated on the basis of at least three data samples, therefore we should take at least three consecutive samples x_{m-1}, x_m, x_{m+1}, $m = \overline{2, M-1}$, assume that on this short interval the signal (6) is sinusoidal and apply mentioned earlier estimation method for obtaining the set of ω_m. Note that frequencies ω_1, ω_M can not be estimated due to limitations of this approach. If one consider this approach in the framework of the array processing, it means splitting of the real array onto several overlapping subarrays. On the other hand, it can be considered as an instantaneous frequency estimation in the running window.

Having a set of local frequencies one can calculate a set of direction angles θ_m and draw several rays from corresponding points like it is illustrated in the Fig. 1.

In this example, the array consists of 31 sensors that give us 29 estimates of angles. Here the black cross in the center illustrates the source position. There are some missed rays at the picture that means that it was impossible to estimate a frequency by the considered algorithm and some beams are pointed far away from the true source position. These facts can be explained by failures and errors of the estimation algorithm due to noise action.

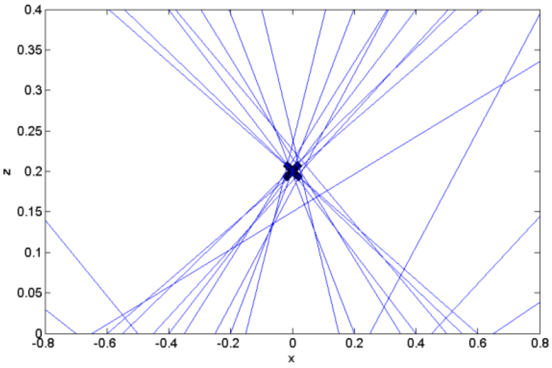

Figure 1: Plot of estimated local directions of arrival for the source located at the boresight.

If we look at the points of rays crossing, we can see that the spatial distribution of them is torn with multiple outliers, but the biggest density is around the true source position. To find the source position the distribution peak position must be estimated as median of all crossing points coordinates.

3 SIMULATION RESULTS

The effectiveness of two proposed approaches was analyzed and the far-field algorithm additionally compared to the ML single-tone estimator and Cramer-Rao lower bound (Rife, Boorstyn, 1974).

Statistical simulations by the Monte-Carlo approach were done under the next conditions: number of sensors in the ULA for a far-field case *N*=11, for near-field *N*=31; due to limited range of the used estimation algorithm, sensors spacing distance is $d = \lambda_c/4$; number of independent runs with single-shot measurements for each plot is 1000.

The first experiments (Fig. 2) represent performance of methods for different directions of arrival under signal-to-noise ratio SNR=20 dB.

One can see that proposed approach works pretty well in the range 0.2–1.4 rad. On the other hand its performance decreases when source is located at boresight or endfire positions because they corresponds to boundaries of the estimation range.

Figure 3 shows performance for different SNR at direction of arrival $\theta = 45°$. One can see, that the proposed method almost reaches the ML-estimator, especially at high SNR.

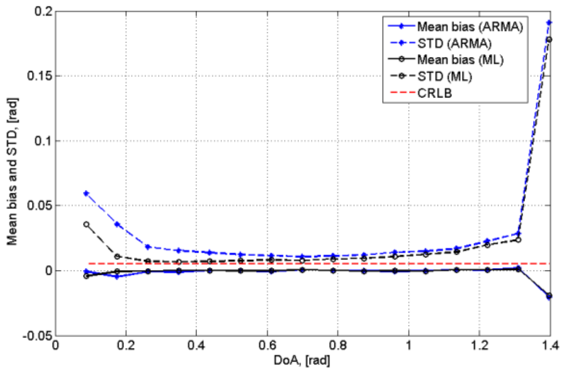

Figure 2: Dependence of the DOA estimation precision of far-field algorithms on the its value.

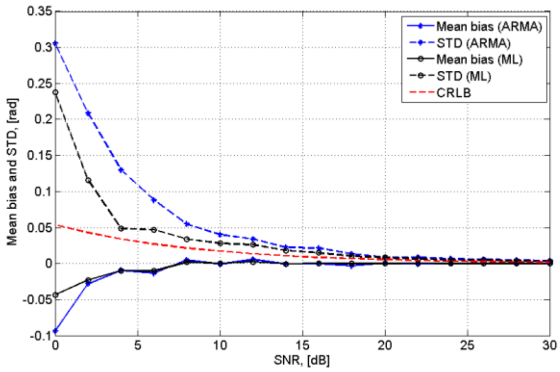

Figure 3: Dependence of the DOA estimation precision of far-field algorithms on the SNR.

The near-field estimation algorithm was analysed under the SNR=20 dB for the source located $2\lambda_c$ forward and $2\lambda_c$ right from the beginning of the ULA. The signal was simulated as a real part of the model (Bai et al., 2013, p. 15). Figures 4 and 5 shows precision indicators for the range and the DOA. One can see, that the proposed approach requires quite high SNR (>30 dB) for decent estimation quality, even with comparatively big number of sensors. This can be explained by the fact that local spatial frequency is estimated only in 3-point running window and under this condition the algorithm is pretty sensitive to the noise action. On the bigger distance to the source using of bigger windows becomes possible and precision increases.

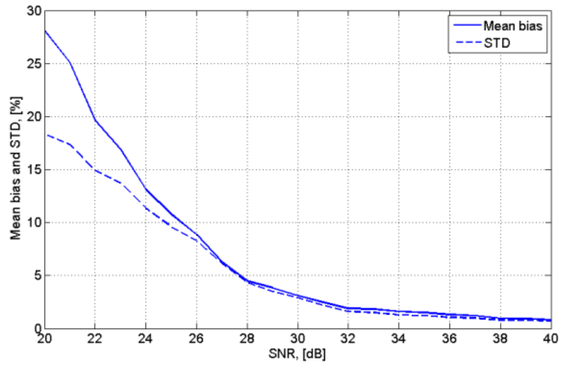

Figure 4: Dependence of the range estimation precision of the near-field algorithm on the SNR.

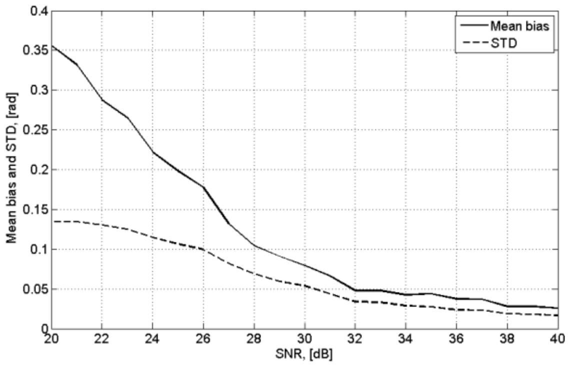

Figure 5: Dependence of the DOA estimation precision of the near-field algorithm on the SNR.

4 CONCLUSIONS

The proposed far-field method shows performance close to the maximum likelihood estimator in the range between boresight and endfire source positions, when SNR is bigger than 5 dB for few sensors. The near-field method generally requires bigger amount of sensors in comparison to far-field method and gives relatively unbiased estimates only at SNR higher than 30 dB.

ACKNOWLEDGEMENTS

The research work reported in the paper was partly supported by the Project AComIn "Advanced Computing for Innovation"; grant 316087, funded by the FP7 Capacity Programme.

REFERENCES

Bai, R. M., Ih, J. G., and Benesty, J. (2013). Acoustic Array Systems. Singapore: John Wiley & Sons Singapore Pte. Ltd.

Chen, C. E., Lorenzelli, F., Hudson, R. E., and Yao, K. (2008). Stochastic maximum-likelihood DOA estimation in the presence of unknown non-uniform noise. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (Las Vegas, USA, March 31 -April 4 2008), ICASSP 2008, 2481–2484.

Gao, F., Gershman, A. B. (2005). A Generalized ESPRIT Approach to Direction-of-Arrival Estimation. IEEE Signal Process. Letters, 12 (3), 254–257.

Grosicki, E., Abed-Meraim, K., and Hua, Y. (2005). A weighted linear prediction method for near-field source localization. IEEE Trans. on Signal Process., 53, 3651–3660.

Johnson, R. C. (2006). Antenna Engineering Handbook (3rd ed.). McGraw-Hill, Inc.

Prokopenko, I. G., Omelchuk, I. P., and Chyrka, Y. D. (2012). Robust frequency estimation. International Radar Symposium (Warsaw, Poland, May 23–25), IRS–2012, 319–321.

Rife D. C., Boorstyn, R. R. (1974). Single-tone parameter estimation from discrete-time observations. IEEE: Trans. on information theory. 5, 591–598.

Stoica, P. and Besson, O. (2000). Maximum likelihood DOA estimation for constant-modulus signal. IEEE Electronics Letters, 36(9), 849–851.

Stoica, P., and Nehorai, A. (1989). MUSIC, Maximum Likelihood, and Cramer-Rao Bound. IEEE Trans. on Acoustics and Signal Process., 37 (5), 720–741.

Swindlehurst A. L., and Kailath T. (1988). Passive direction of arrival and range estimation for near-field sources. IEEE Spec. Est. and Mod. Workshop, 123–128.

Wax, M. (1982) Detection and estimation of superimposed signals, Ph. D. Dissertation, Stanford University.

Wu, Y., Liu, H., and So, H. C. (2009) Fast and accurate direction-of-arrival estimation for a single source. Progress In Electromagnetics Research C, 6, 13–20.

Yang, D., Shi, J., Liu, B. (2009) DOA Estimation for the Near-field Correlated Sources with Interpolated Array Technique. 4th IEEE Conference on Industrial Electronics and Applications (Xi'an, China, 25–27 May 2009), ICIEA 2009, 3011–3015.

Zhi, W., and Chia, M. Y-W. (2007). Near-field source localization via symmetric subarrays. IEEE Signal Process. Letters, 14(6), 409–412.

SPECIAL SESSION ON eHEALTH SERVICES AND TECHNOLOGIES

Modelling Processes in Fractalized Hospitals with Multiagent Systems and Data Analytics

Marc Premm, Martin Riekert, Achim Klein, and Stefan Kirn
Department of Information Systems 2, University of Hohenheim, Stuttgart, Germany
{marc.premm, martin.riekert, achim.klein, stefan.kirn}@uni-hohenheim.de

Keywords: Process Modelling in Hospitals, Fractalized Organizations, Multiagent-based Modelling, Data Analytics

Abstract: Most approaches for modelling processes neglect the high degree of distributed decision making in the hospital domain where processes are coordinated by local authorities. The paradigm of fractal organizations combined with the decentralized characteristics of distributed Artificial Intelligence may help to understand and model the problem. This paper presents ongoing research and contributes a meta-model for modelling processes in hospitals with multiagent systems as fractals of a logistics supply network and incorporates data analytics methods to identify dependencies between different fractals. The presented approach is evaluated by analyzing a hospital scenario involving multiple fractals in a patient-centric process.

1 INTRODUCTION

With a constant change towards profit maximization, hospitals are forced to apply new methods. To cope with the increasing cost pressure, approaches from industrial enterprises seem appropriate and, thus, more and more hospitals start to make use of process-orientation on their internal workflow (Cleven et al. 2014). Applications in hospitals have to face complex relationships and dependencies of several departments that follow locally optimized processes. Clinical pathways are a first approach to cope with the interrelations of multiple departments, but neglect the limited suggestibility of intra-departmental processes (Vanberkel et al. 2010). In hospitals, lasting organizational structures are established that partly show a high degree of autonomy on some levels. As hospital units structure their internal processes without external intervention, these local decisions may influence the inter-department processes and lead to a suboptimal efficacy. To address this issue, the decentralized decision making process of hospitals has to be included in process models to form the basis for analyzing dependencies between multiple departments (i.e. process fractals). A well-suited approach to model and analyse such systems with distributed decision making processes comes with the concept of intelligent, cooperative software agents, and multiagent systems (MAS). We further propose data analytics methods to identify interdependent process fractals and predict time-based parameters to improve cooperation among these.

The goal of this paper is (i) to develop a meta-model for modelling interdependent process fractals that is suitable for scenarios in hospitals and (ii) to incorporate data analytics methods to identify the dependencies between multiple process fractals and to predict execution times as a basis for improved cooperation and higher efficacy. The suggested meta-model is based on two major abstractions: (i) logistics (the right material in the right quantity, at the right time and the right place) as well as (ii) the paradigm of fractal companies introduced by Warnecke (1993). The paper presents ongoing research and is based on previous work (Premm & Kirn 2015).

The remainder is structured as follows. Section 2 discusses related work on process management in hospitals and organizational theory. Section 3 develops a logistics-based meta-model to model fractal processes. Section 4 presents a data analytics approach for identifying process fractal dependencies and predicting execution times. Section 5 presents a scenario-based evaluation. Section 6 concludes.

2 STATE OF THE ART

This section presents the state of the art on (i) process management in hospitals, (ii) organizational paradigms that are used in distributed artificial intelligence and may help to structure processes in hospitals, (iii) the concept of fractal organizational units to better understand hospital processes, and (iv) the systematics of logistics tasks and organizational fractals.

2.1 Process Management in Hospitals

In the past, counteracting delays has been performed by adding additional resources. However, a lack of resources is in many cases not the cause for delays in hospitals, but the organization of inter-departmental processes (Haraden & Resar 2004). Decisions are made on a local basis and the actors are not aware of the consequences on other departments or parallel processes. Haraden and Resar (2004) examine this problem and evaluate processes of several hospitals in the United States and the United Kingdom. The authors focus on elective surgeries as well as the surrounding units and found that an overall view on hospital processes may increase the resource efficiency and thus also the financial performance. However, the paper is restricted to an overview on possible improvements and does not suggest how to enforce processes involving multiple units considering the decentralized character of decision making in hospitals.

Vanberkel et al. (2010) survey similar approaches that encompass multiple departments in hospitals focusing clinical pathways, which aim at eliminating the ambiguity of patient care trajectory. While other modelling approaches optimize all patient types in one department, the scope of clinical pathways is one patient type with all relevant departments. However, this point of view neglects different types of patients competing for the same resources. For Vanberkel et al. (2010), the optimization of clinical pathways is the first step before quantitatively optimizing the internal processes of single departments.

The literature also provides work that specifically addresses the distributed nature of the decision problem: Murray et al. (Murray et al. 2014) take a patient-centric perspective and suggest to use software agents representing relevant actors of patient care trajectory (e.g. patient, physician, unit). The agent interaction protocol takes over the coordination of resources as well as involved actors while coping with the decentralized nature of the processes. The distributed nature of the decision making process as well as the huge dynamic caused by the mixture of planned and emergency instances of numerous individual tasks lead to a high complexity of the process landscape in hospitals. The organizational units as well as the processes in their responsibility can be considered as fractals of the organization. Agent-based approaches from Distributed Artificial Intelligence (DAI show great potential to cope with this complexity.

2.2 Organizational Metaphors in Multiagent Systems

In the last decades, researches in DAI have developed a paradigm called MAS, which is suitable for scenarios with multiple actors deciding about their actions on a local knowledge basis. The main focus of research has been on an increased flexibility facing previously unknown environmental circumstances. However, there are also approaches that cope with organizational stability required in hospital scenarios. According to DAI/MAS researchers with backgrounds in management science, “organization” is a metaphor that can be useful in helping to describe, to study, and to design distributed software systems (Malone 1987; Fox 1981). Compared to organizational theories in management, however, MAS/DAI still lacks similar fine-grained concepts and instruments for describing, analyzing, thus understanding and designing organizational phenomena within agent-based systems.

Approaches from DAI that involve single problem solving experts can be compared with the perspective of management science, in which organizations are systems that pool individual resources in order to gain additional benefits for all of their members. From an organizational perspective this approach implements the concept of dividing labor among a set of individuals each possessing a particular capabilities profile (Gasser 1992). As an immediate consequence, distributed problem solving leads us to role concepts – e.g., the role concept of the C-Net system (Davis & Smith 1983), in which manager and contractor roles coordinate the execution of tasks. However, the definition of roles in DAI is quite different to organizational theories in management science. The latter refers to a role as a precise definition of expected behavior of a particular member of an organization.

Management science considers organizations mainly from a social science perspective. This perspective builds upon the basic assumption that humans form an enterprise in order to fulfill a concrete market demand (e.g., production of autonomous cars). Organizational rules and definitions (e.g., definition of positions) are required to coordinate the division of labor, the behavior of employees, and all operational processes to produce, sell, and maintain goods and services. It is well understood, that enterprises need stability with respect to their suppliers and customers, to their employees, and to their infrastructural, technical and financial production factors. It is well understood, too, that the increased dynamics of their environments (e.g., changing consumer behaviors, changing market demand, changing market structures, changing market coordination, etc.) does also require an increase of organizational flexibility. Approaches from distributed artificial intelligence may serve this kind of flexibility, but lack stability in terms of organizational structures. In the healthcare domain, however, both principles are necessary to fulfill patient care. The paradigm of fractal organizations by Warnecke (1993) enables using these two approaches simultaneously.

2.3 From Fractal Processes to Fractal Organizational Units

It has been argued that the enterprise of the future will be radically decentralized, in order to meet the challenges of the increasing complexity of its environment, and the dynamics of world-wide competition. Decentralization involves the allocation of autonomy, resources, and responsibilities to deeper levels of the organizational hierarchy (for instance, see work of Tapscott & Caston (1993) or Warnecke (1993)). This requires enterprises to replace hierarchical planning by more decentralized concepts of coordination like the MAS paradigm introduced above. In turn, autonomous organizational departments need to exhibit improved capabilities in terms of intelligence and self-reference than they do today. This has given rise to the notion of organizational fractals (Warnecke 1993). Organizational fractals are characterized by the following major criteria (Warnecke 1993): (i) self-organization and self-optimization, (ii) goal orientation, (iii) dynamic, as well as (iv) self-similarity.

The last criterion of self-similarity describes the structural characteristics of the organization as well as the modalities of generating added value. The self-similarity between different fractals enables resource sharing especially for informational resources and thus is especially interesting as it enables to build complex systems on simple and reoccurring modules. In the case of hospitals, one can think of several logistic tasks that have to be fulfilled for patient care. Whereas the patient itself undergoes multiple different process steps that show self-similarity in their internal structure. Findings from logistics may be transferred to the hospital domain and may serve to improve processes in hospitals with their fractal organizations.

2.4 Systematics of Logistics Tasks and Organizational Fractals

Logistics aim at supplying a requesting entity with the right good (quantity and quality), at the right time and the right place at minimal costs. The spatiotemporal transformation of goods is the rudimental capability of logistics systems. The involved processes can be distinguished into the following categories (Pfohl 2004): (i) Core processes of goods flow (transport, transshipment and storage processes), (ii) supporting processes, e.g. packaging processes and (iii) order transmission and processing processes. A generic example from the manufacturing industry would be the storage of a resource (temporal transformation) that has to be prepared for pickup (transshipment), transported to the targeted destination (spatial transformation), prepared for further processing (transshipment), physically adapted (production), again prepared for pickup (transshipment) and so on. This elementary example shows that the core logistics processes occur continually

The widespread visualization as a graph is domain-independent and enables also logistics networks as an extension of a logistics supply chain (Domschke 2008). Dependent on the specific modelling goal, there are numerous approaches for formalizing logistics tasks. Besides business driven approaches like the Architecture of Integrated Information Systems (ARIS), which provides general means for business process modelling (Scheer & Nüttgens 2000) and the Supply Chain Operation Reference (SCOR) Model, which is an industry-independent framework for evaluation and improvement of supply chains (Stewart 1997) a huge range of quantitative decision models exist in literature. Quantitatively parameterized mathematical models are mainly used for planning and decision making, but usually involve only a restricted number of parameters (Scholl 2008). With these mathematical models numerous variants of supply chain optimization problem can be addressed. However, these models generally assume some central designer that is able to enforce a production plan to all instances of the supply chain. In real-world scenarios this is usually not the case, especially in hospital scenarios in which single departments remain highly autonomous in their internal processes.

The organizational fractals involve a maximum degree of local autonomy, self-control, and self-organization skills. Aiming to maximize their local utility (for instance, in terms of profit), organizational fractals decide on their own whether they are willing to cooperate, or to collaborate with other organizational units. There is no direct means by which fractals can be compelled to behave in a certain manner. The single acceptable way to control the behavior of an organizational fractal, or of a group of cooperating fractals, is through designing a globally consistent system of aims and objectives (Warnecke 1993). However, due to bounded rationality, organizations are, in most cases, not able to establish consistent goal hierarchies. Instead, the different goals that exist within an organization are more or less inconsistent, the knowledge about goals and relationships between them remains necessarily incomplete, uncertain, fuzzy, and sometimes even wrong.

Organizational fractals form organizationally stable parts of an enterprise and have well-defined interfaces to their environments. They execute locally well-defined production functions (transformations), and they are supposed to guarantee a maximum of internal stability in terms of, e.g., their operations and processes, their requests for resources, their availability, and their responsiveness. Their flexibility results from their capability to cooperate, and even merge with other fractals in order to create a more complex fractal, if required.

3 MODELLING FRACTALS WITH MULTIAGENT SYSTEMS

To address the complex nature of organizing processes in hospitals, this section combines the paradigm of organizational fractals from management science with MAS from DAI and proposes a meta-model for modelling fractals from a supply network perspective.

3.1 A Fractal Supply Network Perspective

The transportation of goods and the systematics mentioned in section 2.4 are independent of a certain domain and the mentioned types of processes show similar characteristics: Goods have to be transported, handled and stored. In general, this is even independent of the fact, whether the good in question is physical or informational. For information goods the border between these core processes and the order transmission or processing might diminish as no physical good is present. In this case, the core process is an information flow just like the order processes.

Independent of the physical presence of a good, it can be observed that supply chains are in many cases divided into different fractals. These fractal are autonomous and cannot be fully controlled from a macro perspective. Depending on the context, these fractals might be whole enterprises (e.g. in a manufacturing supply chain) or different departments (e.g. in a hospital) that show a certain amount of autonomy. Hence, the overall process cannot be planned in detail against the motivation of the single fractals.

3.2 Multiagent Systems

With its focus on distributed decision making, the paradigm of MAS seems well suited for the local authorities in the hospital domain. Since the emergence of the multiagent paradigm numerous MAS have been developed for various domains, e.g. manufacturing and logistics, and in most cases the design is focused on specific issues (Stockheim et al. 2004). Although developed independently, the different MAS cannot be viewed as separated autarkic systems as they interrelate with each other in many ways. The organizational structure between two or more independently developed MAS usually involves the relations between the represented real world organizations. The technical as well as the organizational question has been addressed by the platform Agent.Enterprise in a logistic scenario (Woelk et al. 2006). Agent.Enterprise is not restricted to intra-organizational value chains already represented by MAS, but integrates multiple instances of these into inter-organizational supply chains. This combination of multiple MAS is called a multi-multiagent system and works cross-organizational. Each MAS remains locally controlled, but obtains features of inter-organizational communication and cooperation to further increase flexibility and decrease costs. In Agent.Enterprise each MAS plans and optimizes its logistic and production processes individually, but informs other systems of unforeseen and potentially disturbing events. On the basis of this information exchange, plans of other MAS may be adapted or inter-organizational contracts may be renegotiated (Woelk et al. 2006).

3.3 Meta-Model

In logistics supply chains one can find different levels of organizational structure, e.g. in a manufacturing supply chain, there are usually different companies that work together for one final good. Thus, we can distinguish between intra- and inter-organizational structures, e.g. the intra-organization structure of a company is embedded into the inter-organizational structure of the supply chain that involves various other companies whose behavior is not controllable, but has to be motivated. Analogously, processes in hospitals are characterized by highly autonomous departments that can only be limitedly controlled by the central hospital process management. This leads to

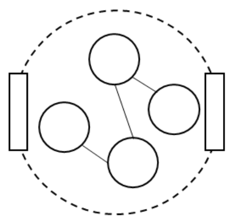

Figure 1: Organizational fractal.

fractal processes within the hospital where each department again can be represented by a single MAS.

Independent of a certain domain, network-wide processes consist of flexibly coordinated fractals being under local control of complex agents, e.g. a single MAS. Two dependent organizational problems evolve: (i) the intra-organizational structure of each MAS that may differ significantly and (ii) the overall inter-organizational structure that aims at a final product and that is not able to fully control the single process fractals. Each fractal has a logistics task based on domain independent types: (i) spatial transformation in form of a transportation process, (ii) temporal transformation in form of storage as well as (iii) physical transformation in form of a production process. The single fractals are represented by a MAS with interfaces to form a supply chain.

The internal workflow of each process fractal is only in a small extent influenceable from an external position. The operational sequences performed by the involved actors may be affected by incentives, but cannot be controlled directly. Hence, for modelling process fractals in logistic processes, it is necessary to have a modelling language that allows to abstract from the internal workflow within a process fractal. Table 1 shows the meta-model for modelling domain independent logistic process that show characteristics of fractalization. Figure 1 shows an example of a process fractal involving the modelling elements described above. The elements are arranged to represent a process fractal with a number of interacting actors and two interfaces.

Table 1: Meta-Model.

Label	Symbol	Description
Process Fractal		A self-contained and self-organized series of activities with a permanent nature that involves a certain number of actors and is available via interfaces
Actor		Smallest organizational entity in a process fractal that has the competency to make decisions with a given scope
Interface		Coupling point of a process fractal that allows for incoming or outgoing products, services or humans from or to another process fractal
Interaction Path		Bidirectional communication link between two actors of a process fractal
Process Flow		Transition of a product, service or human from one process fractal to another one

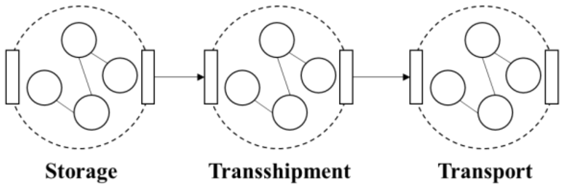

Figure 2: Combination of process fractals.

As described in section 2.4, logistic processes in many domains show self-similarity and can be reduced to three types of processes: (i) storage, (ii) transshipment, and (iii) transport. From a logistic perspective production processes can be interpreted as storage processes, as the product or service has no influence on the logistic system for a certain time and, thus, is transformed in a temporal manner.

3.4 Formalizing Logistics Tasks

The combination of different process fractals is a central feature of the proposed meta-model. The combination of process fractals that are independent from a decision making perspective allows to form logistic supply chains. While each fractal only performs simple tasks the combination of different fractals may serve to solve tasks with a higher complexity. **Error! Reference source not found.** shows an example of a combination of different process fractals: Between transport and storage process fractals, usually, a transshipment process fractal has to be involved to achieve compatibility. In a flow of goods scenario this might be the forklift that allows for transshipping goods in a high-bay warehouse to the transporting truck. However, these process fractals also match for scenarios in a hospital domain, e.g. the patient has to be repositioned (transshipped) from the transportation bed to the surgical table before surgery (see section 5).

These process fractals can be arranged to different kind of processes. Figure 3 shows three elementary types: (i) the single-tier system with only two connected fractals are involved, (ii) the multi-tier system with different interconnected tiers, as well as (iii)

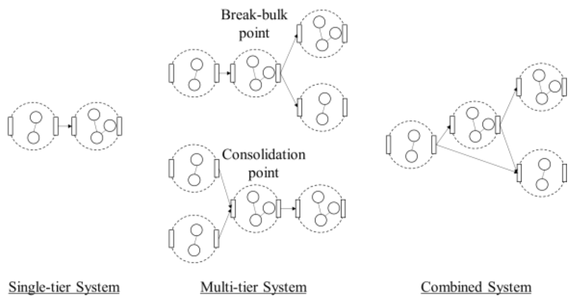

Figure 3: Basic structures of logistics systems.

combined systems that also have connections between non-consecutive tiers. Processes may also split up at break-bulk points and may be joint at consolidation points.

4 DATA ANALYTICS FOR FRACTALIZED PROCESSES

For modelling and better coordinating and supporting logistic processes among fractals with MAS, methods for data analytics can be used to (1) identify fractals in the first place, and (2) predict parameters of the fractals such as the start, duration, and end of individual logistic tasks of different types. The identified fractals can be used for modelling the logistics of an organization with MAS. The predicted parameters can be used by MAS that represent fractals to support and improve coordination among fractals by better anticipating logistic tasks. The most important prerequisite for applying data analytics to the described respect is the availability of large amounts of data that allows describing and predicting the fractals' parameters. Nowadays, this seems less of a problem as more and more data emerges and becomes available due to new types of sensor systems and information systems used in the scope of logistics, e.g. electronic healthcare records and advanced medical devices (Manyika et al. 2011).

For identifying fractals from data, the traces of the logistic tasks within an organization have to be collected and made available for analysis. The data should comprise time stamps and locations of each individual logistic task (i.e., events of starting and completing a logistic task) and a unique reference to the logistic goods across an organization. By sorting the tasks by the time stamps of starting and completing events, the routing of goods can be identified and the duration of tasks can be measured. Aggregating the (most frequent) routes in a graph-based model can help to identify the most important routes and also waiting bottlenecks across fractals can be identified. For conducting this type of data analytics, several software tools are available. For instance, the tool proM can be used (Van Der Aalst et al. 2009).

Nowadays, predicting logistic tasks in an organization is often accomplished by human estimates. Theses are often too coarse-grained and the resulting imprecision leads to bad coordination among fractals and frustration in the implementation of fractalized logistic tasks. Methods for data analytics can be used to more effectively predict all three types of logistic tasks of process fractals.

Predictive data analytics is to create a prediction model in a data-driven way, which maps several predictive variables to the variable to be predicted (here: parameters of logistic tasks). Finding the optimal mapping can be well accomplished by machine learning methods. Machine learning is the ability to improve performance on a task with increasing experience (Mitchell 1997). Performance is measured in terms of the error of the prediction model's output vs. actual outcomes as described in a historic dataset. In the last decade the performance of Machine learning has strongly increased due to the availability of sufficient training data, computational resources and theoretical improvements (Vapnik 2000).

Figure 4 outlines the principle approach of machine learning (Vapnik 2000): the explanatory or predictive input variables created by the *generator* are transformed. The vector transformation makes sure that variables are represented as real numbers. Further types of transformations are also possible that might improve the ability of the method to create an accurate prediction model. The input variables are paired with the variable to predict, which is to be provided by a *supervisor*, e.g. a human annotator. These pairs are used by a so called *learning machine* to create a prediction model, which maps the input variables to the variable to predict $\hat{y}$. With the created model, new data of the input variables can be used to predict the variable of interest.

For the data-driven creation of prediction models, the machine learning method Support Vector Regression (SVR) can be used (Drucker et al. 1997). The SVR method is a Support Vector Machine (SVM; Boser et al. 1992) for regression tasks. The input and output variables are real numbers. But also textual input can be incorporated by means of n-gram based text representations (Joachims 1998). Textual documents are transformed into a vector space representation by means of determining the frequency of occurrence of each n-gram of words within a document and within a whole corpus of documents. Typically, unigrams or bigrams are used.

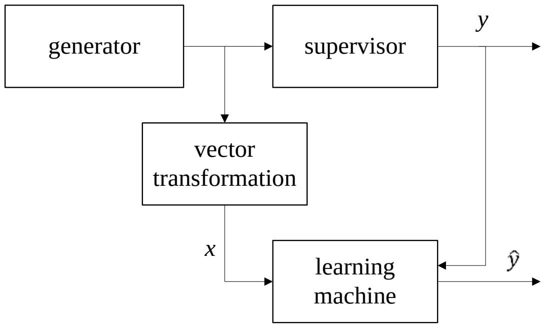

Figure 4: Machine Learning (Vapnik 2000).

The following introduction to SVR is based on Smola & Scholkopf (2004). Given training data $\{(x_1, y_1), \ldots, (x_l, y_l)\} \subset X \times \mathbb{R}$, where X is the input space. SVR determines a function f(x) that is as flat as possible and has at most ε distance from the actual target y_i. To allow a higher distance than ε this algorithm is extended by incorporating a cost parameter (Smola & Scholkopf 2004). For putting in place data analytics for the fractals of an organization, a respective data handling and software architecture is required. The architecture needs to support the desired analytic tasks. Analytics can be conducted either in an offline or online fashion. Offline analytics means to sample large amounts of data, comprising predictive variables and also the variables to predict. The data is used to create the prediction model, which is then applied unchanged on new data. The online approach would try to continuously improve the model once new data becomes available.

5 EVALUATION

Outlined below is a scenario of a patient process, which evaluates the effectiveness of our approach for the improvement of the cooperation among hospital process fractals to improve the overall efficacy. Note that such a process might arise during emergency and regular operations and therefore follows the patterns of reoccurring fractals as described in section 2.4.

The process comprises the following steps: (1) a patient is brought from the ward to the operation section, (2) the patient is moved to a bed in the surgery section, (3) the patient is transported to the operating room, (4) the patient is repositioned to a surgical table, (5) the surgery takes place, (6) the patient is repositioned again to a hospital bed and (7) moved to a postanesthesia recovery.

Figure 5 shows the mapping of the procedural steps into the fractal constructs. The dashed circle represents a fractal, i.e. an autonomously-organized hospital unit. The solid interconnected circles indicate interchangeable agents of the organization. The solid boxes between the fractals represent their interfaces. The procedural steps of the scenario are mapped to following fractal constructs: (1) transport, (2) transshipment, (3) transport, (4) transshipment, (5) storage, (6) transshipment, and (7) transport.

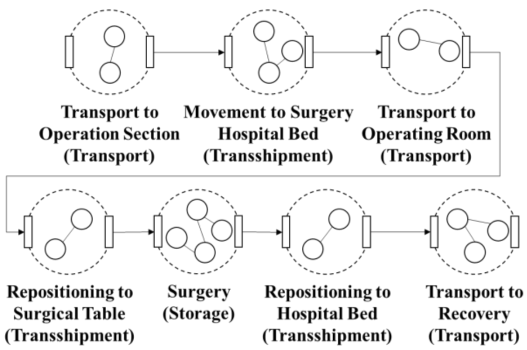

Figure 5: Fractal model of hospital scenario.

In this scenario several problems occur if the prediction of the process time is imprecise. First, all steps are subsequent and therefore an imprecise prediction of the duration of a step will directly suspend the earliest initiation of the following steps. Second, interdependencies of resources like specialized surgeons, medical devices and operating rooms further delay surgeries in this or other operating rooms. Third, due to the previously named problems, the planning of the hospital time is difficult due to the high variance in the actual execution of plans, which leads to the allocation of fewer resources to planning and also decreases commitment of the staff to the plan, which further increases the prediction error.

By means of the fractal based modelling approach it is possible to understand the limits of process planning. One can easily recognize that processes may only be planned on a certain level of abstraction. On a more detailed level, the process execution is always performed by a certain set of involved agents and, thus, can only be indirectly influenced. However, these fractals contain dependencies among each other that have to be recognized to optimize process execution, e.g. the execution time of process steps within a fractal may also be relevant for process steps within other fractals.

The usage of data analytics for the prediction of process times improves the hospital organization by following aspects. First, due to machine learning the start and end time of the fractals can be predicted with low prediction error. Therefore, the planning error is directly reduced. Second, the confidence for the prediction can be estimated. This information allows scheduling surgeries with low prediction confidence in spots that have as few as possible interdependencies with other surgeries. Third, process times can be predicted up to the minute. These predicted process times can be communicated to other affected process fractals without involvement of a human, which allows the automatic updating of process times of emergency and regular surgeries when new information becomes available.

6 CONCLUSION

This research contributes a meta-model for fractalized organizations from a logistics perspective, which is used for modelling hospital processes. The proposed meta-model forms the basis for data analytic methods aiming to identify dependencies between multiple fractals. The contribution has been evaluated by a scenario-based evaluation and is planned to be validated in a field study in future work. However, first results show great potential for modelling hospitals with the paradigm of fractal organizations. With mostly independently organized units, hospitals show a high level of fractalization and, thus, are predestined for modelling processes following the paradigm of organizational fractals.

Together with data analytics focused on hospital needs, the dependencies between different fractals can be identified and parameters of fractals such as process duration can be predicted for the benefit of increasing patient throughput as well as to improve patient care significantly. A detailed investigation will be subject to further research.

ACKNOWLEDGEMENTS

This work has been supported by the project InnOPlan, funded by the German Federal Ministry for Economic Affairs and Energy (BMWi, FKZ 01MD15002).

REFERENCES

Van Der Aalst, W.M.P. et al., 2009. Prom: The process mining toolkit. In *Business Process Management Demonstration Track*.

Boser, B., Guyon, I. & Vapnik, V., 1992. A training algorithm for optimal margin classifiers. In *5th Annual ACM Workshop on Computational Learning Theory*. pp. 144–152.

Cleven, A.K. et al., 2014. Process management in hospitals: an empirically grounded maturity model. *Business Research*, 7, pp.191–216.

Davis, R. & Smith, R.G., 1983. Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence*, 20(1), pp.63–109.

Domschke, W., 2008. *Grundlagen der Betriebswirtschaftslehre – Eine Einführung aus entscheidungsorientierter Sicht* 4th ed., Springer.

Drucker, H. et al., 1997. Support vector regression machines. *Neural Information Processing Systems*, 9, pp.155–161.

Fox, M.S., 1981. An organizational view of distributed systems. *IEEE Transactions on Systems, Man and Cybernetics*, 11(1), pp.70–80.

Gasser, L., 1992. DAI Approaches to Coordination. In N. M. Avouris & L. Gasser, eds. *Distributed Artificial Intelligence: Theory and Practice*. Kluwer, pp. 31–51.

Haraden, C. & Resar, R., 2004. Patient flow in hospitals: understanding and controlling it better. *Frontiers of health services management*, 20(4), pp.3–15.

Joachims, T., 1998. Text categorization with support vector machines: Learning with many relevant features. In *10th European Conference on Machine Learning*. Lecture Notes in Computer Science. pp. 137–142.

Malone, T.W., 1987. Modeling Coordination in Organizations and Markets. *Management Science*, 33(10), pp.1317–1332.

Manyika, J. et al., 2011. *Big data: The next frontier for innovation, competition, and productivity*,

Mitchell, T., 1997. *Machine Learning*, McGraw Hill.

Murray, J., Widmer, T. & Kirn, S., 2014. Agent-based process coordination among hospital units. In *Multikonferenz Wirtschaftsinformatik*. pp. 763–774.

Pfohl, H.-C., 2004. *Logistiksysteme – Betriebswirtschaftliche Grundlagen* 7th ed., Springer.

Premm, M. & Kirn, S., 2015. From Cooperating Agents to Cooperating Multiagent Systems: A Survey. In *Proceedings of the 13th German Conference on Multiagent System Technologies*. Springer.

Scheer, A.-W. & Nüttgens, M., 2000. ARIS Architecture and Reference Models for Business Process Management. In *Business Process Management: Models, Techniques, and Empirical Studies*. pp. 376–389.

Scholl, A., 2008. Grundlagen der modellgestützten Planung. In D. Arnold et al., eds. *Handbuch Logisitk*. Heidelberg: Springer.

Smola, a J. & Scholkopf, B., 2004. A tutorial on support vector regression. *Statistics and Computing*, 14, pp.199–222.

Stewart, G., 1997. Supply Chain Operations Reference Model (SCOR): the First Framework for Integrated Supply-Chain Management. *Logistics Information Management*, 10(2), pp.62–67.

Stockheim, T. et al., 2004. How to build a multi-multi-agent system the agent.enterprise approach. In *ICEIS 2004 - Proceedings of the Sixth International Conference on Enterprise Information Systems*. pp. 364–371.

Tapscott, D. & Caston, A., 1993. *Paradigm Shift: The New Promise of Information Technology*, New York: McGraw Hill.

Vanberkel, P.T. et al., 2010. A Survey of Health Care Models that Encompass Multiple Departments. *International Journal of Health Management and Information*, 1.

Vapnik, V.N., 2000. *The nature of statistical learning theory*, Springer.

Warnecke, H.-J., 1993. *Revolution der Unternehmenskultur* 2nd ed., Heidelberg: Springer.

Woelk, P.-O. et al., 2006. Agent.Enterprise in a Nutshell. In S. Kirn et al., eds. *Multiagent Engineering – Theory and Applications in Enterprises*. Heidelberg: Springer.

Big data in Neurosurgery: Intelligent Support for Brain Tumor Consilium

Karol Kozak[1,2]
[1]*Medical Faculty, Dresden University of Technical, Fetscherstraße 74, D-01307 Dresden, Germany*
[2]*Wrocław University of Economics, Komandorska 118/120, Wrocław, Poland*
karol.kozak@uniklinikum-dresden.de

Keywords: Big data in medicine, Machine learning, Neurosurgery, Consilium, Radiology

Abstract: A brain tumor occurs when abnormal cells form within the brain. Medical imaging plays a central role in the diagnosis of brain tumors. When a brain tumor is diagnosed, a medical team will be formed (consilium) to assess the treatment options presented by the leading surgeon to the patient and his/her family. Using historical evidence-based healthcare data and information directly extracted from images to categorize them may support to increase decision for treatment of patient with brain tumor. Due to its complexity, cancer care is increasingly being dependent on multidisciplinary tumor consilium. That is why it is very important to avoid emotional and quick decisions done by members of consilium. Few studies have investigated how best to organize and run consilium in order to facilitates important decision about patient therapy. We developed and evaluated a multiparametric approach designed to improve the consilium ability to reach treatment decisions. In particular the use of discriminative classification methods such as support vector machines and the use of local brain image meta-data were empirically shown to be important building blocks as support for therapy assign. For efficient classification we used fast SVM classifier with new kernel method.

1 INTRODUCTION

Brain tumors have mainly two types. First is benign tumors are unable of spreading beyond the brain itself. Benign tumors in the brain generally do not essential to be treated and their progress is self-limited. Sometimes they can cause complications because of their position and surgery or radiation can be helpful. And second is malignant tumors are typically called brain cancer. These tumors can extent outside of the brain. Malignant tumors of the brain will always change into a problem if left untreated and a violent approach is almost always warranted. Brain malignancies can be divided into two categories. Primary brain cancer originates in the brain. Secondary or metastatic brain cancer extents to the brain from another site in the body. Cancer arises when cells in the body (in this case brain cells) divide without control. Generally, cells divide in a structured manner. If cells keep separating uncontrollably when new cells are not needed, a mass of tissue forms, called a progress or tumor. The term cancer generally refers to malignant tumors, which can attack nearby tissues and can extent to other parts of the body. A benign tumor does not extent. Last year, an estimated 22,850 adults (12,900 men and 9,950 women) in the United States will be diagnosed with primary cancerous tumors of the brain and spinal cord. It is estimated that 15,320 adults (8,940 men and 6,380 women) will die from this disease this year. About 4,300 children and teens has been diagnosed with a brain or central nervous system in last year. More than half of these are in children younger than 15 (Cancer.net, 2015).

Thanks to the rapid development of modern medical devices and the use of digital systems, more and more medical images are being generated. This has lead to an increase in the demand for automatic methods to index, compare, analyse and annotate them. Care for brain tumors is increasingly complex and often requires specialized expertise from multiple disciplines. Brain tumor consilium reviews provide a multidisciplinary approach to treatment planning that involves doctors from different specialties reviewing and discussing the medical condition and treatment of patients (National Cancer Institute, 2012).

In large university hospitals, several terabytes of new data need to be managed every year. Typically, the databases are accessible only by alphanumeric description and textual meta information through the standard Picture Archiving and Communication System (PACS). Data mining can be defined as the process of finding previously unknown patterns and trends in existing tumor images and using that information to build predictive models for consilium decision support (Kincade, 1998).

Alternatively, it can be defined as the process of data selection and exploration and building models using vast data stores to uncover previously unknown patterns (Milley, 2000).

The underlying Digital Imaging and Communication in Medicine (DICOM) protocol supports only queries based on textual content and limited number of parameters present on the DICOM file and defined by the modality (DICOM, 2011). DICOM files contain their modality as part of the meta-data. We suggest to use that information, feature extraction mechanism can take context into account into the feature extraction and decision support process.

The purpose of our study is to automate extraction of DICOM Metadata from the PACS over patients population for specific brain tumor and support brain tumor consilium in making decision for applied therapy. An Support Vector Machine (SVM) classification technique is proposed to recognize in reasonable malignant and benign MRI brain image from historical database in PACS.

2 METHOD

The first step is to automatically extract dicom images semantic and similarity information and expose that information to a classifier in a very efficient way.

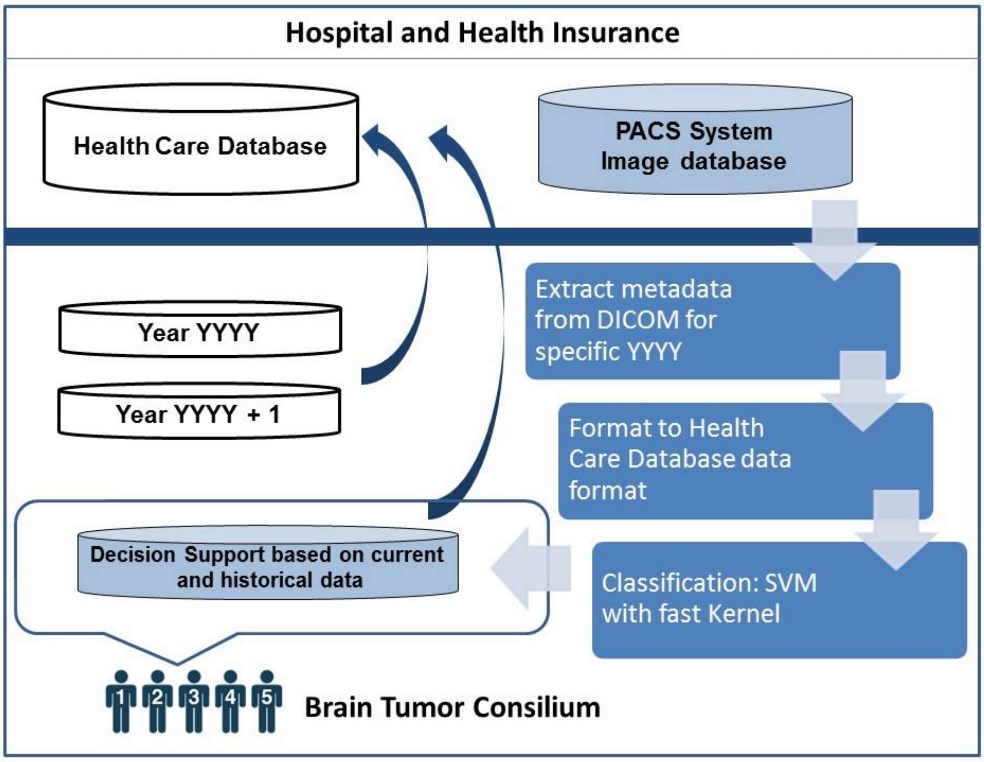

Figure 1. Decision support system for brain tumor consilium. Metadata information from large population of dicom images is analysed by SVM.

The most direct approach to get decision based on images is to match image volume features directly. In this context, content means some property extracted from the image such as color and intensity distribution, texture, shape, or high level features such as the presence of nodes or objects of interest. This approach however is generally not feasible as it may not be clear which volume from one dicom image correspond to which volume in the other image. DICOM objects consist of sets of attribute-value pairs that allow nesting (the values can be other DICOM objects). There are several thousand official attributes, an extension mechanism for private attributes and 27 data types called value representations (VR) for the values (DICOM Part 5, 2011). The data type for each official at tribute is fixed.

Official attributes are identified by a group and element number (16bit unsigned integers usually in hexadecimal notation). Attributes can also represent some kind of real world entity that is only implicitly defined by DICOM or some kind of abstract entity created by the particular hospital. There are important metadata such as pixel parameters, acquisition index, patient dose and geometric information that are generated by the modality and transferred to the PACS database as DICOM metadata.

We have divided metadata into feature sets. General dicom image features, which can be extracted from PACS and can therefore be applied to queries over brain tumor category, and modality specific features. Our concept relies on the automatic extraction of attributes from a dicom image to provide the multiparameters for classifier (Fig. 1).

2.1 Classification

An Support Vector Machine classification technique is proposed to recognize malignant and benign tumors from MRI brain images (meta-data).

a) b)

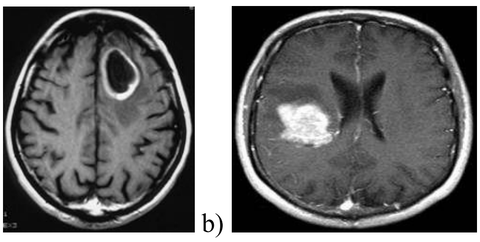

Figure 2: DICOM images of a a) benign and b) malignant brain tumor.

Benign tumors have well defined edges and are more easily removed surgically. Malignant tumors have an irregular border that invades normal tissue with finger-like projections making surgical removal more difficult. Image source: a) http://neurosurgery.ufl.edu and b) http://cdn.phys.org

2.2 Fast SVM

SVM is one of the successful approaches to multiparametric data analysis. In supervised classification we have a set of data samples (each consisting of measurements on a set of variables) with associated labels, the class types (malignant, benign). These are used as exemplars in the classifier design. The classification experiments in dicom analysis were carried out with a support vector machine (SVM) (Vapnik, 1995).

Discriminative approaches to recognition problems often depend on comparing distributions of features, e.g. a kernelized SVM, where the kernel measures the similarity between histograms describing the features. In many practical cases where performance of classification is significant SVM with standard kernel function like Gaussian Kernel (GK) or Radial Basis Function (RBF) are not suitable.

Recently, the use of kernels in learning systems has received considerable attention. The main reason is that kernels allow mapping the data into a high dimensional feature space in order to increase the computational power of linear machines (see for example Vapnik, 1995, 1998, Cristianini and Shawe-Taylor, 2000).

SVM can be optimized for performance via the kernel methods adapted for dicom image datasets. In Kernel methods, the original observations are effectively mapped into a higher dimensional non-linear space. For a given nonlinear mapping, the input data space X can be mapped into the feature space H:

$$\phi : X \rightarrow H \ \ where\ x \rightarrow x : \phi(x). \qquad (1)$$

Linear classification in this non-linear space is then equivalent to non-linear classification in the original space. Require Fisher LDA can be rewritten in terms of dot product.

$$K(x_i, x_j) = \phi(x_i) \bullet \phi(x_j) \qquad (2)$$

Unlike Support Vector Machine (SVM) it doesn't seem the dual problem reveal the kernelized problem naturally. But inspired by the SVM case we make the following key assumption,

$$w = \sum_i \alpha_i \phi(x_i) \tag{3}$$

In terms of new vektor the objective J (α) becomes,

$$\arg\max_{\alpha \in R^n} J(\alpha) = \frac{\alpha^T S_B^{\phi} \alpha}{\alpha^T S_W^{\phi} \alpha} \tag{4}$$

Table 1: Most popular kernels used for SVM classification.

Kernels	Formula
Linear	K(x, x')= x. x'
Sigmoid	K(x, x') = tanh(a x. x'+b)
Polynomial	K(x, x') = (1+ x. x')^d
RBF	K(x, x') = exp(- γ || x - x'||^2)
Gaussian	K(x, x') = exp(- γ || x - x'||)

Table 1 present most popular kernel methods. Correspondingly, a pattern in the original input space R^n is mapped into a potentially much higher dimensional feature vector in the feature space H.

The scatter matrices in kernel space can expressed in terms of the kernel only as follows:

$$S_{\mathbf{B}}^{\phi} = [K_1 K_1^T - KK^T] + [K_2 K_2^T - KK^T] \tag{5}$$

$$S_W^{\phi} = K^2 - (N_1 K_1 K_1^T + N_2 K_2 K_2^T) \tag{6}$$

$$K_1 = \frac{1}{N_1} \sum_{im \in psoitive} K_{im} \qquad K_2 = \frac{1}{N_2} \sum_{i,m \in negative} K_{im} \tag{7}$$

$$K = \frac{1}{N} \sum_{i,j \in N} K_{ij} \tag{8}$$

Popular choice is the Gaussian kernel

$$K(i,j) = \exp\left(-\frac{\| i - j \|^2}{2\sigma^2}\right) \tag{9}$$

with a suitable width of kernel and must σ > 0.

So, we have managed to express the problem in terms of kernels only which is what we were after. Note that since the objective in terms of has exactly the same form as that in terms of w.

In this project the input dicom image is not directly fed into SVM as inputs. Instead, a set of simple features is first extracted from meta-data, and then the features are used as inputs. It will be assumed that each dicom image meta-data set z = {b_1, . . . , b_M} is composed of a set of range-bearing measures b_i = (αi, di) where αi and di are the bearing and range measures, respectively

Each training example for the SVM algorithm is composed by one observation z_i and its classification υ_i. The set of training examples is then given by

E = {(z_i, υ_i) : υ_i ∈ Y = {benign, malignant}} (10)

where Y is the set of classes. In this paper it is assumed that the classes of the training examples are given in advance (benign, malignant). The objective is to learn a classification system that is able to generalize from these training examples and that can later classify day/night in laboratory environment.

Kernel SVMs have become popular for real-time applications as they enjoy both faster training and classification speeds, with significantly less memory requirements than non-linear kernels due to the compact representation of the decision function (Subhransu et. al, 2008). The crossplane kernel, KH_I (t_a, t_b) = n_i=1 min (t_a(i), t_b(i)) is often used as a measurement of similarity between histograms ta and t_b, and because it is positive definite (Odone et.al, 2005) it can be used as a kernel for discriminative classification using SVMs. Recently, crossplane kernel SVMs (call CPSVMs), have been shown to be successful for detection and recognition (Grauman and Darrell, 2005 and 18. Lazebnik et.al, 2006). We based on kernel Intersection Kernel (Subhransu et. al, 2008). Given feature vectors (parameters from DICOM meta-data) of dimension n and learned support vector classifier consisting of m support vectors, the time complexity for classification and space complexity for storing the support vectors of a standard CPSVM classifier is T (p u).

.

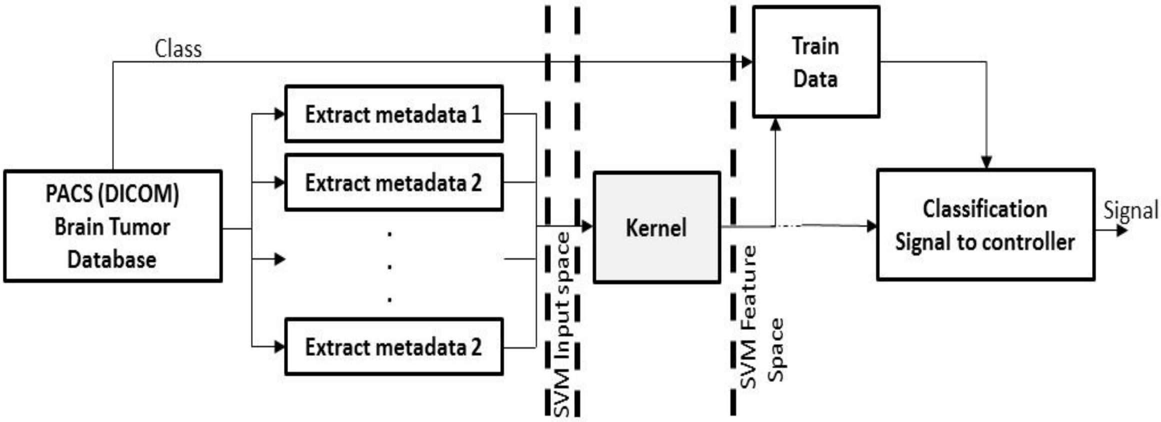

Figure 3: Classification model.

We apply an algorithm for CPSVM classification with time complexity T (u log p) and space complexity T(pu). We then use an approximation scheme whose time and space complexity is T (u), independent of the number of support vectors. The key idea is that for a class of kernels including the crossplane kernel, the classifier can be decomposed as a sum of functions, one for each histogram bin, each of which can be efficiently computed. In dicom anaylsus with thousands of support vectors we also observe speedups up to 2000× and 200× respectively, compared to a standard implementation.

Now we show that it is possible to speed up classification for CPSVMs. For feature vectors x, z $\in R_{+}^{n}$, the crossplane kernel is:

$$K(x,z): K(x,z) = \sum_{i=1}^{n} \min(x(i), z(i)) \quad (11)$$

and classification is based on evaluating:

$$(x) = a_0 + \sum_{i=1}^{m} a_l\, y_l K(x, x_l) + b = \sum_{i=1}^{m} a_l\, y_l \left(\sum_{i=1}^{n} min\big(x(i), x_l(i)\big)\right) + b \quad (12)$$

Thus the complexity of evaluating h(x) in the naive way is O(pu). The trick for crossplane kernels is that we can exchange the summations in equation 12 to obtain:

$$\begin{aligned} h(x) &= a_0 + \sum\nolimits_{l=1}^{m} \alpha_l y_l K(x, x_l) + b = \\ &= \sum_{i=1}^{m} a_l\, y_l \left(\sum_{i=1}^{n} min\big(x(i), x_l(i)\big)\right) + b = \\ &= \sum_{i=1}^{m}\left(\sum_{i=1}^{n} a_l y_l min\big(x(i), x_l(i)\big)\right) + b \\ &= \sum_{i=1}^{n} h_i\big(x(i)\big) \end{aligned} \quad (13)$$

Rewriting the function h(x) as the sum of the individual functions, hi, one for each dimension, where

$$h_i(s) = \sum_{l=1}^{m} \alpha_l y_l \min(s, x_l(i)) \quad (14)$$

So far we have gained nothing as the complexity of computing each hi(s) is T(p) with an overall complexity of computing h(x) still T(pu). We now show how to compute each hi in T(logp) time.

Consider the functions hi(s) for a fixed value of i. Let $\bar{x}_l(i)$ denote the sorted values of $x_l(i)$ in increasing order with corresponding α' s and labels as $\bar{\alpha}_l$ and $\bar{y}_l$.

If $s < \bar{x}_1(i)$ then $h_i(s) = 0$,

otherwise let r be the largest integer such that

$$\bar{x}_r(i) \leq s.$$

Then we have,

$$h_i(s) = \sum_{l=1}^{m} \bar{\alpha}_l \bar{y}_l \min(s, \bar{x}_l(i)) \quad (15)$$

$$= \sum_{1 \leq l \leq r} \bar{\alpha}_l \bar{y}_l \bar{x}_l(\mathrm{i}) + \mathrm{s} \sum_{r \leq l \leq m} \bar{\alpha}_l \bar{y}_l$$

$$= A_i(r) + sB_i(r)$$

Where we have defined,

$$A_i(r) = \sum_{1 \leq l \leq r} \bar{\alpha}_l \bar{y}_l \bar{x}_l(\mathrm{i}), \quad (16)$$

$$B_i(r) = \sum_{r \leq l \leq m} \bar{\alpha}_l \bar{y}_l \quad (17)$$

Equation 17 shows that hi is piecewise linear. Furthermore hi is continuous because:

$$h_i(\bar{x}_{r+1}) = A_i(r) + \bar{x}_{r+1} B_i(r) = A_i(r+1) + \bar{x}_{r+1} B_i(r+1). \quad (18)$$

Notice that the functions A_i and B_i are independent of the input data and depend only on the support vectors and α. Thus, if we precompute $h_i(\bar{x}_r)$ then $h_i(s)$ can be computed by first finding r, the position of $s = \bar{x}_l$ in the sorted list $\bar{x}$ (i) using binary search and linearly interpolating between $h_i(\bar{x}_r)$ and $h_i(\bar{x}_r+1)$. This requires storing the $\bar{x}_l$ as well as the $h_i(\bar{x}_l)$ or twice the storage of the standard implementation. Thus the runtime complexity of computing h(x) is T(u logp) as opposed to T(pu), a speed up of T(u/logp). In our experiments we typically have SVMs with a few thousand support vectors and the resulting speedup is quite significant.

3 RESULTS

In order to show the validity and classification accuracy of our algorithm we performed a series of tests on few dicom benchmark data sets. Data sets are presented in table 2. We tested a proposed extension of Intersection Kernel in experimental datasets from sample collection dicoms. We use the standard SVM algorithm for binary classification described previously. The regularization factor of SVM was fixed to C = 10. In order to see the effect of generalization performance on the size of training data set and model complexity, experiments were carried out by varying the number of training samples (30, 60, 120, 180, 240) according to a 5-fold cross validation evaluation of the generalization error. The data was split into training and test sets and normalized to minimum and maximum feature values (Min-Max) or standard deviation (Std-Dev).

Table 2: DICOM images dataset for astrocytoma brain tumors from demo dataset. Datasets are divided in malignant tumors and benign tumors.

	Total	Training data	Test data
Images	840	420	420
Malignant tumors	260	130	130
Benign tumors	580	290	290

Results for our classifier are presented in table 3.

Table 3. Classification results for two datasets from two patients using two kernel methods with c = 20.

Dataset 1: Malignant tumors

Training set	Kernel RBF	Training Time/ Classification Time	Classification Accuracy	Intersection Kernel	Training Time/ Classification Time	Classification Accuracy
30	C = 20	16 s / 3s	83.6%±6.7	C = 20	11 s /3s	84.7%±1.6
60	C = 20	27 s / 6s	84.2%±2.6	C = 20	12 s / 4s	85.5%±6.7
120	C = 20	35 s / 10s	85.6%±5.5	C = 20	24 s / 10s	86.2%±1.5
180	C = 20	38 s / 18s	87.2%±1.5	C = 20	22 s / 13s	88.3%±4.3
240	C = 20	48 s / 19s	82.4%±3.9	C = 20	33 s / 17s	82.6%±2.5

Dataset 2: Benign tumors

Training set	Kernel RBF	Training Time/ Classification Time	Classification Accuracy		Intersection Kernel	Training Time/ Classification Time	Classification Accuracy
30	C = 20	14 s / 5s	83.6%±6.7		C = 10	12 s /3s	84.4%±2.6
60	C = 20	27 s / 9s	82.3%±2.5		C = 10	16 s / 4s	82.9%±5.5
120	C = 20	32 s / 12s	85.0%±4.4		C = 10	23 s / 11s	87.1%±1.7
180	C = 20	40 s / 15s	83.1%±1.5		C = 10	27 s / 13s	85.2%±3.4
240	C = 20	45 s / 18s	85.5%±2.7		C = 10	35 s / 13s	82.8%±1.6

4 CONCLUSION

The accuracy of the SVM for classifying malignancies was by average 85.4% (28s) and the negative bening tumors predictive value, 83.61% (24s). The SVM proved helpful in the decision based on imaging diagnosis of brain tumor. The classification ability of the SVM with fast Kernel is nearly equal to that of the standard SVM model, but the SVM with fast kernel has a much shorter training and prediction time (1 vs 189 seconds). Given the increasing size and complexity of data sets, the SVM is therefore preferable for computer-aided decision support. Our method has the potential to predict therapy strategy in fast time, saving a significant amount of time to consilium experts, giving them suggestions, enabling them to quickly move from a single observation object image to a set of similar ones, potentially containing historical decisions in therapy. These supporting decisions, when compared to the current patient dicom image, may strengthen the case for the diagnosis or provide the consilium with additional insight.

REFERENCES

(2011) Digital imaging and communications in medicine (DICOM) part 7: Message exchange. Section 9.1.2. National Electrical Manufacturers Association.

(2011) Digital Imaging and Communications in Medicine (DICOM), Part 5: Data Structures and Encoding, Section 6.2 http://medical.nema.org/ Dicom/ 2011/11_05pu.pdf

Cancer Center http://www.cancer.net

CDN Physics: http://cdn.phys.org

Cristianini, N. and J. Shawe-taylor, (2000). An introduction to Support Vector Machines. 200.11/year

Cortes, C. and Vapnik, V. (1995). Support-vector networks. Machine Learning. 213. 94

Grauman, K. and Darrell, T. (2005). The pyramid match kernel: Discriminative classification with sets of image features. ICCV, 2.

Kincade, K. (1998). Data mining: digging for healthcare gold. Insurance & Technology, 23(2), IM2-IM7

Lazebnik, L., Schmid, C. and Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In CVPR.

Milley, A. (2000). Healthcare and data mining. Health Management Technology, 21(8), 44-47

National Cancer Institute. Definition of Tumor Board Review, (2012).

Cancer http://www.cancer.gov/dictionary?cdrid=322893.

Neurosurgery http://neurosurgery.ufl.edu and

Odone, A., Barla, F. , Verri, A. (2005) Building kernels from binary strings for image matching. IEEE T. Image Processing, 14(2):169–180.

Subhransu, Z. M., Berg, A. C. Malik, J. (2008) Classification using Intersection Kernel Support Vector Machines is Efficient. IEEE Computer Vision and Pattern Recognition.

Vapnik, V.N. (1995). The Nature of Statistical Learning theory, Springer Verlag, New York.

Mapping of Terminology Standards
A Way for Interoperability

Sven Van Laere, Frank Verbeke, Frederik Questier, Ronald Buyl and Marc Nyssen
Department of Biostatistics and Medical Informatics – Public Health,
Vrije Universiteit Brussel, Laarbeeklaan 103, 1090 Jette, Belgium
{svvlaere, fverbeke, fquestie, rbuyl, mnyssen}@vub.ac.be

Keywords: Terminology standard, interoperability, eHealth

Abstract: Standards in medicine are essential to enable communication between healthcare providers. These standards can be used either for exchanging information, or for coding and documenting the health status of a patient. In this position paper we focus on the latter, namely terminology standards. However, the multidisciplinary field of medicine makes use of many different standards. We propose to invest in an interoperable electronic health record (EHR) that can be understood by all different levels of health care providers independent of the kind of terminology standard they use. To make this record interoperable, we suggest mapping standards in order to make uniform communication possible. We suggest using mappings between a reference terminology (RT) and other terminology standards. By using this approach we limit the number of mappings that have to be provided. The Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT) can be used as a RT, because of its extensive character and the preserved semantics towards other terminology standards. Moreover, a lot of mappings from SNOMED CT to other standards are already defined previously.

1 INTRODUCTION

In medical practice a lot of standards are used (Gaynor, Myung, Gupta, Rawn, & Moulton, 2008), not only for exchanging information between medical instances, i.e. communication standards, but also for documenting and coding of medical data, i.e. terminology standards. In this paper we will focus on terminology standards and their variety. Different terminology standards are used, even for referring to an identical concept: GP's use the International Classification of Primary Care, 2nd edition (ICPC-2), physicians in hospitals use the Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT), coding teams in hospitals use the International Classification of Diseases and Related Health Problems, version 10 (ICD-10) for reimbursement claims, and so on.

Next to the multidisciplinary use of standards by the various health care providers, we also need to deal with differing structures in these standards. In this article we accept the definition of de Lusignan (2005) who makes a distinction between codes, classifications, terminologies and nomenclatures. De Lusignan defined them as follows:

- Codes assign a label to a certain concept.
- A classification groups concepts together, defined by a common characteristic.
- Terminologies assign labels to a certain domain.
- Nomenclatures assign codes to concepts that can even be combined to constitute new complex concepts.

If we look back in history, medical records were represented only using free text for a long time. The reason codes became important was since text-based retrieval is hard. Not much later, the idea of linking similar clinical ideas together resulted in classifications. Nowadays terminology standards use codes that uniquely represent concepts, e.g. code *A00* represents the cholera disorder using ICD-10 encoding. Moreover this terminology makes use of groupings, e.g. the block of codes represented by code *A00* up to *A09*, represents *Intestinal infectious diseases*, the cholera disorder is thus an intestinal infectious disease.

Classification concepts have a single relationship with their direct predecessor; this can be either an *is a* relationship or a grouping relationship that is more generic. For example, the ICD-10 concept *J20* with description “Acute bronchitis” is the parent of the concept with code *J20.0* with description “Acute bronchitis due to Mycoplasma pneumonia”. In ICD-10, the textual representation of the codes expresses the relationship with their parent. Next to classifications, we have nomenclatures, e.g. SNOMED CT, that also include more specialized relationships to express associations such as laterality, finding site, severity, … Moreover, in SNOMED CT, if there is a concept that is not yet included in the standard, one can also rely on its post-coordinated representation by combining SNOMED CT concepts (Cornet, Nyström, & Karlsson, 2013).

Documentation and coding of problems, diagnoses and treatments is getting more valuable to work towards an electronic health record (EHR) (Dickerson & Sensmeier, 2010). The data of this record should not only be used in a uniform way, it should also be possible to interpret the EHR in the same way, from a GP in his medical practice to a physician in the hospital. In healthcare all providers work together in order to deliver the best care to the patient. However, since in the different levels of healthcare, different terminology standards are being used with different underlying structures, we should address the topic on how to align these standards to make them interoperable.

Often a combination of free text and coded text is used in patient record documents. Working towards optimal use of coding inside the health record will lead to better documentation of the patient’s health status and eventually more appropriate treatment will lead to improvement of a patient's health.

This paper proposes to introduce a bridge between different terminology standards using a reference terminology (RT). This RT must fulfill the requirement of being semantic interoperable with other terminology standards. We will further discuss how we can work towards this interoperability in Section 2. We will provide a discussion in Section 3 and give a conclusion in Section 4.

2 MAPPING STANDARDS

Currently, many care providers have to reenter the same data over and over again. When a patient contacts a GP for a problem and needs to be referred to the hospital, the GP’s data is copied by reentering the data in its appropriate form by the physician in a hospital. Instead of entering data over and over again, we propose to introduce a mapping between different terminology standards.

2.1 Mapping

A mapping is a linkage between a concept from one standard to another standard (see Figure 1) based on the equivalence between the two concepts. This is not only done by comparing the syntactical representation of a concept, i.e. the description of a concept. We propose to make mappings between standards, using the following guidelines:

- Consider the place in the hierarchy of a terminology standard;
- Consider the semantics of a concept;
- Consider the relationship with other concepts around a concept (if appropriate).

When we apply the process of mapping concepts of different terminology standards, we do have two approaches: *manual* mapping or *semi-automatic* mapping. In the manual approach we rely on human knowledge of medical experts, linguists, terminology experts… Or we can use a computer algorithm that queries for candidate mappings based on the lexical representation of a concept, i.e. the description. After the candidate mappings are identified, human review of the automatic mapping is still required to evaluate the found mappings according to the abovementioned guidelines. This approach is thus less time consuming w.r.t. the manual mapping approach where experts manually identify and validate the mappings.

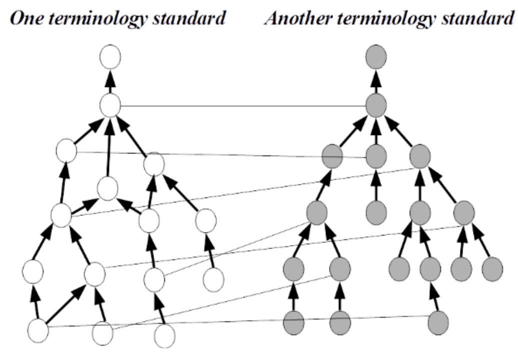

Figure 1: Mapping from one standard to another.

2.2 Semantic interoperability

In healthcare, interoperability is defined as the state which exists between two application entities, when one application entity can accept and understand data from the other and perform that task without the need for extra operator intervention. (Aguilar, 2005) Interoperability can be established at the level of semantics, i.e. semantic interoperability. This means that the information should be understood at the level of domain concepts.

When links between terminology standards are established, we ensure translation of one standard to another is computer processable without losing the aspect of semantics is possible. The degree of semantic interoperability then depends on the level of agreement regarding the terminologies and the conceptual representation behind it.

In the process of communication between a GP and a physician, they both are allowed to use their own proper language (i.e. codes of a terminology standard). By using mappings we can find the equivalent concept in the other terminology standard if a mapping is provided.

2.3 Wide variety of standards

Since a wide variety of medical standards exists and insufficient effort is invested in e-health standard's interoperability, we assume it is worth-while to invest in providing mappings between different standards. We propose to do this in a step by step approach as proposed by the International Health Terminology Standards Development Organisation (IHTSDO, 2014), since this mapping process will be a time-consuming effort. Since clinical terminologies typically consist of several thousands of concept codes per terminology, we propose to provide mappings per medical domain, such as cardiology, nursing, and others instead of mapping the whole terminology at once. We can then evaluate the process of mapping one domain and – if successful – apply it to another domain.

Even if we know the principles for mapping, the problem of the wide variety of medical vocabularies still remains:

- **ICPC-2** classifies patient data and clinical activities in the domain of general/family practice and primary care (Verbeke, Schrans, Deroose, & De Maeseneer, 2006);
- **ICD-10** is a classification used to monitor the incidence and the prevalence of diseases and other health problems (WHO, 2015);
- Logical Observation Identifier Names and Codes (**LOINC**) is the universal standard for identifying medical laboratory observations (McDonald et al., 2003);
- **NANDA** is an international classification used in the nursing domain (Müller-Staub, Lavin, Needham, & van Achterberg, 2007);
- **SNOMED CT** is a nomenclature used for exchanging healthcare information between physicians and other healthcare providers (Donnelly, 2006);
- …

2.4 Reference terminology

If we apply a mapping from each terminology standard to another (see Figure 2a), we end up with $\frac{N\cdot(N-1)}{2}$ mappings that are necessary, where N is the number of standards. We propose the use of a common reference terminology (RT) from which we map to each medical standard (see Figure 2b). The number of mappings needed is then equal to N.

For this reference terminology, it is key to find the terminology that is the most comprehensive in the medical domain, containing also concepts from various domains. SNOMED CT covers more than 310,000 concepts and is likely to be the most appropriate to use as RT. Moreover, SNOMED CT is the only standard providing both pre-coordinated and post-coordinated expressions (Benson, 2010).

Pre-coordination is used when a clinical idea is represented by a single concept id, e.g. *fracture of tibia* is represented by concept id *31978002*. Post-coordinated expressions on the other hand use a combination of concept ids to represent a concept, e.g. *fracture of the left tibia* can be represented as *31978002 : 272741003 = 7771000*, that represents *fracture of tibia : laterality = left*.

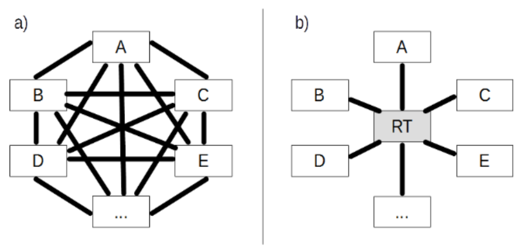

Figure 2a: Mapping each standard to every other standard.
2b: Mapping using a reference terminology (RT)

3 DISCUSSION

In this paper we propose an approach for a more interoperable health record of the patient. We have to explore the possibilities of mappings fully in a certain healthcare domain and evaluate the benefits of it. We propose to use SNOMED CT as a reference terminology (RT).

At the time of writing an EU project, named *ASSESS CT - Assessing SNOMED CT for Large Scale eHealth Deployments in the EU* (ASSESS-CT, 2015), is attempting to evaluate the fitness of the international clinical terminology SNOMED CT as a potential standard for EU-wide e-health deployments. Based on the ASSESS-CT project, an evaluation must be made of the advantages and disadvantages of using SNOMED CT as potential RT standard for the EU. If this fitness evaluation turns out negative, we may need to investigate the possibility of using another RT.

An argument in favor of using SNOMED CT as a RT is that there already exists a lot of mappings from SNOMED CT to other terminology standards:

- ICD-10
- ICD-10-CM (CM indicates a clinical modification of the ICD standard)
- ICD-9-CM
- ICD-O3 (ICD for the oncology domain)
- ICPC-2
- LOINC
- Nursing classifications, such as NANDA, NIC, ICNP, …
- Pharmaceutical classifications such as WHO's ATC and the US National Library of Medicine's RxNorm
- CPT (medical procedure codes)

Another candidate RT is the Unified Medical Language Systems (UMLS) that was designed and is maintained by the National Library of Medicine (NLM) (Humphreys, Lindberg, Schoolman, & Barnett, 1998). UMLS is a collection of vocabularies biomedical health sciences already providing the linkage between them. This system exists of three knowledge sources: the Metathesaurus, the Semantic Network, and the SPECIALIST Lexicon and Lexical Tools. UMLS clusters terms of terminology standards that are equal in one UMLS concept and assigns them a unique id. SNOMED CT is also integrated in this Metathesaurus. Though UMLS does not follow the semantics of SNOMED CT completely (NLM, 2007). As stated by Garla and Brandt (2012) the tool support for using UMLS with respect to SNOMED CT is more robust, whereas semantic differences within UMLS may affect the accuracy of similarity measures. Since the semantics are of utmost importance, we opt to use SNOMED CT instead.

If we use mappings between terminology standards, these mappings are not always bidirectional in use: if a mapping between two concepts of two terminology standards does exist, this is not necessarily the case in reverse. For example the map from SNOMED CT to ICD-10, cannot be reversed since it is common for many details, and different SNOMED CT concepts map to a single ICD-10 code. Reversing the map is not possible since one ICD-10 code would refer to many different SNOMED CT concepts.

We believe working together towards a more integrated EHR, based on a RT, will benefit to the care of patient. The inter-terminology mapping should preferably be an automated background process that is transparent to the health care provider or EHR user and should not interfere with the routine clinical documentation practice.

Since the RT will be used in the shared EHR, the semantics will be implied by the RT. Moreover, by making use of the mappings, care providers will always be able to view the content using the terminology standard that was originally used. Eventually, more extensive use of a RT will also create new clinical decision support opportunities leading to better patient care.

4 CONCLUSIONS

For recording of information in health care, a combination of free text and coded text is often used. In order to improve information sharing for the purpose of patient care or for the management of the hospital, we should invest in mechanisms enabling full and transparent use of coded information in the health record. Most service providers already use one or more terminology standards. However, across different service providers different standards are used. Therefore, sharing information and knowledge about the patient often does not happen in an interoperable way.

This paper proposes a reference terminology based mapping approach in order to meet this requirement. A reference terminology (RT) has the advantage of limiting the number of mappings that must be made. The proposed RT is SNOMED CT, because it is the most extensive medical terminology in use, it supports both pre- and post-coordination and the semantics are preserved with respect to other terminology standards. Another reason for choosing SNOMED CT is the amount of resources that are available. There already do exist a lot of mappings from SNOMED CT to other terminology standards, respecting the accuracy of similarity metrics between different terminology standards.

Sufficient effort should be invested in making the mappings database more complete. This is a time consuming process and therefore a step by step approach is suggested. Start by testing the idea in one domain and then apply it in another one. Eventually this will lead to a shared EHR ensuring interoperability between care providers. Large collections of structured data related to patient health status and health care provider activity can ultimately contribute to EHR systems capable of providing clinical decision support.

REFERENCES

Aguilar, A. 2005. Semantic Interoperability in the Context of e-Health. In *CDH Seminar '05, Galway, Ireland*.

ASSESS-CT. 2015. *ASSESS CT - Assessing SNOMED CT for Large Scale eHealth Deployments in the EU*. Available from: http://www.assess-ct.eu

Benson, T. 2010. SNOMED CT. In Benson, T. (ed.), *Principles of Health Interoperability HL7 and SNOMED* (pp. 189–215). doi:10.1007/978-1-84882-803-2

Cornet, R., Nyström, M., & Karlsson, D. 2013. User-directed coordination in SNOMED CT. In C. U. Lehmann, E. Ammenwerth, & C. Nøhr (Eds.), *Studies in Health Technology and Informatics* (Vol. 192, pp. 72–76). doi:10.3233/978-1-61499-289-9-72

De Lusignan, S. 2005. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Informatics in Primary Care, 13(1), 65-70*. Retrieved from: http://www.ingentaconnect.com/ content/bcs/ipc

Donnely, K. 2006. SNOMED-CT: The advanced terminology and coding system for eHealth. *Studies in Health Technology and Informatics, 121, 279-290*.

Dickerson, A.E. & Sensmeier, J. 2010. Sharing data to ensure continuity of care. *Nursing Management, 41(7), 19-22.* doi:10.1097/01.NUMA.0000384142.04693.58

Garla, V.N. & Brandt, C. 2012. Semantic similarity in the biomedical domain: an evaluation across knowledge sources. *BMC Bioinformatics, 13, 261-273.* doi:10.1186/1471-2105-13-261

Gaynor, M., Myung, D., Gupta, A., Rawn, J., & Moulton, S. 2008. Interoperability of medical applications and devices. In R. H. Sprague Jr. (Ed.), *Proceedings of the 41st Hawaii International Conference on Systems Science, 7-10 January 2008, Waikoloa, Big Island, HI, USA.* doi:10.1109/HICSS.2008.217

Humphreys, B. L., Lindberg, D. A. B., Schoolman, H. M., & Barnett, G. O. 1998. The Unified Medical Language System: An Informatics Research Collaboration. *Journal of the American Medical Informatics Association, 5(1), 1-11*. doi:10.1136/jamia.1998.0050001

IHTSDO. 2014. *Mapping to SNOMED CT. Discussion and Guidance on Mapping Based Solutions Design for Migration or Transformation*. Available from https://csfe.aceworkspace.net/sf/go/doc10065

McDonald, C.J. et al. 2003. LOINC, a Universal Standard for Identifying Laboratory Observations: A 5-Year Update. *Clinical Chemistry, 49(4), 624-633.* doi:10.1373/49.4.624

Müller-Staub, M., Lavin, M. A., Needham, I., & van Achterberg, T. 2007. Meeting the criteria of a nursing diagnosis classification: Evaluation of ICNP®, ICF, NANDA and ZEFP. *International Journal of Nursing Studies, 44(5), 702-713*. doi:10.1016/j.ijnurstu.2006.02.001

NLM. 2007. *Insertion of SNOMED CT® into the UMLS® Metathesaurus®: Explanatory Notes*. Available from https://www.nlm.nih.gov/research/ umls/Snomed/snomed_represented.html

WHO. 2015. *International Classification of Diseases (ICD)*. Available from http://www.who.int/ classifications/icd/en/.

Verbeke, M., Schrans, D., Deroose, S., & De Maeseneer, J. 2006. The International Classification of Primary Care (ICPC-2): an essential tool in the EPR of the GP. *Studies in Health Technology and Informatics, 124, 809–814*. doi: 10.3233/978-1-58603-647-8-809

Interoperability Within E-Health Arena

Darko Gvozdanović[1], Mario Kovač[2], Hrvoje Mlinarić[2], Leon Dragić[2], Igor Piljić[2], Petar Frančec[2], Martin Žagar[2] and Vlado Sruk[2]
[1]Ericsson Nikola Tesla d.d. Krapinska 45, Zagreb, Croatia
[2]Faculty of Electrical Enginnering and Computing, Unska 3, Zagreb, Croatia
darko.gvozdanovic@ericsson.com, mario.kovac@fer.hr; hrvoje.mlinaric@fer.hr; leon.dragic@fer.hr; igor.piljic@fer.hr; petar.francek@fer.hr; martin.zagar@fer.hr; vlado.sruk@fer.hr

Keywords: Personal Healthcare Record, Continuity of Care Document CCD, Electronic Healthcare Record (EHR), Electronic Medical Record (EMR), Interoperability, Information systems in healthcare

Abstract: Integrated care approach and the broader view on a patient's care is something that today's healthcare systems thrive for. Medical information collected from many disparate sources, accessed by authorized users through Electronic healthcare record (EHR) is enabling technology behind. This article gives overview of different interoperability aspects related to data exchange and maps it to usual healthcare business processes. It also comments HL7 CDA being one of today's widely used standards for clinical documents exchange. One concrete approach to Personal Healthcare Record (PHR) to EHR integration using HL7 Continuity of Care Document (CCD) is described.

1 INTRODUCTION

The fact that ICT systems can bring a lot of benefits to all stakeholders within healthcare system is well known (Dobrev, 2009). However, successful implementation and proper introduction of such ICT systems in existing healthcare environment is long and expensive process. Many countries can't increase efficiency in healthcare sector through ICT system introduction. One of three main reasons for this is lack of commonly defined and consistently implemented standards (OECD 2010). Without common standards, one of the biggest advantages introduced by ICT systems in healthcare, which is access to comprehensive and high quality patient medical information in any time or place, remains unsolvable puzzle. Due to the fact that medical data originates from many disparate sources, efficient sharing across organizations, administrative domains or even countries is of utmost importance. Interoperability of implemented ICT systems plays vital role in achieving this goal.

Stroetmann et al. (2009) defined interoperability in healthcare context "as the ability, facilitated by ICT applications and systems: to exchange, understand and act on citizens/patients and other health-related information and knowledge; among linguistically and culturally disparate health professionals, patients and other actors and organizations; within and across health system jurisdictions in a collaborative manner."

In order to meet these requirements, all interoperability aspects, namely legal, organizational, technical (eHealth Governance Initiative, 2010) and semantic (EN13606 Association, 2015) must be addressed (Kovac, 2014). A real life interoperability issues are showed in the example that follows.

Ana is a 22 years old female without any chronic disease. On January 23rd she woke up and felt pain in right side of abdomen. She decided to book an appointment with Dr. Henry Levin, her general physician through the patient portal.

On the date of the scheduled appointment Ana went to the polyclinic to see Dr. Levin. Admission office staff (AO staff) checked if there were any previous visits of Ana in the system using her identity card. AO staff found general data about her in the system (from her previous visits to the same polyclinic.

Since Ana had set access right level for her medical data to “ask patient consent each time” through the patient portal, AO staff couldn’t access medical data and asked Ana to provide consent. Ana refused to provide a consent and signed the document stating that only dr. Levin can access her medical data.

Dr. Levin was logged in to his Hospital Information System (HIS) and chose to review Ana’s electronic medical record (EMR). He decided to check her physical status immediately, opened a new case in the system and invited Ana to step into his office.

Even before physician saw the patient, number of interoperability issues had emerged. The first one was legal - who is the owner of medical information: patient or the physician who generated it; how the access rights are managed; can patient choose which part of medical record will be accessible to medical staff. The second issue is of organizational nature. Healthcare institution might have one central reception, one reception per clinic or completely distributed one. Reception process might be completely administrative where no medical information is needed or it can include triage, taking anamnesis and status in which case access to previous medical information is mandatory.

Dr. Levin noted Ana’s anamnesis, physical status and result of his observation in the system and issued several requests for laboratory tests and additional consultations from the surgeon and gynecologist.

Unless Dr. Levin understands the data within Ana’s electronic medical record (EMR) in the same way that all those users who put information into EMR have wanted, the whole concept of EMR is missed. Having information in free text form is definitely better than having nothing but medical data stored in a structured format can be used for automatic alerting on drug-drug interactions, provision of drug-diagnose contraindications, automatic suggestion of applicable clinical practice guidelines (CPG), automatic reporting, reducing administrative work etc. How clinical documents are structured, what coding systems are used, whether the same or different codes for the same notions are used, are only part of semantic interoperability aspect than needs to be taken care of.

Laboratory order was available through the Laboratory Information System (LIS) at the same moment Dr. Levin sent it through HIS.

Since completely new information system (LIS) appeared in storyboard, technical interoperability issue emerged. Legal aspect returned and became even harder to address since laboratory personnel actually did have Ana in their care but never met her. Ana didn’t give consent to all personnel within hospital to access her medical record so important information that might affect laboratory results was not available for laboratory staff. Since physician and laboratory technician / biochemistry engineer were of different specializations and were using different applications, there must not be any misunderstanding of what test were requested and what results were sent back. Did all healthcare professionals use same coding list or at least some mapping engine (terminology server) existed?

The storyboard ends here since majority of interoperability issue types within one healthcare institution were already mentioned although only three steps were exercised: admission, first examination, and referral to laboratory/consultation. In practice, stakeholders within same institution can efficiently share data because they use the same application or some proprietary integration is done if multiple applications exist. But if Ana were urgently referred to another hospital because of suspected acute appendicitis she would be admitted to the hospital with different internal processes, different specialization and HIS from another vendor. These two healthcare institutions were connected only through national infrastructure if it existed. This means that if integrated care is to be supported, proprietary integration that is possible within one institution has to be properly handled through solving all the interoperability issues mentioned above.

2 E-HEALTH BLOCKS

2.1 Electronic Medical Record

While introducing ICT into the healthcare institutions, system Purchasers (not the users themselves) often prioritize administrative over medical processes. True value of information systems / applications in healthcare provision processes is proper management of medical information. Therefore medical documentation module should be the core of every application within general practitioner and specialist practice application or hospital information system. This core component is called electronic medical record (EMR). In simple terms, EMR is a digital version of the paper charts in clinician offices, clinics, and hospitals (Health information technology, 2015). Market today witnesses thousands of EMR systems. Unfortunately lack of standards led to the situation where most of them are implemented on different information models that are followed by the completely different graphical user interfaces, different ways how data entry is supported and completely different application logic.

2.2 Electronic Healthcare Record

Today's healthcare challenges are numerous and there is myriad of ways how healthcare authorities try to address them. One important tendency in coping with these issues is shift towards so called integrated care. The core of such an approach is broader view on a patient's care. This requires boundaries among multiple EMR's to vanish and much more data about the patient made accessible than it is collected in any single healthcare provider's office. The solution for this is electronic healthcare record (EHR).

EHR contains information generated by all the clinicians involved in a patient's care process, with all these clinicians having also access to it. EHR also shares information with other health care providers, such as laboratories and pharmacies. EHR should be pervasive and follow patients – to the specialist, the hospital, pharmacy, the nursing home, within or out of the country (Health information technology, 2015). Secondary use of information stored in EHR, namely education, research, public health needs etc. is as equally important as its primary continuity of care purpose.

2.3 Personal Healthcare Record

The implementation of different eHealth services brings numerous benefits to the patients even when they do not use the service directly. Example is any service that saves time for the physician, allowing him to spend more time with patients. Nevertheless, final touch on the national eHealth system would be direct patient empowerment where patient portals and personal health records (PHR) play vital role. Personal health records contain the same types of information as EHR – diagnoses, medications, immunizations, family medical histories, and provider contact information, but are designed to be set up, accessed, and managed by patients. Patients can use PHR to maintain and manage their health information in a private, secure, and confidential environment. PHR can include information from a variety of sources including clinicians, home monitoring devices, and patients themselves (Health information technology, 2015).

3 STANDARDIZATION

In order to efficiently use medical information throughout healthcare system, it has to be stored and exchanged in a standardized way. EMR, EHR and PHR in their essence are about documenting different facts. If document is intended for personal use only, than words, grammar and rules are not so important. But if document is intended for use by other persons, all of this must be well defined and collectively accepted. Otherwise, document will be at least partly incomprehensible or what is even worse wrongly understood. In the world of semantic interoperability notion grammar refers to reference model, words/dictionary are codes/coding system and phrases/rules are clinical models, archetypes or templates.

In that sense openEHR and HL7 Clinical Document Architecture (CDA) are two of the most promising standards for storing clinical information and medical documents exchange respectively. Integrating the Healthcare Enterprise (IHE) initiative is the most prominent way to achieve out-of-the-box interoperability at least in specific use cases.

3.1 HL7 CDA

The HL7 CDA is a document markup standard that specifies the structure and semantics of "clinical documents" for the purpose of exchange. A clinical document is a documentation of clinical observations and services, with the following characteristics: persistence, stewardship, potential for authentication, context, wholeness and human readability. A CDA document is defined as a complete information object that can include text, images, sounds, and other multimedia content.

HL7 CDA standard proved to be too generic. In order to refine it, content templates are introduced. One of the most widely known content templates is Continuity of Care Document (CCD). CCD is specification on how to constraint HL7 CDA in accordance with requirements set forward in Standard Specification for Continuity of Care Record (CCR). The CCR is a core data set of the most relevant administrative, demographic, and clinical information facts about a patient's healthcare, covering one or more healthcare encounters (Health Level Seven International, 2007). It provides a means for one healthcare practitioner, system, or setting to aggregate all of the pertinent data about a patient and forward it to another practitioner, system, or setting to support the continuity of care. The primary use case for the CCR is to provide a snapshot in time containing the pertinent clinical, demographic, and administrative data for a specific patient.

Although templates obviously refine underlying standards, one obvious weakness emerges – too many different templates defined by different organizations/vendors/health authorities. Even after content is defined with content standard and refined and constrained with standard templates, overlapping terminologies issue remains. Very representative example is HL7 CDA representation of observation of 108 mg/dL glucose in the plasma of a patient, which is measured in a laboratory setting. There are more alternatives how to exchange this fact within CCD document.

Alternative 1 is that plasma glucose measurement procedure is exercised (SNOMED CT code 119958019) and there was an observation of blood glucose status (SNOMED CT code 405176005), with the actual observed value, which is 108 mg/dL glucose. Alternative 2 is that laboratory test procedure is exercised (SNOMED CT code 15220000), and there was an observation of glucose in serum or plasma (LOINC code 2345-7), with the actual observed value, which is 108 mg/dL glucose.

Although different coding systems and different structure is used, the same medical information is represented and communicated in both instances. So in spite of the fact that communicating applications are capable of using CCD template, interoperability is achieved only partially. When different terminology systems are used in the same structure, it is necessary to semantically mediate them for interoperation. Some of the repositories with mapping information are Unified Medical Language System (UMLS) and Metathesaurus and BioPortal.

There are examples of successful eHealth systems that do not use international terminologies. National information system in Croatia (CEZIH) does not use nor SNOMED CT or LOINC. Local coding systems are defined by professional associations. Since there is national consensus about coding lists used, interoperability on national level is achieved and Croatian eHealth system is perceived as one of the best in Europe.

4 INTEROPERABILITY

The practical approach to solving interoperability problem is one of the goals of project “Information and communication technology for generic and energy-efficient communication solutions with application in e-/m-health” (ICTGEN). In scope of this project we will demonstrate integration of PHR with EHR using HL7 Continuity of Care Document (CCD). Simulation environment, consisting of PHR and EHR, was created at Faculty of Electrical Engineering and Computing. As an EHR system openEMR solution based on openEHR reference model is hosted and adapted to specific needs of the ICTGEN project. Project partner, Ericsson Nikola Tesla d.d. provided their own solution for PHR, Ericsson Mobile Health (EMH).

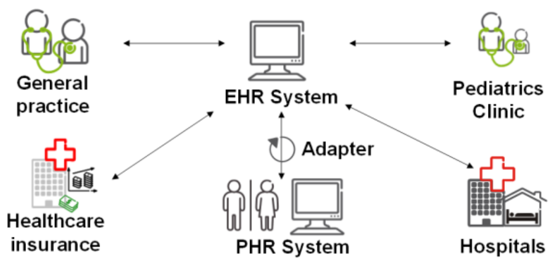

Figure 1: EHR-PHR integration within ICTGEN project.

EMH is one of the numerous PHR solutions offered on the market and its focus is on managing patient's record. Depending on the role, users can access and manage their medical data. EMH provides external access to specific data through Medical Node (MN) API in custom format. Without standardized format of data exchange, integration with any other ICT solution requires system modifications or additional integration components.

First step to solving this interoperability issue was thorough analysis of exchanged medical data format and HL7 Continuity of Care Document (Health Level Seven International, 2007). The analysis lead to classification of medical data into matching categories which were mapped to corresponding CCD elements in the next step. This mapping model was implemented as an adapter component connected to MN API. Since lot of data from PHR is not suitable for EHR, only EHR to PHR communication is implemented. After PHR client is authenticated and authorized for data access, adapter on PHR side receives data from EHR formatted as CCD document. That allows multiple PHR solutions capable of importing CCD to be integrated with EHR. In this project, specific adapter is built and information from CCD document is extracted and stored within EMH database. In that sense EMH is upgraded into interoperable PHR solution capable of importing patient summaries in CCD format presumed limited subset of medical information is exchanged. Although this might seem like unacceptable limitation it is in fact the only realistic way to achieve interoperability. With more than 600.000 concepts within SNOMED CT, it would be illusion to build application that can interpret any of these in the right context. Our approach is to start small and expand adapter making it capable to process more medical information.

5 CONCLUSION

Retrieving all relevant information, utilizing other experiences, exercising team work and looking on things from different perspectives are all aspects of providing high quality healthcare service. Communication is foundation for all of this. The necessary precondition however, is that sender and receiver of information are capable of exchanging it and understanding it in the same way. This is the essence of interoperability. Healthcare, being one of the most complex human domains, poses similarly complex interoperability issues. It actually requires from the participants speaking different complex languages to use one common grammar, words from the same dictionary, to use same phrases or to find one translator that knows all the languages. Neither of this is realistic, especially in short term. Therefore different healthcare interoperability standards and initiatives are introduced but for the time being solution is far away. HL7 CDA without templates is too generic to assure true interoperability. It only allows that clinical documents can be exchanged with appropriate amount of metadata. What's within these documents is not so important to this standard. Templates and constraints narrow this uncertainty a little bit. But even with CCD as one template, same thing can still be expressed in more than one way which makes it very hard for applications to communicate among each other. Nevertheless, ICTGEN project, confirmed that for well-defined subsets, medical information can be efficiently exchanged between different applications like EHR and PHR.

But the main interoperability issue as we see it, is the fact that healthcare professionals does not use same dictionaries (or terminologies / coding list), nor they use the same words (codes) for same events. SNOMED CT as maybe the most comprehensive terminology today is not available in all languages. Mappings to other terminologies are not available at all or are not complete. Process of introducing terminology like SNOMED CT into healthcare system of one country is very long and expensive. Still it does not guarantee that same event will be described with the same code by different healthcare professionals. Until this is solved, no structure, no clinical document definition, no knowledge model (archetype) will bring true interoperability.

The research leading to these results has received funding from the project "Information and communication technology for generic and energy-efficient communication solutions with application in e-/m-health (ICTGEN)" and co-financed by the European Union from the European Regional Development Fund and from project “Carewell” funded by the European Commission within the ICT Policy Support Programme of the Competitiveness and Innovation Framework Programme (CIP).

REFERENCES

Chute, C.G., 2000. Clinical classification and terminology: some history and current observations. J Am Med Inform Assoc 7, 298–303.

Dobrev, A. and Jones, T. (2009) *Report on The socio-economic impact of interoperable electronic health record (EHR) and ePrescribing systems in Europe and beyond – Final study report*. Retrieved from: http://www.ehr-impact.eu/downloads/reports.html

eHealth Governance Initiative – eHGI (2012), Discussion paper on semantic and technical interoperability Retrieved from: http://ec.europa.eu/health/ehealth/docs/ev_20121107_wd02_en.pdf

EN13606 Association, Semantic interoperability of health information. Retrieved June 9, 2015 from: http://www.en13606.org/the-ceniso-en13606-standard/semantic-interoperability

Health information technology (health IT), Retrieved June 9, 2015 from: http://www.healthit.gov

Health Level Seven International (2007), HL7 Implementation Guide: CDA Release 2 – Continuity of Care Document (CCD). Retrieved June 10, 2015 from: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6

Kovac, M. 2014. E-Health Demystified: An E-Government Showcase. Computer 47, 34–42. doi:10.1109/MC.2014.282

OECD, 2010. Improving Health Sector Efficiency. Organisation for Economic Co-operation and Development, Paris.

Stroetmann, V.N.; Kalra, D.; Lewalle, P.; Rodrigues, J.M.; Stroetmann, K. A.; Surjan, G.; Ustun, B.; Virtanen, M.; Zanstra, P. E (2009). *Semantic Interoperability for Better Health and Safer Healthcare: Research and Development Roadmap for Europe: SemanticHEALTH Report*. Retrieved from: http://www.empirica.com/publikationen/2009_en.htm

The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition, 2000. . IEEE Std 100-2000 -. doi:10.1109/IEEESTD.2000.322230

Bulgarian E-Health Overview

Peter Stanchev[1, 2] and Elvira Foteva[3]
[1] *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria*
[2] *Kettering University, Flint, USA*
[3] *Ministry of Health, Sofia, Bulgaria*
pstanche@kettering.edu, efoteva@mh.government.bg

Keywords: E-Health, World Health Organization, Bulgarian Ministry of Health, Health Information Systems, E-Health card

Abstract: The paper presents: (1) the strategy of the World Health Organization and ways of delivering E-Health; (2) discusses the EU Health 2020 policy, EU support, EU projects, goals in the field of E-Health; (3) Bulgarian National Health Strategy (2014-2020), integration and connectivity in healthcare through the national health information system and real access to patient information about their own health; Bulgarian E-Health electronic medical systems, Bulgaria E-Health foundation.

1 INTRODUCTION

According to the World Health Organization (WHO) (World Health Organization, 2015) E-Health is the transfer of health resources and health care by electronic means. It encompasses three main areas:

- The delivery of health information to health professionals and health consumers, through the Internet and telecommunications.
- Using the power of information technology and e-commerce to improve public health services, e.g. through the education and training of health workers.
- The use of e-commerce and e-business practices in health systems management.

The benefits of E-Health for patients include: easier access, better quality, speed and convenience, awareness and participation in control; and for doctors: reducing the administrative burden, eliminating unnecessary documentation, timely and ease of access information for finding evidence-based solutions, easier communication, ways to plan activities more effectively.

The Bulgarian E-Health systems should:

- Improve the quality of information services for all sides involved in health care;
- Define health information standards, regulating the information interaction among all participants in the healthcare process;
- Provide the information necessary for monitoring the health of the population at different levels and the quality of medical services;
- Provide information for planning and control of financial flows in healthcare;
- Attain unification with health information systems of the European Union and WHO.

The paper is organized as follows: Chapter 2 presents the strategy of WHO for E-Health. Chapter 3 discusses the EU E-Health plans. The Bulgarian Ministry of Health E-Health activities are presented in chapter 4. Chapter 5 describes the current state and the future of the Bulgarian E-Health electronic medical systems. The paper ends with conclusions.

2 THE WORLD HEALTH ORGANIZATION STATEGY

The World Health Organization published in 2012 its National E-Health Strategy Toolkit (World Health Organization and International Telecommunication Union, 2012). It reflects the growing impact that E-Health is bringing to the delivery of health care around the world today, and how it is making health systems more efficient and more responsive to people's needs and expectations. It is divided into three parts: National E-Health vision, National E-Health action plan and National E-Health monitoring and evaluation. WHO's key undertaking within E-

Health is to provide independent advice and assistance to countries towards the long-term development of sustainable national E-Health solutions in relation to strengthening health systems performance and the capacity for countries to gather and analyse health information. WHO delivers E-Health in 3 ways: (1) As a knowledge-broker and facilitator between nations and the International Community at large; (2) By developing and sharing best practices and standards precipitated from successful E-Health implementations; (3) By working directly with Ministries of Health to address their technical and strategic needs for E-Health and Health Information.

3 EUROPEAN COMMISSION E-HEALTH PLANS

The European Commission published in 2012 its E-Health Action Plan 2012-2020 - Innovative healthcare for the 21st century (European Commission, 2012). The vision of the Action Plan is to utilize and develop E-Health to address several of the most pressing health and health systems challenges.

Health 2020 is the new European health policy framework. It aims to support action across government and society to: significantly improve the health and well-being of populations, reduce health inequalities, strengthen public health and ensure people-centred health systems that are universal, equitable, sustainable and of high quality. It has two strategic objectives, constructed around equity, gender and human rights and improved governance for health.

European Commission support can be divided into three main areas: co-financing of projects, support to awareness-raising events (in particular to the annual High-Level Presidency E-Health conferences) and running structures for awareness and networking building, best practice sharing and policy development.

The goals of the main EU E-Health projects are to:

- Test and demonstrate new models and tools for health and care delivery;
- Support the translation of findings into the clinic and other health and care settings to: improve health outcomes, reduce health inequalities, and promote active and healthy ageing.

Some of the EU projects are connected with the overview of the national laws on electronic health records in the EU Member States (European Commission, 2013, Overview of the national laws on electronic health records in the EU Member States National Report for Bulgaria).

Common E-Health challenges observed in Europe are:

- Ownership and governance of E-Health, developing and delivering on national E-Health strategies and managing the burden of ongoing system development and maintenance;
- Developing and adopting appropriate legislation to allow for E-Health;
- Ensuring security, privacy, identity management and ethical issues;
- Workforce issues: education, awareness and retraining;
- Acceptance of solutions by health professionals;
- Regionalization (within country) can be both a strength and a weakness;
- Digital literacy issues are contributing to a delay in E-Health adoption.

4 THE BULGARIAN MINISTRY OF HEALTH E-HEALTH ACTIVITIES

The Bulgarian Ministry of Health published in 2013 a National Health Strategy (2014-2020) (The Bulgarian Ministry of Health, 2013). It includes the development of E-Health as an essential tool for ensuring the effective functioning of the health system. A National Health Information System (NHIS) will be developed to provide public access through an electronic identifier managing health records, electronic prescription, etc. This will take place through the following steps:

- Development of mandatory standards for health information and statistics;
- Development of policies for security and interoperability of the healthcare information systems;
- Establishment of a national health information system and ensuring public access to the system through an electronic ID.

Some of the main system functionality will allow:

- The use of a secure user interface to exchange information in real time between physicians, patients, laboratories, insurance company;
- Management of the electronic health record of the patient;
- Providing information to the public and health education;
- Distribution of telemedicine services in diagnosis, treatment and surgery.

Development of E-Health is a government priority and this is clearly indicated in the Government Programme for sustainable development of the Republic of Bulgaria (Government Programme for sustainable development of the Republic of Bulgaria 2015).

The main EU recommendations for including E-Health in Bulgaria are:

• Need of a well-established health strategy, as a precondition for access to the structural funds (ex-ante conditionality criteria);

• Need of an action plan of this strategy before planning the investments needed;

• Adoption of EU standards for interoperability should guide the investments in E-Health;

• Continuous and strong involvement in EU initiatives, such as joint actions and the eHealth Network.

In E-Health, a reform facilitator final report on Health Financing Diagnostic and Review of Envisaged Reforms (Final report, 2015) is written that Bulgaria is rich in technology resources, but it seems that the country is now falling behind its peers. The Ministry of Health is well placed to play a significant role in encouraging cooperation among all health stakeholders, but before that it must resolve the fundamental issue of the National Health Insurance Fund system.

Expected results of applying the plan are: Integration and connectivity in healthcare through a national health information system and real access to patient information about their own health.

There is also a non-profit, non-governmental organization functioning in Bulgaria - The Bulgaria E-Health Foundation. It was established with the purpose of boosting the development of E-Health on national level as part of the electronic government of the Republic of Bulgaria. The necessity of speeding up the health reform in Bulgaria requires the development of electronic healthcare as a cornerstone in our health system. In this process the Foundation works together with all participants and interested parties in the healthcare process – the Ministry of Healthcare, the National Health Insurance Fund, private health insurance funds, hospitals, GPs, pharmacies, laboratories, medical doctors and patients (E-Health Bulgaria Foundation, 2015).

5 BULGARIAN E-HEALTH ELECTRONIC MEDICAL SYSTEMS

An essential tool for ensuring the effective functioning of the health system is the integration and connectivity of healthcare by building a national health information system and ensuring public access to the system through an electronic identifier. A unified information system of health care is the basis on the development of which E-Health will be built with its main components: electronic health records, electronic prescription, electronic referral, electronic portal, etc. Its implementation allows for more online administrative and health services sector activities, providing access to information about the patient‘s own health, improving relations between different levels of the system, improving the quality of medical services and making the spending of public funds for health care more effective.

The Bulgarian E-Health electronic medical system has to include information about:

• Patients – to improve treatment, thanks to evidence-based medicine and provide a way to involve patients in decisions about their health;

• Medical and health professionals - for quick and easy access to information, diagnosis and for performing complex remote interventions, as well as access to specialized education and training resources; with the support of medical research, effective management and dissemination of medical knowledge;

• Managing the health care system - to improve access and dissemination of best practices for planning and management of healthcare for the benefit of patients and society;

• Civil society - in order to better health education, a healthy lifestyle, prevention, information, health resources and opportunities for the health system to be able to influence the management of the health system at the local and national level.

The Bulgarian Ministry of Health needs systems in health policy planning including:

• “What-if” systems that model different policy parameters and attempt to compare results;

• Simulation systems that similarly provide guidance on policy decisions based on the simulation of random epidemiological or environmental or macroeconomic events;

• Geographic Information Systems that can create a full “health map” of the country;

• Systems aimed at monitoring and enforcing quality standards across the health sector;

• Budgeting systems, National Health Accounts analytic systems and accounting systems.

The structure of the current National Health Information System is given in Figure 1. Processing of the claims includes: electronic files submitted by hand or by e-mail. Different databases are used by regions and types of claims. There is one point missing, a uniform mechanism and database for processing, approval and storage of all claims. There is no control of funds and optimizing payments software. The system has two cores, an Enterprise Resource Planning (ERP) core (Core 1) for managing the financials, and Core 2, which provides the centralized claims-processing step at the central site in Sofia. Hospital claims, the largest per-unit cost claims must now bypass the claims adjudication system altogether and proceed to the payment module with seemingly little scrutiny. Some of the regional modules used in processing claims are given in Figure 2.

The current problems of using NHIS (P. Moskov, 2015) could be summarized as:

• Lack of a comprehensive information picture of the activities in the healthcare system;

• No mandatory application of health information standards;

• No statistically valid planning;

• No quantitative rational control of the quality and effectiveness of medical activities.

The following steps will be taken in the next five years:

• Creation of an Electronic Health Record for each Bulgarian citizen;

• Creation of a National Health Information System and its infrastructure (portal data center, electronic identification) available to all entities of the health system (natural and legal, public and private);

• Creation of electronic prescriptions;

• Construction of health and medical standards;

• Construction of pharmacy-therapeutic practices;

• Development of a unified medical ontological database;

• Creation of a Statistical Center;

• Further development of statistical standards;

• Creation of sets of metrics for quality management.

Work will be based on the following principles:

• Leading role of the state;

• Use of existing initiatives and systems;

• Providing a gradual process of construction in accordance with our long-term vision;

• Protection of existing investments;

• Security and access to regulated information in accordance with the regulations;

• Mandatory use of health information standards;

• Uniform requirements and equal access to information;

• Broad consensus and involvement of all stakeholders.

The steps in the next five years are given in Figure 3.

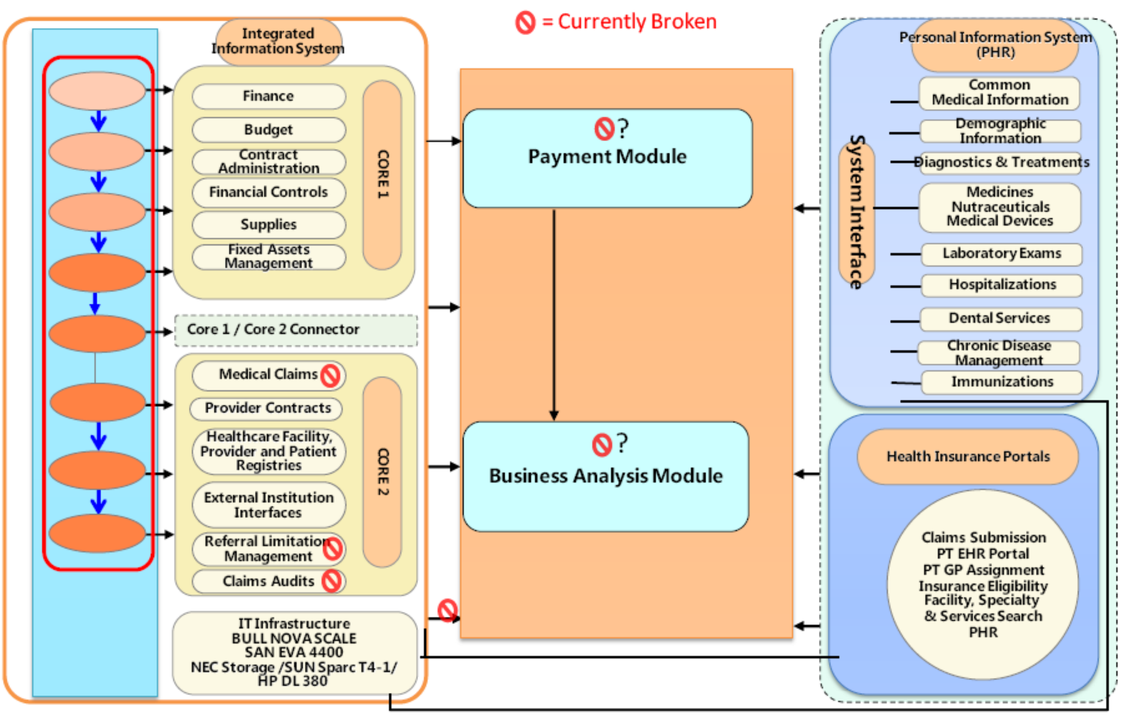

Figure 1: Main modules of the current National Health Information System

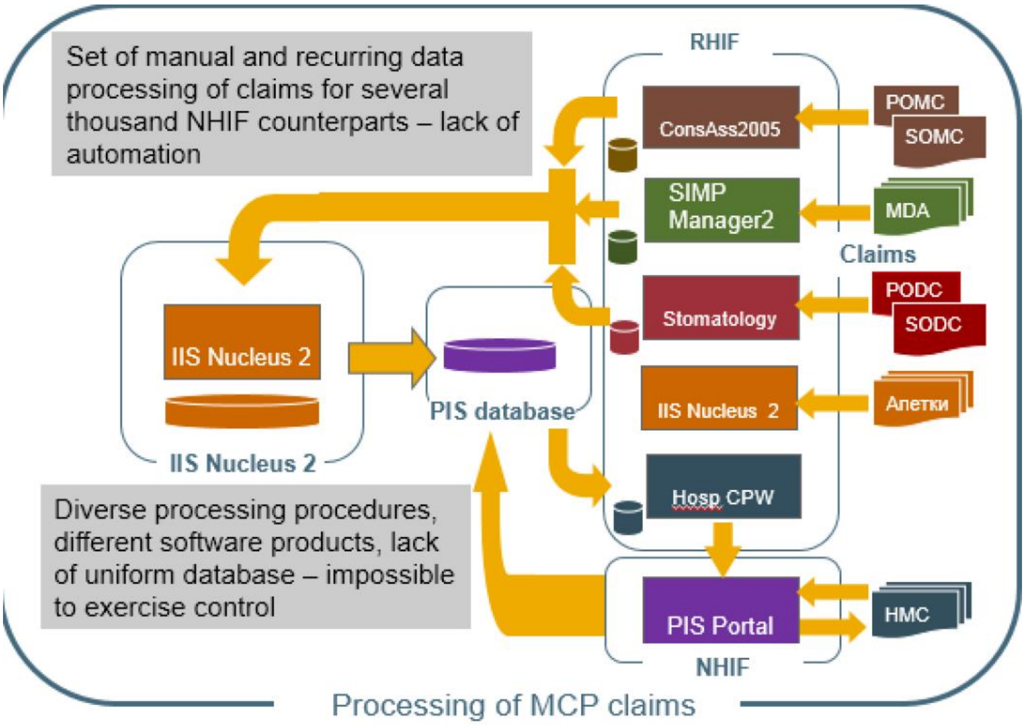

Figure 2: Regional modules of the current National Health Information System

2015	2016	2017	2018	2019	2020
Ministry of Health Database 1		Ministry of Health Database 2		National E-Health Portal	
Medical Nomenclatures		National Statistics Centre		Medical Ontologies	
Medical Standards		Standard Information System 1		Standard Information System 2	
Medical Good Practices		Medical Records 1		Medical Records 2	
Health Legislation		Medical Administrative Services 1		Medical Administrative Services 2	
Health System Analysis		Medical Software Certification		Telemedicine	
NHIS - core1 + E-medical Records + E-Prescription		NHIS - core 2		NHIS - core 3	
Technical Support 1		Technical Support 2		Technical Support 3	
Human Relation 1		Human Relation 2		Human Relation 3	
Project Management 1		Project Management 2		Project Management 3	

Figure 3: Steps to be taken in the next five years for building the new National Health Information System

6 CONCLUSIONS

WHO has both a normative and a supportive role to play in the member states in the field of E-Health. WHO and the European Commission are committed to bringing together representatives of all member states to advocate for the development of a national E-Health strategy, the adoption of standards for interoperability and promoting the implementation of E-Health with the aim of attaining Universal Health Coverage.

The Bulgarian Ministry of Health is working to bring E-Health to citizens. E-Health is essential to an efficient and sustainable health system. Bulgaria needs interoperable systems to attain E-Health solutions.

At present in Bulgaria there exist hospital information systems, software for providers of outpatient care, small databases for different medical practices, an NHIF system at different centres and agencies.

There is no single system at the moment that unifies and enables communication between different information systems and databases. There is an action plan and a road map for building a National E-Health system over the next five years.

REFERENCES

World Health Organization, 2015. Trade, foreign policy, diplomacy and health, E-Health, retrieved August 10, 2015, from: http://www.who.int/trade/glossary/story021/en/

World Health Organization and International Telecommunication Union 2012, National eHealth Strategy Toolkit Overview, retrieved August 10, 2015, from: http://www.who.int/ehealth/publications/overview.pdf

European Commission, 2012. Digital Agenda for Europe, eHealth Action Plan 2012-2020: Innovative healthcare for the 21st century, retrieved August 10, 2015, from: https://ec.europa.eu/digital-agenda/en/news/ehealth-action-plan-2012-2020-innovative-healthcare-21st-century

European Commission, 2013, Overview of the national laws on electronic health records in the EU Member States National Report for Bulgaria, retrieved August 10, 2015, from: http://ec.europa.eu/health/ehealth/projects/nationallaws_electronichealthrecords_en.htm

The Bulgarian Ministry of Health, 2013, National Health Strategy (2014-2020), retrieved August 10, 2015, from: http://www.mh.government.bg/media/filer_public/2015/04/08/nacionalna-zdravna-strategia_2014-2020.pdf

Government Programme for sustainable development of the Republic of Bulgaria 2015, retrieved August 10, 2015, from: http://www.government.bg/cgi-bin/e-cms/vis/vis.pl?s=001&p=0211&n=132&g=).

Final report, 2015, E-health—a reform facilitator Final Report on Health Financing Diagnostic and Review of Envisaged Reforms, May 2015, advisory services agreement between the Ministry of Health of the Republic of Bulgaria and the International Bank for Reconstruction and Development, retrieved August 10, 2015, from: http://www.mh.government.bg/media/filer_public/2015/06/16/final-report-on-health-financing-diagnostic-and-review-of-envisaged.pdf

E-Health Bulgaria Foundation, 2015, retrieved August 10, 2015, from: http://www.ehealth-bg.org/en/

P. Moskov, 2015, The reform of the healthcare system, retrieved August 10, 2015, from: http://www.mh.government.bg/bg/novini/aktualno/dnes-ministr-d-r-petr-moskov-iznese-prezentaciya-p/

Building a National Clinical Data Warehouse

Ustun Yildiz[1] and Ayse Keles[1,2]
[1]Agency of Health Informatics, Ankara, Turkey
[2]Yildirim Beyazit University, Ankara, Turkey
{ustun.yildiz, ayse.keles}@saglik.gov.tr

Keywords: EHR, Clinical Data Warehouse, Data Privacy, Health Transformation Program

Abstract: Turkey has created an e-health vision along with the Health Transformation Program. In this framework, standard coding systems such as standard definitions of the institutions, databank of healthcare providers, standard disease, drug and medical supplies classifications have been developed and a national clinical data warehouse (Health-NET) was established. Health-NET is an integrated, safe, fast and expandable health information system which aims to improve efficiency and quality of health services by collecting all kinds of data produced in the health institutions in line with the standards and generating information adequate for all stakeholders of the collected data.

1 INTRODUCTION

Access to healthcare is regarded as a fundamental element of social development. While classic theories of development explain development in terms of many socio-economic and cultural indicators such as GNP per capita, level of industrialization and employment, and consumption level of primary goods and services, today, it is explained primarily in terms of indicators related to the access to education and healthcare. At the end of 2002, the status of the Turkish health system made it necessary to undertake radical changes in many areas from service delivery to financing and from human labor to information systems. In order to improve the quality of service in these areas, Turkey has gone through a Health Transformation Program (HTP) started as early as 2003(OECD, 2014)(Akdag, 2008)(Akdag, 2011). As a result of the health infrastructure rehabilitation efforts, Turkey has observed a rapid decline in the Under-5 Mortality Rate (U5MR) from 29 in 2003 to 7.7 in 2011 per 1000 births (T.D., 2010) Similarly, the life expectancy at birth has increased to 75 in 2011 from 70 in 2004. Consequently, the satisfaction with the government provided health services was measured 75.9% in 2011 comparing to 39.5% in 2003(WHO, 2012). In the context of HTP, the effective collection and use of nation-wide Electronic Health Records (EHR)s became a primordial goal. In parallel with the development of the relevant legislative framework, a national clinical warehouse (CDW) that collects EHR and other operational data from all health organizations in the country had been established in 2012 and became operational. As analogous to the worldwide approach of putting the health information at the center of decision processes along with the patient, the data collected from the healthcare facilities are used to structure and manage the new health infrastructure of the country (De Mul, 2012)(Yoo, 2014). Generally speaking, CDW is used to empower traditional application software in order to analyze public health behavior and support several different decision workflows such as clinical quality improvement (Weiner, 2012), pay-for-performance(Van Herck, 2010) and evidence-based medicine systems (Sacklett, 2000). The integration and use of diverse healthcare data from various sources into the same clinical repository is a challenging problem when implemented at a nation-size level. The problem becomes even more complex when Hospital Information Systems (HIS) that create the collected data are autonomous and implemented with different technologies. The current HIS market in Turkey includes more than 150 private firms that implement their customized software in more than 1500 healthcare facilities operating with more than 200.000 beds capacity (Kose, 2013). In this paper, we share our experience in implementing CDW and its use in the critical decision processes of the Ministry of

Health of Turkey. We summarize the design and operation processes, the software architecture, its use in Health Transformation Program and the lessons learned.

2 BACKGROUND

In this section, we describe some of the main aspects of a large-scale CDW that needs conceptual and technical considerations.

2.1 Data Collection

The health institution that provides CDW with data is composed by Level I Family Practitioners, Level II Public and Private Hospitals and Level II University Hospitals and Research Centers. It is obvious that the data collection might not be achieved with 100\% accuracy due to the complexity of data and the widespread use of the system. Data collection services are frequently updated due to the updates in the data packages definitions. This continuous process requires a tight coordination with Hospital Information System (HIS) providers that will implement client component for data upload to collection services. Due the difficulties in the integration process that needed continuous support, a help desk has been established by the Ministry of Health. The help desk provided regular data on the amount of data collected by CDW, the amount and type of upload errors to HIS firms. It has been recognized that the success of integration capabilities of institutions that belong to different levels are different. Level I institutions were 99% successful in uploading their data as described in the integration kits of CDW public website. Level II institutions were less successful in sharing their data. The main reason behind the successful integration of Level I institutions was based on the underlying business model which makes possible the calculation of practitioners salaries based on the collected data by CDW. While Level I data collection was successful, Level II data collection had not been as expected and it was below 60% average for the first year. The main reason for the lack of data was that Level II institutions were not subject to any business model implemented by CDW. The second year of the implementation of CDW, the data collection rate was increased to 77% as the parameters for Service Quality Standards were started to be calculated with CDW data. The integration of Level III research institutions were even less successful given that they were not managed by the Ministry but by independent universities. During the third year, the integration of Level II has improved by the increasing calculation of different healthcare service parameters using available CDW data.

2.2 Data Quality

It is important that we should not confuse the concept of 'data quality' to the aforementioned concept of 'clinical quality data'. While 'clinical quality data' is specific data based on clinical quality indicators which help to understand the clinical quality of the services provided to patients with specific diseases e.g. diabetes, stroke etc. On the other hand, 'data quality' is about the quality of any data collected by CDW and based on certain criteria sets e.g. complete, valid/correct, timely, without duplication (Kahn, 2012)(Arts, 2002). In this section, we detail the issues and our solutions to improve the ‘data quality’ of the ‘clinical quality indicators’ collected for CDW. The major issue that has been encountered in establishing the data quality was the data wrongly packed that do not fulfill data package acceptance rules. We identified the main reason was the difficulties associated with the establishment of data packages to be sent to CDW. On the other hand, we identified HIS users use different ICD-10 (WHO, 2004) codes for certain diagnosis and diseases because HIS require sophisticated data input interfaces for the latters. One particular aspects that needs special attentions is the geographic and temporal properties of the clinical data. It is observed that public health indicators could be misleading based on certain periods and on locations (e.g. Temporary Refuge Spaces), these problems are configured by expert knowledge.

2.3 Data Privacy

The privacy of EHR had been a high priority concern in the implementation of CDW and the tools that manage its data. In the collection process of EHR, the definition of data packages was mainly defined by the public health surveillance necessities and the establishment of personal health records. The idea behind the establishment of personal health records is to support the continuity of healthcare and prevent redundant services such as radiology. One other advantage of the involvement of patients in the structuring the records is the elimination of inconsistencies. Turkish citizens have a unique and publicly available 10 digit number. The use of this number considerably facilitates the consolidation and access of personal health records but in the same time could be a major privacy concern.

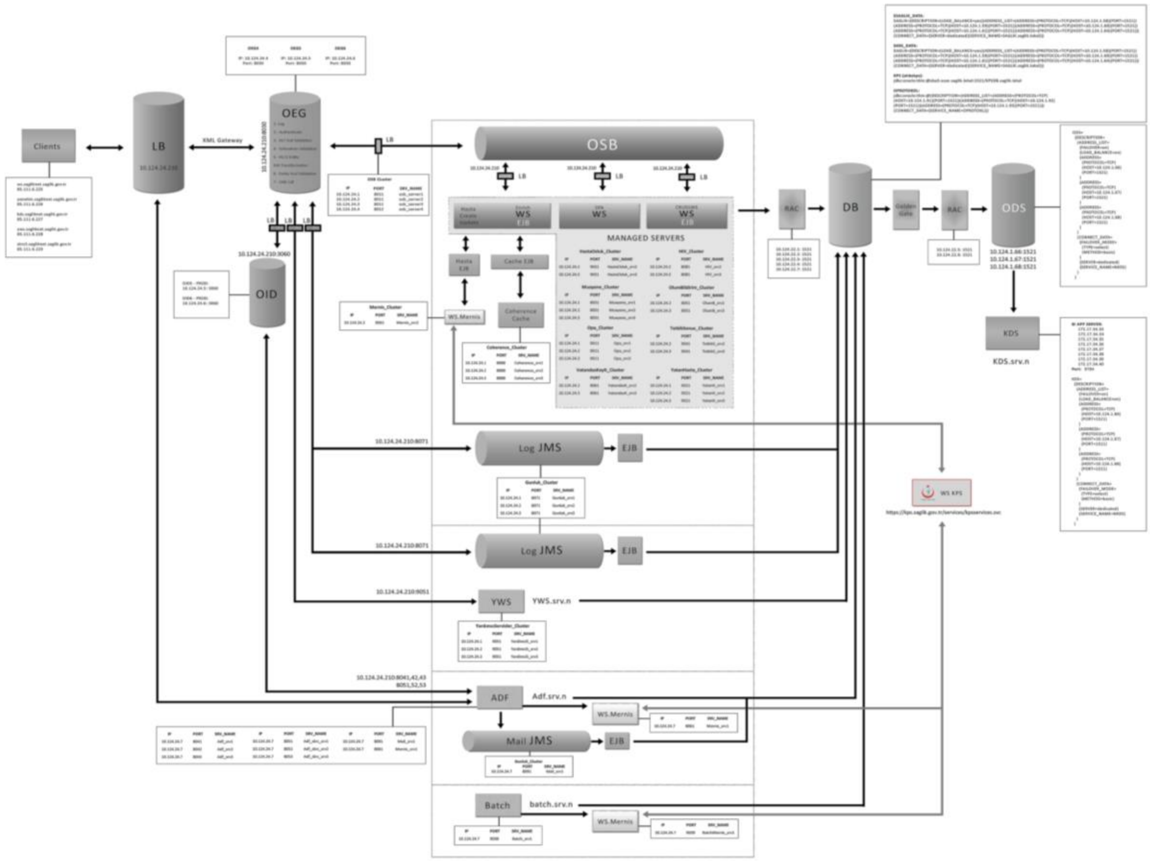

Figure 1: Overview of Architecture

In order to overcome with privacy concerns, an electronic consent form has been implemented in personal health record system (MoH, 2015). Patients can login and configure their preferences on the use of their records. The preferences are overwritten only in the case of emergency room services. It is recognized that the collection private health information and its computation after de-identification of data still can be a privacy concern as the identification of patient is possible with the combination of different queries (Fernandez, 2013) For certain diagnosis and diseases codes (e.g. HIV), we have decided to collect the records without identifiable information and a cryptic number without publicly available citizenship number

3 SOFTWARE ARCHITECTURE

Figure 1 describes the overview of the architecture of CDW with components related to data collection through web services, storage and data analyze modules. Messages coming from healthcare facilities are collected under HL7 form by Load Balancer component. Received messages are forwarded to an available XML Gateway. In parallel, HL7 messages are processed asynchronously by the JMS log queue. During its flow throughout the XML Gate way, the Authentication happens with LDAP on OID. Next, HL7 message is validated against XSD schema and Schematron processes business rules (mandatory and optional fields in data packages). After the validation, HL7 messages are transformed to a local data format to be saved in the database. The transformed messages are also validated against XSD Entity.

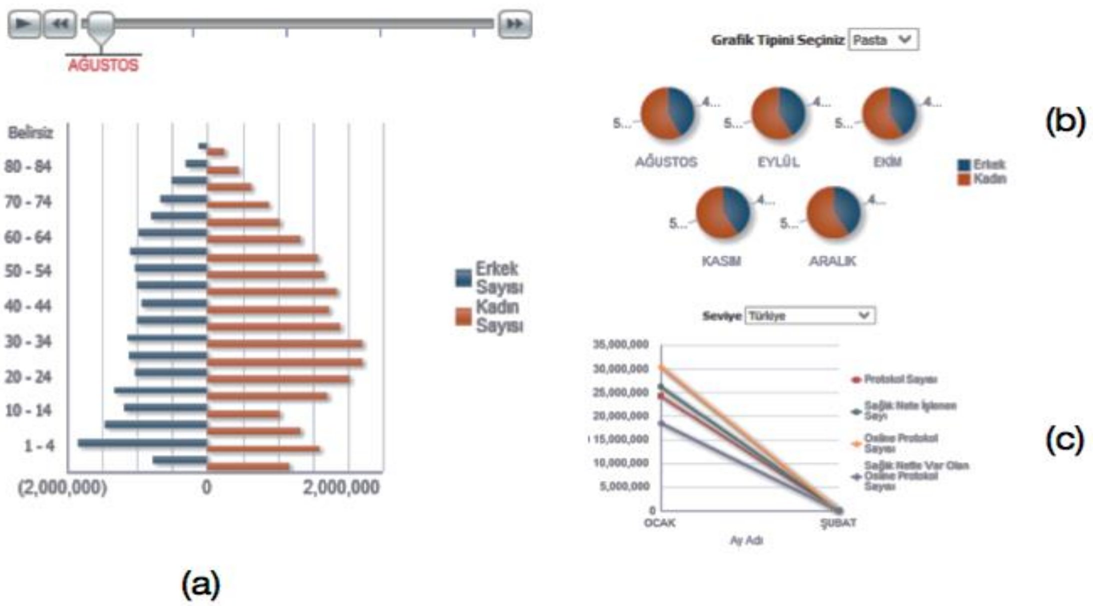

Figure 2: Functions available on dashboard

Collected messages are transferred to Service Bus. In the latter, the data is unpacked and the patient ID is verified through an external ID verification service (WS-Mernis). Figure 2 describes some of the functions available to users.

4 CONCLUSIONS

This paper has presented a small set of lessons learned from the establishment of a National Data Warehouse. The main research problems that we took the challenge on were: Data quality, data collection and data privacy issues. Although the patient count and basic computational problems were manageable with relation data models and systems, the advanced computation of patient records such as disease correlation analysis, organization of cohorts for evidence-based medicine applications require the use of big data solutions.

REFERENCES

OECD, Oecd reviews of health care quality: Turkey 2014doi:http://dx.doi.org/10.1787/9789264202054-en. URL /content/book/9789264202054-en

R. Akdag, Turkey health transformation program 2003-2010, Tech. Rep. 839, Re- public of Turkey, Ministry of Health (June 2011).

R. Akdag (Ed), Health transformation program in turkey and primary health care services, Tech. Rep. 770, Republic of Turkey, Ministry of Health (November 2008).

T. D. V. D. Lilia Jelamschi, Decline in the under-5 mortality rate (u5mr) in turkey: A case study, Tech. rep., UNICEF, http://www.unicef.org.tr/files/bilgimerkezi/doc/declin-in-the-under-5-mortality-rate-u5mr-in-turkye-a.pdf (March 2010).

W. H. O. R. O. for Europe, Successful health system reforms: The case of turkey, Tech. Rep. WHO LIS e96508, World Health Organization: Regional Office for Europe (May 2012).

M. De Mul, P. Alons, P. Van der Velde, I. Konings, J. Bakker, J. Hazelzet, Development of a clinical data warehouse from an intensive care clinical information system, Computer methods and programs in biomedicine 105 (1) (2012) 22–30.

S. Yoo, S. Kim, K.-H. Lee, C. W. Jeong, S. W. Youn, K. U. Park, S. Y. Moon, H. Hwang, Electronically implemented clinical indicators based on a data ware-house in a tertiary hospital: its clinical benefit and effectiveness, International journal of medical informatics 83 (7) (2014) 507–516.

J. P. Weiner, J. B. Fowles, K. S. Chan, New paradigms for measuring clinical performance using electronic health records, International Journal for Quality in Health Care 24 (3) (2012) 200–205.

P. Van Herck, D. De Smedt, L. Annemans, R. Remmen, M. B. Rosenthal, W. Sermeus, Systematic review: effects, design choices, and context of pay-for- performance in health care, BMC health services research 10 (1) (2010) 247.

D. L. Sackett, Evidence-based medicine, Wiley Online Library, 2000.

R. Kose, Health statistics year book, Yearbook 978-975-590 509-9, Agency of Health Research (2013).

M. G. Kahn, M. A. Raebel, J. M. Glanz, K. Riedlinger, J. F. Steiner, A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research, Medical care 50.

D. G. Arts, N. F. De Keizer, G.-J. Scheffer, Defining and improving data quality in medical registries: a literature review, case study, and generic framework, Journal of the American Medical Informatics Association 9 (6) (2002) 600–611.

W. H. Organization, et al., International statistical classification of diseases and health related problems (the) icd-10, Ph.D. thesis, World Health Organization (2004).

T. M. of Health, E-nabiz personal health records system, , https://enabiz.gov.tr/Giris.aspx (2015).

J. L. Fernández-Alemán, I. C. Señor, P. Á. O. Lozoya, A. Toval, Security and privacy in electronic health records: A systematic literature review, Journal of biomedical informatics 46 (3) (2013) 541–562.

Using Togaf for Building a National Implementation Strategy for E-Health Services and Technologies in Burundi

Frank Verbeke[1], Marc Nyssen[1], Sandrine Kaze[2] and Etienne Mugisho[3]
[1]Department of Biostatistics and Medical Informatics, Vrije Universiteit Brussel, Brussels, Belgium
[2]Burundi Health Informatics Association, Bujumbura, Burundi
[3]Belgian Technical Cooperation, Bujumbura, Burundi
frank.verbeke@vub.ac.be , marc.nyssen@vub.ac.be , Sandrine.kaze@openit-burundi.net, etienne.mugisho@btcctb.org

Keywords: e-Health enterprise architecture, TOGAF, Health information systems, Burundi

Abstract: In order to better align existing and future ICT implementations in the health domain with the strategic options defined by the National Plan for Health Development, the Ministry of Health (MoH) of Burundi initiated in 2014 the development of a national e-health enterprise architecture based on the TOGAF methodology. A first part of the development cycle consisted of a detailed analysis of regulatory documents and strategic plans related to the Burundian health system. In a second part, semi-structured interviews were organized with a representative sample of relevant MoH health structures. The study demonstrated the donor driven unequal distribution of hardware equipment over health administration components and health facilities. Internet connectivity remains problematic and few health oriented business applications found their way to the Burundian health system. Paper based instruments remain predominant in Burundi's health administration. The study also identified a series of problems introduced by the uncoordinated development of health ICT in Burundi such as the lack of standardization, data security risks, varying data quality, inadequate ICT infrastructures, an unregulated e-health sector and insufficient human capacity. The results confirm the challenging situation of the Burundian health information system but they also expose a number of bright spots that provide hope for the future: a political will to reclaim MoH leadership in the health information management domain, the readiness to develop e-health education and training programs and the opportunity to capitalize the experiences with DHIS2 deployment, results based financing monitoring and hospital information management systems implementation.

1 INTRODUCTION

In 2005, the Ministry of Public Health and Fight against Aids (MoH) of Burundi has developed a National Health Policy covering the period 2005 to 2015. This policy was later translated by the MoH and its technical and financial partners into a series of objectives and results in the National Plan for Health Development 2011-2015. Amongst the objectives were the reinforcement of the National Health Information System and the restoration of the MoH leadership in the field of health information management. Therefore, a number of priority actions have been identified:

- The development of an e-health strategic plan for strengthening the national health information system
- The development of an integrated and competitive health information management system
- The development of effective tools for planning, monitoring and evaluation
- Increasing the availability of ICT tools (hardware, networks and software) at all levels of the Burundian health system
- The promotion of data driven research activities in the health sector

Integrating e-health in the national health policy yielded from the beginning enthusiasm from the donor community and in the course of the past decade, a growing number of ICT tools have found their way to the Burundian health sector. But most of these tools have been introduced for supporting projects lead by NGOs and foreign technical and financial partners whereas hardware and software solutions almost systematically served well the donor objectives, but inter-project coordination and interfacing remained exceptional. Several successful e-health tools remained hidden in silo-projects and only produced a fraction of their potential benefits. Without corrective action, the Burundian health sector threatens to evolve towards a cacophony of divergent non-integrated health informatics implementations. In order to cope with this threat, the MoH initiated in 2014, with financial backup of the Belgian Technical Cooperation, the development of a national e-health enterprise architecture based on The Open Group Architecture Framework (TOGAF). In a first step, an initial analysis of human resources, business processes, hardware, software, communication and networking infrastructure related to health information management, had to be established. This study describes the objectives, methods and findings of this analysis.

2 MATERIALS AND METHODS

The main objective of the study was to provide a reliable estimation of the existing human and material resources and issues related to health information management in Burundi. The research hypothesis was that an industrial framework like TOGAF could be used for this, even in the challenging environment of one of the poorest countries in the world. If successful, the study results were to become the first step in a complete e-health enterprise architecture development cycle according to the TOGAF methodology, and therefore needed to provide data for the development of 4 essential sub-architectures:

- **Business architecture**: what are the MoH business needs in terms of health information management?
- **Application architecture**: which health information management applications have already been implemented in the field and to what extent do they address the business needs?
- **Data architecture**: what data is needed and collected today by the MoH and what is the quality of it?
- **Technology architecture**: what are the necessary technologies (software, hardware, and networking) and which ones are used today in the health domain in Burundi?

A first part of the study consisted of a detailed analysis of regulatory documents and strategic plans related to the Burundian health system.

In a second part, field visits and semi-structured interviews were organized with a sample of relevant structures of the MoH. A standardized study-specific interview guide was developed and systematically used by the interviewers.

3 RESULTS

3.1 Mission analysis and field visits

The study of regulatory documents and strategic plans took place in October and November 2014. After that, a series of field visits and interviews have been organized with 39 relevant MoH and -related structures in the Bujumbura province (the permanent secretary and all MoH directorates, major health programs, donor agencies, NGOs, public and private health facilities and educational institutions). In the period from November to December 2014, the e-health architecture development team also visited 5 other provinces (Muramvya, Gitega, Ruyigi, Kirundo and Ngozi), covering 5 provincial health offices, 5 health district administrations and 12 hospitals. In total, management staff of more than 15% of the MoH structures have been questioned about the mission, the mandate and the vision of their organization, their objectives and the way their work is organized. After that, a detailed analysis was made of health information management related human resources, ICT solutions and non-ICT (paper based) instruments at their disposal and procedures used for exchanging health information with other (MoH or non-MoH) organizations. Finally, an analysis was performed of health information management problems, expected benefits and potential threats of health ICT for each component of the MoH.

3.2 Hardware

The study showed that computer hardware has most often been supplied to the MoH by donor-driven intervention programs. There is no organization-wide management of computer equipment and distribution of hardware over the different MoH directorates, provincial- or district administrations and hospitals is very heterogeneous: some structures which are supported by several donors are over-equipped, others remain without any computer hardware at all. Under impetus of recent national and provincial policies and international hype, a growing number of health centers in Burundi started buying computer hardware with their own funds, unfortunately without having a clear idea of how to integrate such tools in their existing activities.

Generally speaking, hardware specifications are low standard: desktop PCs with Windows XP and Windows 7 operating systems, of which a large number have limited functionality due to computer virus infections (no budget is available for keeping antivirus software databases up to date and many of the PCs have no access to internet for performing these updates anyway). PCs are almost always accompanied by an uninterruptible power supply (UPS) but due to the lack of battery maintenance, the protection offered by these UPSs is minimal.

Many of the executive health staff make use of laptop computers which in about half of the cases are their personal privately owned equipment.

Printers are rarely shared in a network and toner or ink cartridge supply is problematic due to the unavailability of toner cartridges on the Burundian market or the lack of budget for operational costs.

Electronic files and documents are commonly transferred between computers using USB memory sticks, which constitute an infamous source of virus infections.

3.3 Networks

Most of the central MoH structures in the Bujumbura region have access to a wired or Wi-Fi based LAN. Many times, these networks are only connected to the internet by grace of donor funding, which is always limited in time (and sometimes also in data volume). Few larger structures (central MoH site, reference hospitals) have been connected to a national optical fiber network offering reasonable internet connectivity. However, for most of the small and medium-sized health facilities, broadband internet prices remain prohibitive and bandwidth offered by local ISPs in Bujumbura is poor and unstable although considerable improvement has been seen in the past few years.

Installation of internet connections is hardly coordinated, with some structures sometimes accumulating several (poorly performing) parallel connections on the same site: 4 different wired internet connections have been identified at the site of the national blood transfusion site, in addition to the numerous individual 3G-USB modems already offered by several donor programs. In spite of the inadequate internet bandwidth, most central level MoH structures still state that an internet connection has become indispensable for their daily activities.

Away from the national and provincial capitals, the situation is worse: wired internet connections are unavailable and performance of 2G and 3G wireless data networks is unpredictable. Some donor agencies (such as EU) have equipped MoH structures with VSAT connections which provide stable and reliable bandwidth but come with high operational costs. Many of these satellite internet connexions remain unavailable part of the time due to inappropriate use (downloading movies or audio) consuming all of the monthly foreseen VSAT credit in only a few days.

3.4 Software

Almost all of the end user computers run Microsoft Windows operating systems completed with Microsoft Office applications, with the exception of a number of desktop and server computers running Linux Mint or Ubuntu at the directorate of the national health information system.

Although health specific software implementations remain rare, a clear tendency towards web-based business applications is being noted, often based on Linux/Apache, MySQL databases and PHP or Java development:

- The MOH started in 2014 pilot implementations of the DHIS2 data warehouse as a replacement for the outdated MS Access based GESIS health data collection solution.
- iHRIS human resource information system deployment also started end 2014 with the first implementation pilots scheduled early 2015.
- Hospital information system (HIS) implementations remain rare (less than 10% of the hospitals), with all of the health facilities in our study sample running OpenClinic GA. The majority of the HIS solutions are concentrated in third level reference health facilities.
- OpenRBF has been implemented for monitoring of results based financing (RBF) programs at the central and provincial levels.
- Joomla and Drupal are the most popular solutions for website content development.

Some successful m-Health applications (the RapidSMS based KIRA Mama project and SIDA-info) provide promising results today.

Epi-Info and SPSS are the leading statistics software solutions. General and analytical accounting systems are used by several health sector structures of Burundi: Asyst and QuickSoft (local development), SAGE Saari, Popsy, and Banana were found in about half of the interviewed health facilities while Tompro was recently introduced for project-oriented accounting at the central MoH level.

3.5 Paper based instruments

The vast majority of the provincial and health district administrations are using ICT-tools for reporting health data to the central level (GESIS), but a number of hospitals and almost all health centers still rely on paper based instruments for routine data collection. Information is written down in registers by administrative clerks and clinical staff and sent on a monthly basis to the health district administration (emergency surveillance information is sometimes reported more quickly using SMS). Health districts then forward compiled health facility data to the provincial level, where eventually provincial reports are sent to the central level in Bujumbura.

A minimum of 25 paper registers must be kept by all health centers and around 75 registers are in use in an average district hospital. Additionally, donors and health intervention programs claim parallel and redundant reporting from the health facilities and district administrations they support, all of which represents an impressive administrative overhead.

Paper based instruments are also predominant for health record keeping in most (90%) of the hospitals. All of them are facing health information quality management issues.

3.6 Health information management problems detected

Over the past 10 years, the existing health sector ICT landscape of Burundi grew organically, with most of the project-oriented solutions being provided by donors and health programs. This happened in an uncoordinated way, leading to:

- Lack of standardization: health information representation is hardly standardized and few international classifications or coding systems are in use (with the exception of some of the DHIS2 and HIS modules using ICD-10).
- Data availability risks: many databases are hosted in donor countries outside Burundi, with true data accessibility risks for the MoH. Also, many MoH agents use personal computer equipment without appropriate backup procedures or anti-virus protection.
- Data protection risks: data access rights are not being organized in layers according to the role people fulfil in the health administration; most often one has full access to all of the information or no access at all.
- Varying data quality: multiple issues explain the poor quality of data collected in the field. There is (1) the lack of intrinsic motivation of MoH staff that don’t produce data for their own purpose; (2) the important administrative burden caused by redundant health data collection processes; (3) many MoH agents don’t have the necessary qualifications for producing reliable data; (4) the absence of personal consequences linked to the production of erroneous information; (5) donor funding for the collection of project specific health data at the same time compromising the global and systemic collection of routine data for which no financial incentives exist (RBF).
- Varying data promptness: the lack of reliable (electronic) communication instruments delays the transmission of health information between different levels of the health system.
- Lack of data completeness: data is sometimes considered a factor of power and the lack of perceived personal interest in information sharing may hinder the effective, complete and systematic exchange of health sector data.
- Defective and insufficient computer equipment: a number of MoH structures have no access to appropriate ICT hardware and due to the lack of maintenance procedures, many of the existing equipment has become defective. Computer virus infections also constitute a major problem for the MoH administration.
- Inadequate ICT infrastructure: today, access to stable electric power is out of reach for many of the MoH structures, even in the larger cities. UPSs have been provided with most of the computers, but their defective batteries often don’t provide any protection against power failures (sometimes power failures can last for several days, which heavily compromises the reliability of electronics in every day’s work). Affordable broadband internet is unavailable for most of the MoH components. Donor project-funded internet connectivity is always limited in time and does rarely bring a sustainable solution.
- Unregulated e-health market: although e-health solutions are being considered “medical devices” by WHO, no standards or regulations

have been put in place for introducing ICT-tools in Burundi's health system. E-Health solutions deployment therefore escapes today from any health authority control.

- Lack of health applications: most of the software solutions deployed in the health sector are generic office applications, statistical analysis applications or aggregate data reporting instruments. Few health application implementations such as hospital-, laboratory- or pharmacy information systems have found their way to Burundi's health system.
- Insufficient human capacity: on the one hand, qualified staff who are capable of effectively using ICT-tools in their work environment are missing in most of the MoH structures. On the other hand, there is a plethora of unmotivated and underqualified staff occupying positions in the MoH administration preventing young and better qualified workers from being recruited. Additionally, health-ICT related training and education opportunities are not aligned to the needs expressed by the different directorates and health facilities.
- Organizational problems: the organizational structure of the MoH reflects in no way the important transversal role of ICT in today's healthcare. The statute of ICT professionals at the MoH is far from attractive; they are considered an administrative burden rather than a valuable asset of the organization.
- Ineffective dissemination of information: the absence of a reliable communication network limits the dissemination of regulations, practice guidelines and policies from the central MoH level to the peripheral structures.

4 CONCLUSIONS

The TOGAF methodology, after applying some simplifications, offered the appropriate instruments to quantitatively and qualitatively describe the status of health ICT tools deployment in the health sector of a low-resource country like Burundi. The output of the study was later used as a starting point for the further development of an e-Health Enterprise Architecture for Burundi's MoH, which has been officially validated on July 29th, 2015.

The study results more or less confirmed the challenging situation of the Burundian health information system, but they also exposed a number of bright spots that provide hope for the future:

- There is a political will to reclaim MoH leadership in the health information management domain by enforcing compliance with international consensus and standards for future e-health initiatives, putting the MoH in a regulator/gatekeeper position.
- The human resource deficit in health informatics is huge and many of the country's education institutions will have to collaborate on national and international levels to provide necessary ICT training programs. Burundian academic institutions and the donor community seem to be willing to invest in this.
- DHIS2 implementation got substantial support from the government and donor agencies. Extensive training programs have started in December 2014 and a lot of enthusiasm exists to make the implementation of a flexible national health data warehouse a reality.
- Hospital information management systems implementation has been convincingly successful in several hospitals and provides clear evidence for the feasibility of HIS implementation in Burundi.

An important challenge remains to capitalize the experiences from the few success stories and to integrate these in a new coordinated, well adapted and appropriately funded e-health strategy for the country in the next 5 to 10 years. According to the architecture vision developed in this study, such a strategy should account for:

- The creation of a national MoH datacenter in Bujumbura that centralizes shared databases and applications and provides a professional infrastructure with stable electricity, access control, data backup and redundancy.
- The development of a multi-technology (optical fiber, 3G and VSAT) VPN-based health care intranet connecting central, provincial and district level structures.
- The implementation of shared generic applications for the public health sector: accounting software, workflow management, a unique central website, a virtual library, a geographic information system and an MoH owned mail server (preventing loss of valuable information when staff using gmail.com of yahoo.fr accounts leave the organization)
- The implementation/strengthening of a series of health specific business applications such as DHIS2, iHRIS, OpenRBF, OpenClinic GA HIS,

LMIS and a series of health resource registries (a facility registry being one of them)

- The implementation of tablet and smartphone based patient oriented health data collection tools in health centers and at the community level (KIRA Mama and SIDA-Info)
- The implementation of an SMS-to-IP gateway enabling health facilities that have only access to plain GSM and SMS connectivity to participate in the country's electronic data collection mechanisms.
- The development of 3 health informatics teaching programs to cope with the important human capacity building needs: (1) a Master in Health Informatics program in collaboration with universities from neighbouring countries, (2) a specialization program in applied health informatics for health professionals and (3) the creation of a biomedical technician bachelor program.
- The creation of an autonomous health informatics directorate at the MoH with departments in charge of (1) standardization and regulation, (2) health informatics infrastructure management (datacenter and intranet), (3) health informatics education and promotion and (4) helpdesk and support functions.

REFERENCES

Plan National de Développement Sanitaire II 2011-2015, Ministry of Public Health and Fight against Aids, Bujumbura, Burundi

The Open Group Architecture Framework (TOGAF), http://www.opengroup.org/togaf/

Politique Nationale de la Santé 2005-2015, September 2004, Ministry of Public Health and Fight against Aids, Bujumbura, Burundi

Politique Nationale de Développement des Technologies de l'Information et de la Communication du Burundi 2010-2025, November 2010, Council of Ministers, Bujumbura, Burundi

Plan Stratégique du Système National d'Information Sanitaire 2011-2015, Ministry of Public Health and Fight against Aids, Bujumbura, Burundi

Plan Stratégique National de Développement des Ressources Humaines pour la Santé 2010-2014, Ministry of Public Health and Fight against Aids, Bujumbura, Burundi

Human resources for health: overcoming the crisis. WHO; 2004. Available from: http://www.who.int/hrh documents/JLi_hrh_report.pdf

Kruse S-E. *Health Information Systems in Africa: Resources, Indicators, Data Management, Dissemination and Use*. Technical papers. Algiers: Centre for Health and Social Development (HeSo); 2008. Available from: http://www.tropika.net/specials/algiers2008/technical-reviews/paper-8-en.pdf

Kamadjeu RM, Tapang EM, Moluh RN. *Designing and implementing an electronic health record system in primary care practice in sub-Saharan Africa: a case study from Cameroon*. Informatics in primary care. 2005;2005(13):179–86.

Verbeke F, Karara G, Nyssen M, *Evaluating the Impact of ICT-tools on Health Care Delivery in Sub-Saharan Hospitals*, Stud Health Technol Inform, August 22, 2013

KIRA Mama project, Unicef, 2014, https://unicefstories.files.wordpress.com/2014/05/projet-kira-mama-pdf-final-1.pdf

SIDA Info, SWAA Burundi, http://www.sida-info-service.org

OpenClinic GA, http://sourceforge.net/projects/open-clinic

Suggestions for the Elicitation of Seniors Involvement in ICT and Socially Innovative Solutions to Tackle IT

Adrianna Nizinska[1], Antoni Zwiefka[2] and Karol Kozak[3,4]
[1]*Dean Educational Affairs, University of Lower Silesia, Wroclaw, Poland*
[2]*Departament Polityki Zdrowotne, Urząd Marszałkowski Województwa Dolnośląskiego, Wroclaw, Poland*
[3]*Medical Faculty, Dresden University of Technical, Fetscherstraße 74, D-01307 Dresden, Germany*
[4]*Data management Center, Fraunhofer IWS, Winterbergstraße 28, 01277 Dresden, Germany*
adrianna.nizinska@dsw.edu.pl, antoni.zwiefka@dolnyslask.pl, karol.kozak@uniklinikum-dresden.de

Keywords: Seniors, ICT, Human Computer Interaction, Social

Abstract: Human-computer interaction (HCI) issues for older people are extremely important in the light of the rapidly ageing population in developed countries. In addition, technology offers great potential for this age group, but it will be only useful if it can be used effectively by its target users. We will therefore examine how HCI can address the needs and situations of this increasing older population and how their involvement can be maximized in order to support participation and empowerment. In recent years, HCI technologies have manifested their potential to enhance the autonomy and quality of the life of elderly people, through boosting the elicitation of seniors. This paper is describing a methodology and challenges that will support healthcare professionals' action in the further effective usage of systems based on HCI.

1 INTRODUCTION

Innovations in human-computer interaction research have revealed effective methods for people with various disabilities to use computers or to receive computer-aided medical treatment (Hendrick, Schooley & Gao 2013).

Since the inception of medical computing three decades ago there has been an extensive discussion of the value of developing an interactive computer-based clinical record system for the practitioner, not only to provide routine decision support for patient care, but for the capture of both contemporaneous and longitudinal data important to clinical epidemiology, quality assurance, risk management, and the development of increasing varieties of experiential based reasoning. The goal of regular clinical use requires to put the user in focus or – as HCI (human-computer interaction (HCI) researchers would call it – to adopt a user-centered design approach (Norman, Draper 1986). This is not what already has been implemented and is our recommendation.

Human-computer interaction testing is being commonly used in many commercial settings to form better human computer interface scenarios (Gosbee, Ritchie 2007). The users of HCI are both patients and medical professionals. This means the HCI testing needs to include both types of users. Patients come with all levels of background and experience in technology. Medical professionals are naturally very patient-focused and not necessarily focused on medical devices. We suggest that HCI testing needs to meet the needs of all potential users.

Nevertheless, there has been a growing attention to the development of more powerful social methodologies, identifying barriers and focusing on challenges with new interaction modalities. It is also crucial for these methodologies to address the issue of participation and empowerment. Since WHO Global Strategy for Health for All by the year 2000 (1981) these categories have been of high importance in improving health and still remaining a challenging one. This paper is describing a methodology and challenges that will support healthcare professionals' action in further effective usage of systems based on HCI.

2 PROBLEM DEFINITION

As people get older they will also want to remain active in ways previous generations did not. They will want to continue to be a part of the workplace, to drive and travel the world. Seniors in 15-years´time will expect and demand from suppliers to be able to use new technologies as they appear on the market. Seniors will want to use computer technology to stay connected to society, work colleagues, friends and children. They will want them to support their health and well-being. And as they get older, friends and family may want computers to keep an eye on their aging family members. There will be an increasing number of social communities, where people will be able to upload their personal health data or send photos of what they have eaten and a log of their activities to online doctors who will be able to give them up-to-date and personalized assessments. Medical records of physical and psychological health may also become resources for new ways of sharing and documenting the medical problems of older life, allowing for more customized and reassuring health care.

The last 15 years and the era of smartphones and IP-based devices have seen not only an enormous growth in the number of devices, but also an almost explosive diversification in the nature of these devices, operating systems and terminals as they have entered every aspect of our lives. In this paper, we face senior population where they will need to live with an ever growing and dynamic set of interconnected digital devices. End- user interfaces will be close to seniors and even attached to them, while others will be invisibly built into their private ecosystem (home, senior house).

Design recommendations for senior-targeted technologies or rendering more elderly-friendly common technologies have been common in the literature on aging and technology (Discipulus Project, May 2013). Changes particularly in the cognitive, sensory and psychomotor functioning of adults as they grow older mean that the elderly often require interfaces that correspond their specific needs. What are the key technological setup guidelines for designing a computer system for senior users?

The reason can be discussed from two different points of view. The first is their age-related problems; the physical and mental decline make it hard to adapt to new products. Another perspective is a software logical and interface design; software architects do not take into account older users' characteristics and personalization. For a more specific group of users such as seniors, it is necessary to find out wider and deeper design parameters, which should be a more specific and enchanted version of technological guidelines.

3 SUGGESTIONS

3.1 Seniors as users and co – designers of ICT based e-health solutions

Most often the innovative business processes in health care results in shortening the healing process through the support areas. This support is mainly based on effective self-management and integrated IT systems. Such innovation systems can also include the implementation of telemedicine Innovations (Fig 1). These methods rely on the analysis of business processes and organizational structure in terms of increasing the efficiency of the care organization. There are many processes for improving the patient care with the possibility to increase efficiency in health care. In this group integrated care pathways are a typical patient-centric approach to the care of frail people realized by CareWell Project . It is an innovative approach and consists of a platform supporting information and communication needs of the patient. Every service to be active should also be gathered from the relevant experts. It supports what needs to be in place for legal, technical, organizational and financial requirements as well as the different needs of patients and healthcare professionals whilst taking into account elements that may bias them, such as their current access attitude and knowledge of ICT. It is also important how they perceive the impact the service will have on them and what they expect to be the service's advantages and disadvantages.

A new vision of elderly care is based on the patient's focusing and on high-quality equitable health care for all patients in the Europe. Providing a strong and integrated service for senior patients in the area of health, social inclusion and life and personal fulfillment is possible through the use of e-health solutions.

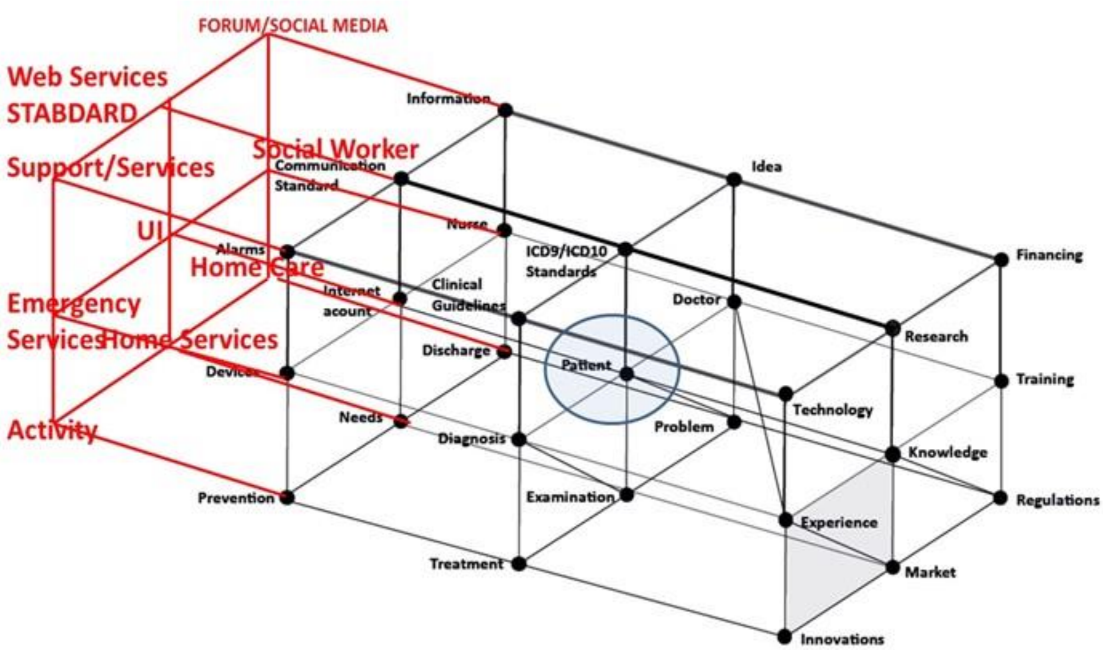

Fig 1. Graphical representation of health care system

According to the EU health policy priorities in the field of e-health it is necessary to put the patient into the center of attention and activity. The main risk factors and efforts in empowering seniors activity are based on the acceptance by patients and doctor relationship (with an emphasis on understanding and trust). Patient empowerment should focus on social responsibility to take into account long-term costs of telemedicine (after pilot stage the costs will be shifted to the patients), and the risk that non-users of Telehealth will have some serious reservations about the lack of face-to-face contacts. Telecare in the future should fully replace "conventional health care". For the effective implementation of TeleHealth it is necessary to consider more barriers on the side of technology than patients. Another influence of Telecare, which should be considered, is based on patient safety in communications with healthcare professionals. This is why many of EU projects, such as CareWell project, are focused on the impact of e-health solutions concerning the safety of the patient. Based on the CareWell experience it is very important to design networking with other key stakeholders at the EU level to put the patient at the center of attention. In accordance with the principles of equality, prevention is the process by which people can gain greater control over decisions and actions affecting their health (WHO 1998). These key issues in CareWell Projet were mainly based on collaboration, networking, and mutual talks according to the knowledge and expertise of patients, resources and styles of action. Now, patients should be able to plan what to do next, how to go forward and how to build their own solutions to health issues. There is a need to develop initiatives and educational interventions (Cooper et al. 2011) which utilize the principles of adult education who wish to have influence on the feelings, thoughts and behaviour. According to this, it is necessary to use teaching strategies to enable patients to make choices, so that they can transform their knowledge into practice. Therefore, theoretical models of behavior should be incorporated into models of education. Interventions should be designed taking into account the expected effects such as quality of life, personal models of the disease and the empowerment regardless of physical conditions and in accordance with the instructions of self-care. For the entire patient support process and education, it is important that the patient is able to evaluate the quality of life. Therefore, questionnaires about the quality of life (QoL or QOL) will be filled-in as a part of patients enrollment in the CareWell project. Furthermore, the emphasis is put on the importance of questionnaires, which generally relate to the quality of everyday life of the individual, or they asses their well-being (Table 1). This includes all emotional, social and physical aspects of life of an individual in connection with health. In healthcare the quality of life and its assessment is very important. However, every person may be limited in time through an illness or a disability. In this way it can influence the interface of the IT system.

Table 1: List of existing questionnaires.

Questionaire	Description
PIRU	Questionnaire on user experience of Integrated Care
GDS	Geriatric Depression Scale
Barthel Index	Barthel Index Of Activities Of Daily Living
IADL	Instrumental Activities of Daily Living Scale
eCCIS	Questionnaire on service utility and costs.

On the basis of surveys such as PIRU, GDS, the patient empowerment initiative was possible to be taken by the three EU projects: SmartCare, BeyondSilos and CareWell. Implementation of these projects will inevitably deepen the possibility of wide adoption of the technology. Moreover, partners/regions will implement integrated care services with adequate incentives. Currently general care pathways should evolve in the direction of each individual track. This type of interaction will allow the patient to improve the integrated care. Integration is an extremely important aspect of health and social care for millions of people. It is not about the structure, organization and roads, it concerns the achievement of better patient outcomes.

The comprehensive transformation of health care, which is currently delivered to the patient, is based on technological innovation. The high demand for deeper studies concern the improvement of multidisciplinary IT infrastructure, better communications, high-quality data and the highest quality tools. This translational vision is building a vision of Digital Patient (ang. Discipulus Project 2013) .

3.2 Participatory action research model of seniors involvement in informing the design and improving the functionality of the systems.

One of the ways to respond to the need of profound empowerment and involvement of elderly people in designing and using ICT based health systems can be an exploration of participatory action research (PAR), which is a methodological approach based on social sciences. It gives a non- researcher a status of quasi researcher and increases their involvement in designing, implementing and maintaining vital social innovations. Such approaches respond well to the needs of all adults as a learners, social actors and proactive beings.

According to Baum, McDougal and Smith (2006) PAR differs from most other approaches to public health research, because it is based on reflection, data collection and action that aims to improve health and reduce health inequities through involving the people who, in turn, take actions to improve their own health. An attempt of applying the PAR approach in ICT-based health intervention was the idea behind the design of STAMFoRd concept (consortium of H2020 PHC – 21-2015, Advancing active and healthy ageing with ICT: Early risk detection and intervention, http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2268-phc-21-2015.html led by Italian company, ENG, currently in the process of reviewing). The concept intends to realize an innovative, unobtrusive, intuitive and cost effective Indoor Falls Risk Monitoring and Management System for older adults living in nursing homes and/or alone at home. The main goal to achieve is to develop and validate an integrated ICT solution for predicting indoors falls aimed to tailor ICT interventions.

In this concept the interests and the balance of power between the different social groups involved are decisive in determining whether or not the end-users become real participants in the design process. Standard protocols will be developed in STAMFord according to state of the art in the field:

- the level of regulatory controls will be commensurate with the potential risks associated with the technological ecosystem to protect the end-user while ensuring continued access to new technologies and an innovation-friendly business environment;
- regulatory framework addressing a life cycle of the project will include definition and classification of the end-user devices, essential principles of safety and performance, quality system requirements, vigilance system requirements and the use of international integration standards;
- the use of international nomenclature and standards will be encouraged, due to the need of exchange technology between two end-user locations of partners in the project. The level of regulatory control for the end-user should increase with the rise of the degree of risk, taking account of the benefits offered by use of the device.

Since real end-user participation is essential for the successful outcome of the observation and prediction process, another factor will be introduced - a social-science based PAR approach will be used to involve seniors, to empower them and to increase their independence and control over the whole process of engagement in the project's activities, as a complementary protocol to the full compliance with technical requirements.

What would be the added value of PAR, on top of standard protocols of ensuring end- users engagement?

- PAR’s main goal is to enable action. Action is achieved through a reflective cycle, whereby participants collect and analyze data, then determine what action should follow.
- PAR’s approach pays careful attention to power relationships, advocating for power to be deliberately shared between the researcher and the researched: blurring the line between them until the researched become the researchers. The researched cease to be objects and become partners in the whole research process.
- unlike other static approaches, PAR does not remove data and information from their contexts. Most health research involves people, even if only as passive participants, as “subjects” or “respondents”. PAR advocates that those being researched should be involved in the process actively. The degree to which this is possible in health research will differ as will the willingness of people to be involved in research (Baum at al., 2006 , p. 854)

4 CONCLUSIONS

The recommendation for improving future practices in the elicitation of seniors´ involvement in ICT-based health solutions, based on the experiences of ongoing and future projects in the field can be summarized as follows. With aging population it is a vital social and economic purpose to develop complex and, sophisticated, yet easy to use and reliable, ICT systems supporting wellbeing, healthcare and interventions. It calls for sustainable development: in order to make such systems useful and productive there is a need for advancement in the field of technology and a need for advancement in the methods of involving active seniors in designing such services. Both fields need to be balanced, but they cannot be developed separately – there is a call for synergetic and socially innovative approaches. One of the ways to achieve that is to take multidisciplinary approach and equally respect all the sciences: medical, technical, social and humanities - to build integrated systems supporting and empowering all involved participants. The biggest challenge so far is to successfully communicate between all the disciplines, while designing, testing and developing products and services, but without tackling it, the profoundly integrated approaches will not be achievable.

REFERENCES

Baum, F., MacDougall, C., & Smith, D. (2006). Participatory action research. Journal of Epidemiology and Community Health, 60(10), 854–857. doi:10.1136/jech.2004.028662

Cooper at all 2001

Barker, M., Baird, J., Lawrence, W., Jarman, M., Black, C., Barnard, K., Cradock, S., Davies, J., Margetts, B., Inskip, H. & Cooper, C. (2011). The Southampton Initiative for Health: a complex intervention to improve the diets and increase the physical activity levels of women from disadvantaged communities. J Health Psychol. 2011 Jan;16(1):178-91, doi: 10.1177/ 1359105310371397

Digital Patient Roadmap, Discipulus Project, DISCIPULUS Roadmap for the Digital Patient FINAL DRAFT, May 2013

WHO (1981). Global Strategy for Health for All by the year 2000. Geneva: World Health Organisation. Retrieved July 17th, 2015 from: http://whqlibdoc.who.int/ publications/9241800038.pdf

Gosbee, J. and Ritchie, E. (2007). Human–Computer Interaction and Medical Software, Interactions, July/August.

Hendrick, E., Schooley, B. and Gao C. (2013). CloudHealth: Developing A Reliable Cloud Platform for Healthcare Applications, IEEE International Workshop on Consumer eHealth Platforms, Services and Applications, 2013 (accepted)

Norman, D.A. and Draper S.W. (1986). User centered system design: New perspectives on human- computer interaction. Lawrence Erlbaum Associates, Hillsdale, NJ.

Martin, L. R., Williams, S. L., Haskard, K. B., & DiMatteo, M. R. (2005). The challenge of patient adherence. Therapeutics and Clinical Risk Management, 1(3), 189–199.

WHO (1998).The World Health Report 1998 Life in the 21st century A vision for all Report of the Director-General, World Health Organization.

The Application of PEST Analysis to the Creation of the Profile of an IT Product Designed to Activate and Support Senior Citizens in Poland

Beata Butryn, Wiesława Gryncewicz, Robert Kutera and Maja Leszczyńska
Institute of Business Informatics, Wroclaw University of Economics, Komandorska St. 118/120, 53-345 Wroclaw, Poland
{beata.butryn, wieslawa.gryncewicz, robert.kutera, maja.leszczynska}@ue.wroc.pl,

Keywords: IT product, elderly people activation and support, PEST analysis

Abstract: The purpose of the paper is to create a profile of an IT product, using IT tools to collect and to analyse information and enable communication between elderly people to support and activate them. PEST Analysis method was used to determine different factors of influence on the domain of elderly people activation and support Having evaluated the factors and defined the product characteristics, the Authors focused on defining the profile of the IT product by referring the aforementioned characteristics to the structure of the product in the marketing understanding thereof.

1 INTRODUCTION

Demographical analyses in Poland indicate that our society is aging at a very fast speed. The percentage of senior citizens in the Polish population is going to increase in the years to come (GUS, 2014, GUS, 2015, AGE Platform Europe, 2013, Strzelecki, Kowalczyk, 2014). This trend is accompanied by the globalization of the economy and by the development of the information technologies (IT). The two aforementioned processes are highly connected and dependant on each other. It is impossible to imagine social and economic relations without modern IT tools. On one hand, they integrate all the entities operating on the market but on the other, they constitute a reason for the ever-deepening and permanent social divide and exclusion. Those who are particularly vulnerable in this respect are elderly people since they lacked the ability to acquire digital competencies throughout their professional careers (Ferrari, 2012, Frąckiewicz, 2009, Kucharska, 2013). This means the necessity to undertake measures designed to activate and support this social group, for example, by developing IT products suitable for their needs (MPiPS, 2014).

The **purpose** of this paper is to create the profile of such a product. The test method which was used was PEST Analysis. PEST Analysis is a tool designed to analyse macro environmental factors. The factors analysed in the basic form of the method in question are classified into four different categories which cover Political, Economic, Technological and Social factors. The application of the said method allowed to identify and validate the macro environmental factors which influenced the analysed product (Duncan, 1972) (Ward and Rivani, 2005). For the purpose of this article the Authors adopt the assumption that the IT product in question is a platform which uses IT tools designed both to collect and to analyse information as well as to enable the communication between the users.

The structure of the article is following: section 2 – methodology, section 3 – technological factors of the PEST analysis, section 4 - IT product profile and in the last section – conclusions and recommendations.

2 METHODOLOGY – THE ESSENCE OF PEST METHOD

PEST Analysis is also referred to as PESTLE Analysis, PESTEL Analysis , PESTLIED Analysis, STEEPLE Analysis, SLEPT Analysis and LONGPESTLE Analysis (Mindtools, 2015)

(Makos, 2015) (Clulow, 2005) (Voros, 2001) (Narayanan and Fahey, 1994), depending on what factors are taken into consideration. The PEST Analysis is applied to specify political, social, economic and technical conditions in Poland in the context of their influence on the profile of the IT product dedicated to elderly people. In this paper the Authors characterize in details only the technological conditions because of this conference topics. The purpose of the analysis is to provide conclusions on the general characteristics of the IT product designed to activate and support this social group but not to specify detailed functional parameters of the product. What will be indicated in this paper are its basic characteristics resulting from a thorough understanding of the socio-political, economic and technological context. The aforementioned characteristics shall cover such factors as: the living standards of elderly people in Poland, their digital competencies and the access to the Internet in their households.

The Authors regard PEST Analysis as the first stage of research on defining the detailed profile of the IT product dedicated to elderly people in Poland.

The examination procedure involved the following steps:

- Specifying through brainstorming the most significant factors to be taken into account in PEST Analysis.
- Verification of the available research reports, including the statistical ones, and of the available resources in order to carry out a detailed analysis of the factors and in order to specify their impact and likelihood.
- Specifying the influence of the factors on the IT product profile by defining the characteristics thereof.

The summery of the PEST Analysis results is presented in Table 1.

Table 1: PEST Analysis Results

Factor	Impact (from -2 to 2)	Likelihood (from 0 to 1)	Influence (impact x likelihood)	Product characteristics
Political factors				
European and domestic policies facilitating activation of senior citizens	+2	0,5	1	Free
Domestic legal framework	-1	1	-1	Compliant with the legal framework
Prolonging working life	1,8	0,6	1,08	Oriented at offering services and entering into transactions
Economic factors				
Structure of income/expenses	-1	0,7	-0,7	Free
Level of wealth	1,2	0,5	0,6	Free
Social factors				
Age-related biological, mental, social barriers in IT perception	-1,6	0,8	-1,28	Adjusted to elderly people's perception Help desk support provided
Social mobility	1,4	0,3	0,42	Community oriented (relations and communication) Mobile Integrated with popular messengers
Technological factors				
Informatisation level in Poland	2	1	2	Available online
Level of acceptance of technology by citizens	-1,8	0,3	-0,54	Help desk support provided
Condition of telecommunication market in Poland	0,6	0,9	0,54	Using popular communication channels (text messages, e-mail)
Software provision method	1,5	0,7	1,05	Available in SaaS model
Easiness of software developing	0,4	0,9	0,36	Using web standards Open to integration
Technological progress	0,6	0,7	0,42	Easily expandable

The Authors, on the basis of in-depth resource studies, research and statistical reports as well as their own experience evaluated the factors which had been selected at the first stage of PEST Analysis by specifying their impact (on a scale from -2 to 2, where -2 was means factors with a very negative impact, +2 means factors with a very positive impact and 0 means factors of a neutral character) and likelihood (on a scale from 0 to 1, where 0 means unlikely phenomena and 1 means phenomena certain to happen). The influence was evaluated by multiplying one of the aforementioned factors by another.

Furthermore, the Authors determined in discussion the characteristics of the IT product dedicated to elderly people. The said characteristics constitute a response to a given factor and if the factor is negative, they constitute an antidote which is capable of eliminating its influence. At this stage, it is also possible to notice that according to the Authors, what influences the IT product dedicated to elderly people in the strongest, positive way is a high level of informatisation in Poland and what has the most negative influence are the biological, psychological and social barriers in the IT perception related to aging

Having evaluated the factors and defined the product characteristics, the Authors focused on defining the profile of the IT product by referring the aforementioned characteristics to the structure of the product in the marketing understanding thereof. The results of such an analysis are presented in the section 4.

3 TECHNOLOGICAL FACTORS

Activating elderly people in Poland is closely related to their involvement in the functioning within the framework of the information society, which is dependent on the IT development. Therefore, what constitutes a very important element of PEST Analysis in this case is to determine technological factors which can influence the examined field.

The first factor is a **general level of informatisation in Poland**, which determines the ability to use technologies suitable to support elderly people. Three quarters of households in Poland had the access to the Internet in 2014 and in 77.1% of households there was at least one computer. The percentage of households with the access to the Internet and to a computer varied depending on both the type of a household and the class and degree of urbanisation of the place of residence and the region of Poland. These were the households with children which much more frequently had the access to the Internet and computer (the difference between them and the households without children but with the Internet access amounted to 30%). What gained the market was the fixed broadband (59.7%) and wireless broadband (24,4%) (GUS, 2015, Bucki, 2014, UKE, 2014). Moreover, both at the state level and self-governmental level, public services, such as administration, health, public finances, social security etc,. become gradually more computerised. It has been made possible to settle official matters (entirely or partially) via the Internet (eDeclarations, Electronic Services Platform ZUS, the Electronic Public Administration Services Platform - ePUAP, Integrated Patient's Guide – ZIP etc.). What is also available is the electronic stationary equipment designed for the applicant's self-service. The web pages of the government, self-governments and other institutions get significantly enhanced so that they are capable of playing a role of information and communication centres for citizens. Implementing new technological solutions is entirely dependent on the suitable infrastructure. The data presented above clearly indicate that the Polish households as well as the state and self-governmental structures are ready to make use of common technological solutions.

What is related to the informatisation is another important factor, namely the **level of acceptance of technology by the citizens**, especially by elderly people. 39.4% of people at the age between 55 and 64 and 20% of people at the age between 65 and 74 declare that they regularly use a computer (in both groups growth by 12% approximately as compared to the year 2010). Furthermore, among retired people and other professionally inactive people (but not among the unemployed) this rate amounts to 32.4% (in this group among other professional activity groups the highest growth was observed between the year 2010 and 2014 - growth by 11.2%). Nearly 40% of people at the age between 55 and 64 use the Internet regularly (the highest growth by 14.1% as compared to the year 2010) and nearly 20% of people at the age between 65 and 74 declare such a use. It is also worthwhile to note that in the group of retired people and other professionally inactive people 30% use the Internet regularly. In the aforementioned age groups people most frequently use the technology in question at home (GUS, 2015). The foregoing data indicates that the Internet and computer are more and more accepted and used in practice by the analysed population between 55 and 74 years of age. People belonging to this group already have skills necessary to use the operating system, the Internet browser or the e-mail. However, the fact that the majority of elderly people still lack suitable digital competencies has a substantial and negative impact on the effectiveness of implementations of IT solutions which support everyday consumption and administrative matters.

As far as more advanced technologies are concerned, cloud computing (e.g. virtual disks) still remains not a very popular solution. The said technology is used by 1.8% of people between 55 and 64 years of age and 0.9% of people between 65 and 74 years of age (it is used by 8.7% of people as compared to the test sample). Smartphones are used by 7% of people between 55 and 64 years of age and 3% of people between 65 and 74 years of age (generally 25.5% of users as compared to the entire group of respondents). The second popular advanced technology was the Smart TV, which is a television set with integrated Internet capabilities and advanced functions –14% of the total number of respondents owned a Smart TV in 2014 (GUS, 2015). These observations can determine the selection of available communication media which will respond to the expectations of elderly people.

Another factor which is worthwhile to notice is the **condition of the telecommunication market in Poland**. Saturation with telecommunication services on the individual customer market is high. Only 3% of people who are 15 years old or more use no such a service. These are the mobile phones that are most often used –88% of respondents have a mobile phone (in the group of people who are older than 60 this rate amounts to 58% and among pensioners and retirement pensioners it is 60%). A fixed telephone line is connected in 23% of households. Such telephones are much more frequently used by people who are 60 years old or more (48%) and by pensioners and retirement pensioners (46%) (UKE, 2014). What should be emphasised here is the fact that there exists a gradual trend among elderly people towards switching from fixed lines to mobile telephony. This entails the cost reduction and other practical implications (e.g. instant communication or the ability to call for help). Furthermore, there appeared a range of mobile devices which respond to the expectations of elderly people. They are equipped with suitable, clear displays, large fronts and ergonomic buttons. Certainly, what should be taken into account in every case while developing electronic communication with elderly people are the trends which exist on the telecommunication market. A large selection of communication channels and easiness in implementing new solutions result in the fact that this factor does not play a crucial role in the process of implementing technological solutions dedicated to the social group in question.

Technological development is determined by another factor, namely by the **choice of a method of providing software** dedicated to elderly people and to the entities which provide services for this group. What is gaining importance in Poland is the *software as a service* delivery model (SaaS), which is gradually replacing dedicated or boxed software. According to the data provided by PMR, in the year 2013 the value of the Polish market for cloud services increased to 300 million PLN and more than 60% of the market share was seized by the online access to software, which is regarded as the most prospective in the service market. Audytel, which at the end of 2013 conducted a survey among 50 biggest domestic companies, indicates that 20% of the said companies already use the SaaS model and 7% declared the intention to do so (Marszałek, 2014). This model allows the purchasers to use the software in a customized way and to pay for the actual usage, which considerably reduces the costs of purchase and maintenance of the software. From the end-users' point of view, such solutions result in their easier adaptation to a new service if this service uses popular software available in this model (e.g. online shops based on the same software). It is particularly important as far as elderly people are concerned due to the fact that their ability to adapt themselves quickly to new solutions is limited. An informed and right choice of a software provision method constitutes in the majority of implementation cases a crucial factor which proves to be decisive for the successful implementation.

What is connected with the abovementioned issue, is the **easiness of software developing**, which results mostly from the application of the generally accepted standards (e.g. data exchange in XML format) and from the ability to integrate open source applications, dedicated or boxed, with the applications available in the SaaS model. Furthermore, the aforementioned open source software is absolutely free and it has open source code, what results in the fact that new solutions, which integrate such applications with the most popular productivity apps and Internet services, are created very quickly. The vast majority of new technologies improve the software development process (e.g. due to the application of design patterns or frameworks) but also for this reason the aforementioned easiness in the context of technological support of the activation of elderly people cannot be regarded as a factor which has a significant influence on the implementation process.

What also constitutes an important factor is **keeping up with advances in technology** among both the software suppliers and its users. From the suppliers' point of view, this results from the necessity to be competitive and to optimize their activities through a higher level of efficiency and functionality/safety. Users also expect novelties, interface improvements aiming at better, more ergonomic solutions, faster applications and a higher level of safety. This mostly applies to people with higher digital competencies, including some of the elderly people, who are aware computer and Internet users – therefore, this factor is not decisive for the activation of elderly people with the application of technological support.

The technological factors in the field of activating elderly people significantly influence both the shape of information and communication processes dedicated to this social group and the selection of suitable technologies according to the level of digital competencies of senior citizens. In other words, these factors determine a range of characteristics of the IT product which is addressed to this target group.

4 IT PRODUCT PROFILE

According to the marketing definition, every object of the market exchange or anything offered to a market is deemed to constitute a product. Therefore, the notion of product shall not be interpreted as referring exclusively to material objects. Consequently, a product can be regarded as a bundle of benefits offered to the purchaser in order to satisfy their needs (Pindakiewicz, 1997). A lot of contemporary products are a mixture of material and non-material elements (Altkorn, 2000). This is particularly true of IT products. In this sense, the IT product is composed of: a specified software code compilation, the license to use the software and a set of IT services connected with its implementation and maintenance.

From the marketing perspective on the IT product, it is possible to repeat after T. Levitta his remarks on a specified structure of a product. According to him, this structure is composed of four layers similar to the rings which can be observed in a tree cross section (Altkorn, 2000):

- generic product
- expected product
- augmented product
- potential product.

Below you will find the characteristics of the particular structural elements of the IT product dedicated to elderly people. Together they will form the profile of such a product. The description was created on the basis of the conducted PEST Analysis, presented in this paper.

Generic product is a notion which specifies what the purchaser really acquires. In the case of the IT product, it is a particular compilation of its code which is made available to the user in different forms, i.e. in a form of a desktop application or in a form of a service available via an Internet browser. While bearing in mind the results of PEST Analysis as presented above, what seems reasonable in the case of the IT product dedicated to elderly people is to offer the product for the lowest possible price, if not to resign from charging senior citizens at all. Such an approach is recommendable due to the fact that senior citizens constitute a relatively underprivileged group whose expenses in most cases satisfy only their basic needs. On the other hand, the existing possibilities to obtain financing for the development of such products from the UE domestic programmes also facilitate such an approach. Nevertheless, what should be taken into consideration is the fact that because of the limited number of financing sources and because of a high level of competition among the beneficiaries, it might be difficult to obtain financial support. In order to reduce the costs even more, the IT product for senior citizens can be offered as a service in the SaaS model. This will eliminate the necessity to install software and respond to the low level of digital competences among senior citizens as confirmed by PEST Analysis. At the same time the fact that the Internet and computer equipment are common in the Polish households also advocates such a solution. It is essential for the product to be consistent with the local legal regulations and to take into account the rules on the settlement of transactions and the tax treatment of transactions. Furthermore, it should be oriented toward service offerings and entering into transactions in a way which will enable senior citizens to offer their own services and to undertake employment in the form of a part-time job with flexible working hours.

Furthermore, what should be taken into consideration while making the decision to provide senior citizens with the IT product via the Internet is the immensity of offers which can be encountered. From the consumes' perspective, this can result in the complexity of the decision situation and in the difficulty in selecting a suitable service provider. What can constitute an additional barrier for elderly people is the complexity of a web page where the offer is presented. Furthermore, such pages are not prepared to be used by elderly people, who for example, might suffer from sight problems. Therefore, it seems justified to equip the IT product dedicated to elderly people with a plain interface, tailored to the needs of elderly people. The web page designed for senior citizens should be as simple as possible and it should give the possibility to enlarge the text or to use a voice-guided navigation. This will allow to eliminate the factor which has the highest negative influence, i.e. the age-related barriers in IT perception. Furthermore, a web page dedicated to senior citizens should be available mostly via a computer because, as the presented research shows, senior citizens relatively seldom use mobile devices such a smartphones.

What constitutes another element of the product structure is the **expected product,** understood as any additional, useful feature related to the generic product. In the case of the IT product available via a browser, there exists no implementation service which could constitute an expected product but it seems justified to provide senior citizens with the help-desk which will perform a function of a direct, telephone support for them. Thus, the help-desk, which will provide both substantive and technological help, will play a role of an expected product. It will allow to overcome the senior citizens' reluctance to new solutions and to ensure a proper level of trust in the product. The help-desk will provide senior citizens with assistance and introductory training on the use of the IT product and it will perform a function of an intermediary in transactions between senior citizens and service providers. This will allow to ensure a proper credibility of service providers. Ultimately, the role of the help-desk can be broadened to include the support provided to senior citizens in the negotiations with the service providers and in the financial settlements. This will allow to eliminate the lack of trust among elderly people as well as the lack of skills necessary to benefit from online transactions. This is consistent with the marketing concept according to which the content of the expected product is not constant but it is highly dependent on the competitiveness of the market which offers a given product.

It is important to observe here that the elements which over time are perceived by the purchasers as elements of the expected product are initially elements of the augmented product. **The augmented product** is understood as any additional element of the product designed to facilitate the consumption of a given item. While analysing this structural aspect of the IT product dedicated to elderly people, it should be noted that many of them need a training or support before they order a service via the Internet for the first time so that they could overcome their fears related to both the use of modern IT solutions and the lack of trust in the service provider or the lack of knowledge. This problem can be eliminated by the help-desk; however, what can also be offered to elderly people as a part of augmented product are trainings which will be more product-detail orientated or such trainings whose main purpose will be to increase the general IT competencies level among senior citizens.

The term **potential product**, on the other hand, means everything which can make a product more attractive and which can attract the recipient's attention in the future. In the field of IT, what can be regarded as a potential product are its **development visions** presented in the form of the announcements of the new versions. The potential product is composed of the elements of inventiveness, imitation and adaptation ant its shape depends on both the users themselves, who suggest functional improvements and on the producers, who are mostly in charge of the product adjustment to the fast changing economic and legal realities. What can constitute an example of a potential product in the case of the IT product dedicated to elderly people is its extension in the form of the smartphone or smart TV applications which will appear when their current users enter the old age.

5 CONCLUSIONS AND RECOMMENDATIONS

This paper specifies the macro environmental factors which should be taken into consideration while creating a profile of the IT product dedicated to senior citizens.

On the basis of PEST Analysis, the Authors observed that what has the strongest positive influence on such a product is a high level of informatisation in Poland. Due to the existence of a developed infrastructure and due to the falling costs of its use, elderly people have a better access to modern IT products and to the Internet. Furthermore, it should be pointed out that both the domestic and European policies support initiatives dedicated to the analysed social group by providing the source of financing. Thus, IT companies have a possibility to provide senior citizens with a free access to their products and services. At the same time, the SaaS model is gaining importance as a form of software delivery.

At the same time negative influences of several factors of the analysed environment were observed. While creating the profile of the IT product dedicated to elderly people one should take into consideration biological, psychological, social and legal barriers which constitute an obstacle for the users in question. Therefore, it is important to create suitable IT products which will respond to the needs and perception of elderly people and to undertake measures designed to educate them in this field. This will allow senior citizens to benefit from their intellectual capital, experiences and skills. Furthermore, this will help to eliminate the generation gap as well as the digital exclusion of elderly people.

Further research will concentrate around an issue of a strategic analysis of the environment and they will be designed to elaborate elderly people activation scenarios with the usage of IT tools.

REFERENCES

AGE, 2013 *AGE Platform Europe*. Retrieved July 9, 2015, from: http://www.age-platform.eu.

Altkorn, J., 2000. *Podstawy marketingu*, Wydanie III zmienione i poszerzone, Wydawnictwo OO. Franciszkanów, Instytut Marketingu, Kraków.

Bucki, P., 2014. *Coraz więcej seniorów na świecie korzysta z Internetu i mediów społecznościowych*, Retrieved July 10, 2015, from: http://www.biznes.newseria.pl.

Clulow, V., 2005. *Futures dilemmas for marketers: can stakeholder analysis add value?*, European Journal of Marketing, 39 (9/10).

Duncan, R.B., 1972. *Characteristics of organizational environments and perceived environmental uncertainty*. Administrative Science Quarterly 17(3)

Ferrari, A., 2012. *Digital competence in practice: An analysis of frameworks*, Publications Office of the European Union.

Frąckiewicz, E., 2009. *Wpływ Internetu na aktywność seniorów z punktu widzenia budowania strategii społeczeństwa i2010*, Zeszyt - Centrum Europejskie Natolin, Centrum Europejskie Natolin, Warszawa.

GUS, 2014. *Sytuacja demograficzna osób starszych i konsekwencje starzenia się ludności Polski w świetle prognozy na lata 2014-2050*, GUS, Warszawa.

GUS, 2015. *Społeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2010-2014*, Warszawa.

Kucharska A., 2013. *Pracownik 60+*, Główny Inspektorat Pracy, Wydanie pierwsze 13027/01/00, Warszawa, Retrieved July 9, 2015, from: http://www.pip.gov.pl

Makos, J, 2015. *What is STEEP Analysis*, Retrieved July 9, 2015, from: http://pestleanalysis.com.

Marszałek, K., 2014. *SaaS a sprawa polska*, Retrieved July 9, 2015, from: http://www.crn.pl/artykuly.

Mindtools, 2015. *PEST Analysis. Identifying "Big Picture" Opportunities and Threats*, Retrieved July 9, 2015, from: http://www.mindtools.com.

MPiPS, 2014. *Rządowy Program na rzecz Aktywności Społecznej Osób Starszych na lata 2014–2020*, Retrieved July 9, 2015, from: https://www.mpips.gov.pl.

Narayanan, V. and Fahey, L., 1994. *Macroenvironmental analysis: understanding the environment outside the industry*. In Fahey, L. and Randall, R. (Eds.), *The Portable MBA in Strategy*, 2nd edition, New York: Wiley.

Pindakiewicz, J.,1997. *Podstawy Marketingu*, Oficyna Wydawnicza SGH, Warszawa.

Strzelecki, Z, and Kowalczyk, E., 2014. *Przemiany ludnościowe w Polsce. Przyszłość demograficzna*, Warszawa.

UKE, 2014. *Rynek usług telekomunikacyjnych w Polsce w 2014 roku*, Warszawa.

Voros, J., 2001. *Reframing environmental scanning: an integral approach*, Retrieved July 9, 2015, from: http://www.stadsadviseur.nl.

Ward, D. and Rivani E., 2005. *An Overview of Strategy Development Models and the Ward-Rivani Model*, Economics Working Papers.

Home Health Smart TV
Bringing E-Health Closer to Elders

Leon Dragić[1], Igor Piljić[1], Mario Kovač[1], Hrvoje Mlinarić[1], Petar Franček[1], Martin Žagar[1], Vlado Sruk[1] and Darko Gvozdanović[2]

[1]*Faculty of Electrical Enginnering and Computing, Zagreb, Croatia*

[2]*Ericsson Nikola Tesla d.d. Krapinska 45, Zagreb, Croatia*

leon.dragic@fer.hr, igor.piljic@fer.hr, mario.kovac@fer.hr, hrvoje.mlinaric@fer.hr, petar.francek@fer.hr, martin.zagar@fer.hr, vlado.sruk@fer.hr, darko.gvozdanovic@ericsson.com

Keywords: E-Health, Smart TV, Social network, Elders, Healthcare

Abstract: In this paper we present the novel platform for accessing multimedia e-health content adapted for elders. ICT solutions offered on market today are not adapted to elder people. Home Health Smart TV was developed in order to solve this issue. Integration with social network features, like video communication and personalized calendar, solves another concern for elder population - social isolation. Elder population is accustomed to TV devices and with this specially designed platform they can access their medical data and communicate with friends, family or medical staff in a simple way that doesn't require computer literacy.

1 INTRODUCTION

An aging population has become a demographic trend of the majority of developed societies (R. Moody, 2006). Due to this trend, it is a known fact that healthcare costs could double among EU member states by 2060 (Kovač, 2014). This problem is addressed in the European Commission's eHealth Action Plan 2012-2020 (EHealth Action Plan, 2012) which states that ICT solutions should be applied to health and healthcare systems to increase their efficiency, improve quality of life and unlock innovation in health markets.

From 2007 to 2013 percentage of the individuals aged between 16 and 75 who used the internet for seeking health-related information in EU increased from 24 to 44% (Eurostat). This trend shows that people are willing to actively participate in decisions that concern their medical condition and active participation is connected with better health outcomes (Heisler, 2015).

Access to medical data for patients is often provided through patient portals. These solutions, although useful for some patients, require substantial technical knowledge which makes them unusable for elder patients with low computer literacy (Lober, 2006).

Another problem that is of particular concern for older people is social isolation. (Ages, 2013). Finding ways to support people to make and maintain social connections should be a priority for public health particularly in the area of aged care. As such, older adults are potentially at greater risk of becoming socially isolated, and of experiencing the negative health consequences of this. In order to bring ICT solutions closer to elder population it is necessary to address barriers to technology adoption and consequently develop well-designed system that can be used even if the end user is technically illiterate (Independent Age, 2015).

In this paper we present a novel platform "Home Health Smart TV" specifically designed for older people. Using this platform, older people can access their medical data, education materials, history of medical measurements and other medical related information. The platform also addresses problem of social isolation for older people and provides video communication between different users. This way patients can easily communicate with their friends, family and/or medical staff.

In order to minimize the certain fear that older people tend to have regarding technology, the platform is developed for the device older people are familiar with – TV.

2 HOME HEALTH SMART TV

Even though they are technically illiterate, elder people are using TV on daily basis. Studies showed that adults aged > 65 years spent threefold more waking time watching TV than young adults (Depp, 2010). Elder people are interested in their medical conditions but are unable to access that information because technical knowledge is prerequisite (Leist, 2013). Instead of adapting elders to ICT, ICT should be adapted to elders. As a result, Home Health Smart TV platform was developed on Faculty of Electrical Engineering and Computing in Zagreb. Home Health Smart TV is simple-to-use home system, connected to the TV that overcomes the barriers, offers improvement of patient's health care quality and enables easy access to information and data shared between patient and healthcare providers.

Proposed platform addresses two main issues explained in the introduction of this paper. Consequently, the platform can be logically divided into two main domains:

1) E-health domain
2) Social network domain

In e-health domain, Home Health Smart TV acts as a patient portal for accessing medical data using TV device. It connects to different e-health service providers, retrieves the data, filters it and presents it to the patient.

In social network domain Home Health Smart TV enables direct message and video communication between different users in the system. Patient can navigate through the contact list and initiate or join video call. Video and message communication can be used for everyday communication between patient and family but also for communication with medical staff.

2.1 E-Health domain

2.1.1 Platform

Home Health Smart TV is carefully designed package of Android applications deployed on specific Android device. The package consists of 3 applications:

- Home Health Smart TV application
- Auto Update application
- Home Health Smart TV Launcher

Home Health Smart TV application is core part of the Home Health system. It is used for accessing and displaying patient medical data, playing various multimedia educational content, message communication between patients and health-care professionals, etc. The medical information that is showed to the patient is carefully picked by medical experts. Piling too much information can be counterproductive and make the important information less visible, so it is important to filter the most important data before displaying it to user.

Home Health Smart TV application is divided in 5 fragments:

- Home
- Messages
- Education
- Measurements data
- Social networking

Home fragment is used for providing general information to the patient. Patient can check information about the doctor, read the weather report and check if there is any new message in his mailbox.

Message fragment is used for reading received messages from healthcare experts. In the current version, patients cannot reply to the received messages but in future version, set of predefined answers will be available for patient to select. These kind of message exchange can be used for cancellation of appointments or requesting new e-prescription from doctor.

Education fragment is used for browsing educational content stored on the device and displaying the selected content to user. Education material is categorized by types of illnesses which helps patient to focus on specific problem. Education material can be in a form of video, text, photo or presentation. In the future, educational material can be stored on remote server and patients could access it through the internet.

Measurements fragment shows medical measurements displayed as timeline chart. Medical measurements are displayed in a period of 1 year. This way patients can see the progress they are making by regularly exercising, following diets and advices from medical staff which is a great way to empower and motivate the patients. Without having graphical representation of the progress patients can have difficulties to feel any improvements in their medical condition which results in demotivation and thus quitting the regular healthy routines.

Social networking fragment shows the list of contacts user can interact with. Every contact is represented by name, relationship to the user (family, friend, nurse, etc.) and network status (online, offline). Using this fragment, patient can establish video communication with any contact that is online or leave a notification to offline user.

The main application is accompanied with 2 more extensions: Autoupdate Application and Home Health Smart TV Launcher. Autoupdate application, as the name suggests, is used for automatic update of the application without any user input. The application checks the application version deployed on the server. If there is a new version deployed on the server, the Autoupdate application will fetch and install it without any intervention of user. Usual solutions for updating the application, such as Google Play service, can't be applied for Home Health Smart TV platform because updating of the application must occur even if the application is executing. This way we ensure that the latest version of the application is installed on every active device which is very important for users without technical knowledge.

The last application from Home Health Smart TV platform is Home Health Smart TV Launcher. Smart TV Launcher is used as default Android Launcher. Using this custom launcher, we ensure that users cannot exit the Home Health Smart TV application. Making Smart TV Launcher as a default Android Launcher will also provide us automatic start of the application after the boot process.

2.1.2 Design

Primary requirements for application designed for elders is simple design, intuitive navigation, filtered content and user interface adapted to overcome the lack of computer literacy and the physical barriers such as visual impairment. The main input device inside the application is remote control. In order to simplify use of remote control, most of the buttons were disabled, leaving only 8 buttons functional.

Navigation within the application is pretty intuitive. Navigation bar is located at the bottom of the screen. The title located in the middle is the title of currently shown fragment. The titles on the left and the right of the navigation bar show which fragment will be displayed to the user if matching navigation button is clicked.

2.1.3 Integration with e-health service

The application is connected to Ericsson Mobile Health which provides patients personal medical data through Medical node REST API. The interoperability is a great challenge (Vida, Lupse, Stoicu-Tivadar, 2012) when it comes to e-health services. Market is filled with various e-health solutions from different service providers. In order to make the Home Health Smart TV platform adaptable for different environments and e-health services, a special focus is put in adding the module for processing HL7 CCD standard for medical data exchange. The Continuity of Care Document (CCD) is a joint effort of HL7 International and ASTM (HL7 CCD). CCD fosters interoperability of clinical data by allowing physicians to send electronic medical information to other providers without loss of meaning and enabling improvement of patient care. CCD establishes a rich set of templates representing the typical sections of a summary record, and expresses these templates as constraints on CDA. These same templates for vital signs, family history, plan of care, and so on can then be reused in other CDA document types, establishing interoperability across a wide range of clinical use cases.

The platform makes periodic requests to the adapter service which then polls the REST APIs of different service providers. Adapter service is used for adapting internal methods to various external APIs. The communication process is secured using HTTPS protocol. During the first boot and initialization of the device, user enters his credentials for e-health services he is using. The credentials are then locally stored using *AccountManager* class. Entering passwords using remote control is a very difficult and time taking process, so instead of entering passwords on every device boot, patients can secure the access to their medical information using PIN which they can enter after the device initialization. PIN is not mandatory because memorizing it can be an issue to elder people and can be a reason for ceasing the use of the service.

2.1.4 Customizing Android ROM

Even after integrating all previously mentioned modifications, users still need to configure Android OS settings to be able to use Home Health Smart TV platform properly. Although these are minor adjustments, they still require technical knowledge that elder people don't possess. Since the goal of this project is to make platform that is fully adapted to the needs of elder people, Android ROM had to be customized. Default Android ROM is filled with different applications which are not necessary for functioning of Home Health Smart TV. On the other hand all those applications can interfere with normal functioning of the application. All the applications that are not needed for normal functioning of the system were removed. The buttons that are not used inside the application were disabled. The script that executes on first boot of the device was modified with methods for configuring the device settings such as Wi-Fi configuration, time zone configuration, default launcher selection, etc.

2.1.5 Devices

Home Health Smart TV is bundled with 2 devices: Android Set-top box for executing the Home Health Smart TV platform and Wi-Fi 3G Router for internet connection.

Android device was chosen for its ability to connect to different types of TVs. Newer TVs can be connected to the device using HDMI and the older ones can connect using composite video interface. The device is equipped with web camera which will be used in future upgrades of the platform as a way for direct video communication between different stakeholders. Customized Android ROM contains the information about Wi-Fi Router which enables internet access without any additional configuration of the Android OS.

2.2 Social network domain

In a modern society, where social media has altered the basic rules of communication, gap between generations is largely caused by different perception of technology. Main issue of today's communication solutions, such as social networks, smart phones and applications, is that they are adapted to younger population and require technical knowledge. This phenomenon is known as digital divide and is topic of numerous researches done in past years (Norris, 2001) (Broos, Roe, 2006) (Vie 2008).

Statistics show significant difference in percentage between age groups of people that use internet to connect to social networks. While 85% of people between the age of 16 and 25, and 52% of people aged between 25 and 54 used internet to connect to social network in 2014, only 16% of people older than 55 did the same (Eurostat).

Since elder patients are major part of healthcare in every country, integration of e-health with social network has multiple benefits; displaying chosen medical content to patients over internet saves money and time to medical staff, while enriching e-health system with some features that users can find interesting and useful increases usage of the system, but ultimately end user/patient is the one who benefits the most.

2.2.1 Design

Designing social network that doesn't depend on platform is imperative. This way younger population can use smart phones or PC's to access the application, while elders can use TV platform. User interface must provide easy navigation and optimal viewing experience across wide range of devices. Most applications address this issue with responsive web design, which solves the problem of displaying content on different screen sizes, but adaptation to TV platform, where remote control is main input device, is more challenging.

One of the challenges is how to provide content that will attract both, younger and older. Since the final goal is to resocialize elders and make them use modern technologies for communication with their beloved ones focus is set on satisfying their needs, but in order to foster communication between generations younger population must find this application usable as well. Establishing video communication between parties regardless of device type and location is one of the most important built-in functionalities. This is accomplished via high quality communication library. Although it is possible to connect with other users by manually adding them to your network, smart grouping system is applied to facilitate interconnection. Different characteristics, like geographic location, affiliation to certain organization or similar interests, can be used to form a group. Users within a group can chat, share files or make a conference video calls. Personalized calendar for appointments and reminders with external data import support enables easy time tracking. These features along with messaging and news services, unavoidable options for any social network, form the basis of a desired virtual society.

2.2.2 Video communication

Video communication is the cornerstone of introduced social network. It is used for sharing live video and audio streams between two or more actors in the system. By introducing different roles in the system, video communication becomes even more powerful tool. Everyday communication with family and friends can be extended to remote consultations with medical personnel or social services, remote diagnosis, second opinion, remote consultations and therapies to multiple patients at same time and more.

Patients can also join different groups, depending on their medical condition where they can share experiences and advices. It is important to establish community where patients will interact on daily basis and support each other in dealing with their diseases.

Figure 2 shows example of establishing video conference between patient, his doctor and one of his family members.

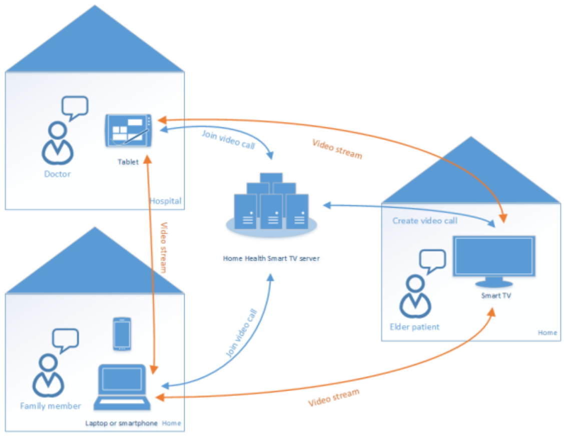

Figure 1: Establishing video communication

2.2.3 Personalized calendar

Following therapies accurately is essential and in a case where there are several therapies prescribed, tracking them can present a challenge, especially for elder patients that tend to be forgetful. The same is true for doctor's appointments that can be scheduled months in advance and patients have to take a note of that.

To enable easy time tracking interactive personalized calendar is available. Every doctor's appointment or therapy is imported in calendar automatically and reminder is triggered before each event.

Some simple processes, like extending continuous therapy, can be very time consuming for both, patient and medical staff. To avoid this problem additional feature is added so the communication between medical staff and patient can be established via calendar.

Extending continuous therapy via calendar is done in four steps:

- Patient chooses therapy he wants to extend
- Medical staff receives request and process it
- Patient is notified after the decision has been made
- If extending continuous therapy is approved calendar is automatically updated and prescription is issued

2.3 Pilot

Home Health Smart TV platform will be tested in pilot project. 50 elder patients that are suffering from at least 2 chronic diseases are included in pilot from which 20 will receive the Home Health Smart TV platform while others will receive mobile device for accessing their medical data. Pilot project will last for 16 months. For this purpose, special service for statistical analysis was developed and deployed on server. Server will receive information about every action users have made. After the 16 months, a lot of valuable data will be gathered from which we will be able to conclude in which direction future work needs to be focused.

3 CONCLUSION

E-health systems and services offer an important complement to routine clinical care. They have become a necessity in most countries and the result is health market filled with numerous ICT solutions. Many of these solutions offer patients access to their medical data through patient portals but the use of these services requires technical knowledge and therefore they are not adapted for elder population. Finding ways to support people to make and maintain social connections should be a priority for public health particularly in the area of aged care. Even though social networks are not complex applications, accessing them can represent a challenge for elders because they require platform-specific knowledge. The solution for these problems is adapting ICT to elders and not vice-versa. A platform was developed on Faculty of Electrical Engineering and Computing in Zagreb in order to address this issue. Home Health Smart TV is simple-to-use home system, connected to the TV that overcomes the barriers and offers improvement of patient's health care quality and enables easy access to information and data shared between patient and healthcare providers. Integration with social network and introduction of new features, such as video communication and personalized calendar, increases usability and functionality of the system and thereby improves well-being of the patient. The functionality of the platform will be tested in the pilot project which will last 16 months and will gather valuable information about the usage of the platform. The information will then be used as guidelines for future work on this project.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the project "Information and communication technology for generic and energy-efficient communication solutions with application in e-/m-health (ICTGEN)" co-financed by the European Union from the European Regional Development Fund and from project “Carewell - Multi-level integration for patients with complex needs” co-funded by the European Commission within the ICT Policy Support Programme of the Competitiveness and Innovation Framework Programme (CIP).

REFERENCES

Harry R. Moody. Aging: Concepts and controversies. Pine Forge Press, 2006

Mario Kovač, "E-Health Demystified: An E-Government Showcase“, Computer, vol.47, no. 10, pp. 34-42, Oct. 2014,doi:10.1109/MC.2014.282

EHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century, 6.12.2012

Heisler, Michele et al. “The Relative Importance of Physician Communication, Participatory Decision Making, and Patient Understanding in Diabetes Self-Management.” Journal of General Internal Medicine 17.4 (2002): 243–252. PMC. Web. 23 Apr. 2015.

Eurostat, Individuals using the internet for seeking health-related information (tin00101) - http://ec.europa.eu/eurostat/web/informationsociety/data/main-tables.

Lober, WB et al. “Barriers to the Use of a Personal Health Record by an Elderly Population.” AMIA Annual Symposium Proceedings 2006 (2006): 514–518. Print.

AGES project, “Activating and Guiding the Engagement of Seniors through Social Media”, T. Morton, A. Genova, Final report, 2013.

Independent Age, Report “Older people, technology and community”, 2015.

Depp, Colin A. et al. “Age, Affective Experience, and Television Use.” American journal of preventive medicine 39.2 (2010): 173–178. PMC. Web. 23 Apr. 2015.M. Young, the Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

Leist A, K, Social Media Use of Older Adults: A Mini-Review. Gerontology 2013; 59:378-384

Vida, M.; Lupse, O.; Stoicu-Tivadar, L., "Improving the interoperability of healthcare information systems through HL7 CDA and CCD standards," Applied Computational Intelligence and Informatics (SACI), 2012 7th IEEE International Symposium on , vol., no., pp.157,161, 24- 26 May 2012.

Continuity of Care Document (CCD®) - HL7 http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6

P. Norris, "Digital divide: Civic engagement, Information Poverty, and the Internet Worldwide" Cambridge University Press, 2001

A. Broos, K.Roe, " The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents",Issues 4-5, 2006 , pp 306-317

S.Vie, "Digital Divide 2.0: “Generation M” and Online Social Networking Sites in the Comp

Kovač M, “SUSTAV ZA PREGLED MEDICINSKIH PODATAKA NA TELEVIZORU“, Patent Pending HR P20140013

Patient with Complex Needs - Experience in Implementation of LSV-Carewell Platform

Kazimierz Frączkowski[1], Antoni Zwiefka[2], Marcin Zaremba[2] and Krzysztof Sikora[3]
[1]*Wroclaw Univesity of Technology, Wyb. Wyspianskiego 27, Wrocław, Poland*
[2]*A. Falkiewicz Specialist Hospital, Warszawska 2, Wrocław, Poland*
[3]*IMMD Health, Institute Medical Mobile Device, Krawiecka 3/10 Wrocław, Poland*
Kazimierz.fraczkowski@pwr.wroc.pl, antoni.zwiefka@dolnyslask.pl, mar_cinzaremba@o2.pl, krzysztof.sikora@immdhealth.com

Keywords: Home telemonitoring, mobile technology, chronic illnesses, quality of life

Abstract: The paper describes telecare procedure concerning patients aged between 65-85 years with at least 2 chronic diseases including hypertension (ICD I10), diabetes (ICD E 11), chronic obstructive pulmonary disease (ICD J44) or heart failure (ICD J50). Ultimately, the project will involve 100 patients qualified on the basis of medical history (last stay in hospital) divided in two groups. In the qualified group, each patient must have at least 60 points according to the Barthl scale. The first group comprises 50 patients taken care of by tele monitoring, which was provided to them with measuring equipment to be able to assess the selected parameters at home (depending on the chronic disease). Then, the results are transmitted automatically via mobile phone network to LSV Telecare (Lower Silesia Voivodeship Telecare) system. The other group of patients consists of those, who were not included in the home monitoring. They are provided with medical care within the current Polish health care system. In this paper we present scenarios and models of business processes, necessary to achieve the objectives of the Care Well project, which is implemented under Competitiveness and Innovation Framework Programme 2007-2013 (project "Multi-Level Integration for Patients with Complex Needs"; grant agreement no: 620983) . The project involves 13 partners from 8 EU countries - project duration is 36 months. The technical parts of the project include tests and examinations of the economic and social effects, as well as indicators of the quality of life based on ICT platform for communication and exchange of medical data, that are essential in the treatment of patients qualified for Telecare.

1 INTRODUCTION

In Poland, among the population of people over 65 a man lives in good health, on average, to 74, and a woman to 78. All scenarios predict that by 2050 percentage of the population aged 65 years and more will double, ie. from 15.8% in 2013 to 31.3% in the low scenario, and to 35.7% in the very high scenario. In the same period the number of the aged 85 and over, is expected to increase five times [http:// Stat.gov.pl/ obszary, 2015; Population Projection 2014-20150, Warszawa 2014].

In the last 50 years the number of people aged 60 and over has tripled and it is expected that it will arrive triple again to 2 billion by 2050. In China, there is a region, that the increase of the proportion of people over 65 years is estimated to reach 22,7% by 2050 [Guy Pare, Mirrou, 2010]. Despite the prolongation of life expectancy, people are not healthy for longer time, on the contrary external conditions and changes of the traditional model of life are often the cause of the emergence of new and chronic diseases of civilization (diabetes, heart disease, hypertension etc.), a list of which according to the Health Minister regulation from 2009 (Dz.Uz 2009, No. 212, item. 1647) in Poland consists of 41 items. And it is estimated that these diseases cause about 60% of all deaths in almost all countries [http://Stat.gov.pl/obszary, 2015; Population Projection 2014-20150, Warszawa 2014].

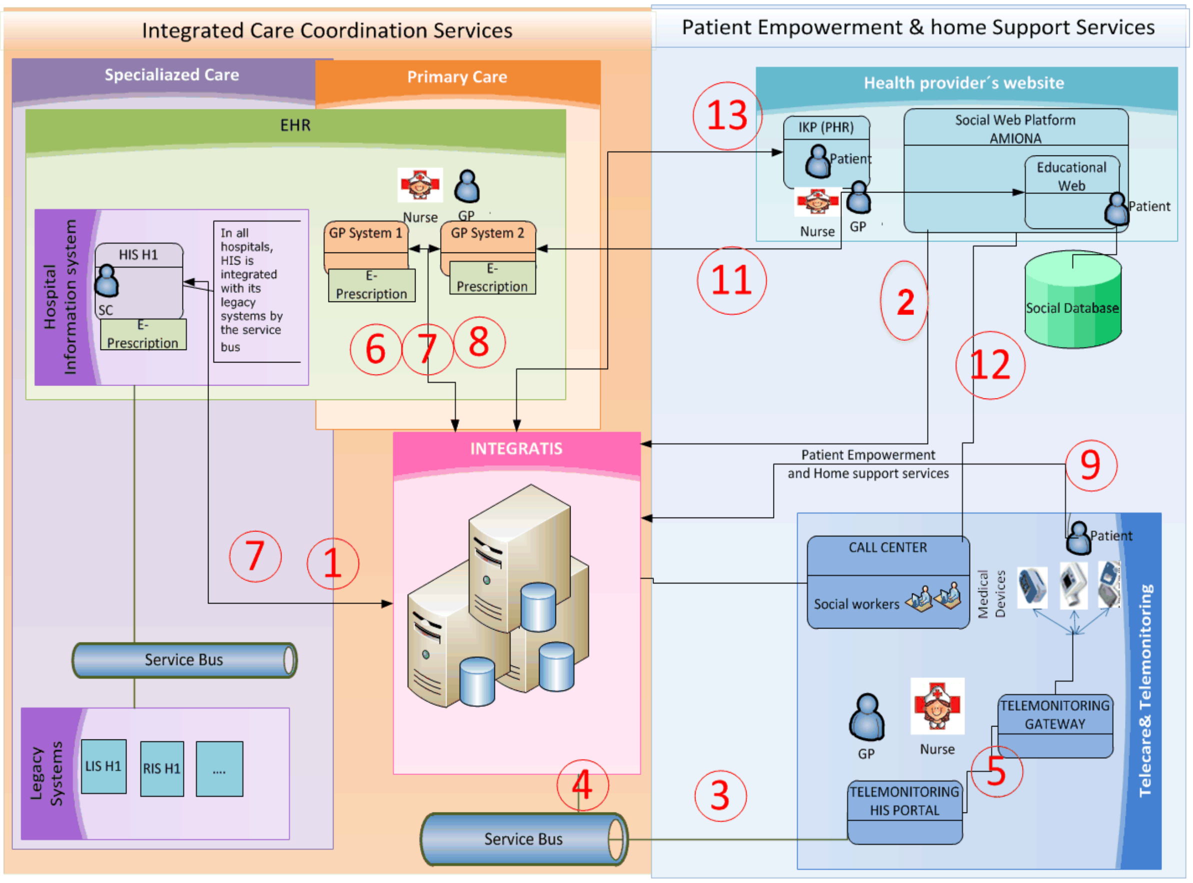

Figure 1: CareWell ICT Architecture Lower Silesia

That is why, there has been set up a number of programs supporting frail patients aged 65+ with more than one chronic disease. One of them is CareWell project, which is aimed at improving the efficiency and quality of medical care. The project is carried out in a consortium with the participation of 13 partners from 8 EU countries (Spain, the United Kingdom, Italy, Croatia, Germany, Belgium, Poland). The main objective of the CareWell-LSV project activities is to prepare ICT platform services based on the modelled business processes, which include 100 patients who are qualified on the basis of the analysis of a disease history (discharge extract from hospital).Therefore in many countries and regions in Europe there are being implemented new systems supporting telecare [Alan Wailer, Tony Maitby 2012]. A new innovative platform, created in Lower Silesia Voivodeship (LSV), for telecare is based on the integration of the already developed applications.

In LSV TELECARE the results from mobile measuring devices are collected in Monitoring

Platform and automatically transmitted to the Integration Platform

To ensure a better quality of life for people aged 65+ with chronic diseases, it is needed to bear the economic burden of chronic diseases, which represents 46% of the global burden caused by these diseases. This phenomenon may be supported by information, communication technologies - mainly mobile ones, and that is an area called m-Health.

All possible questionnaires (including Bartel scale) have been in CareWell Platform within Integration Platform.

A qualified group of patients must have not fewer than 60 points by Barthel Scale. The first group of 50 patients that will be tele monitored, has been provided with measuring device for measuring the selected parameters at patients home. After the measurement is completed the results, are sent (via mobile telephone network) to the CareWell Lower Silesia System. The other group of 50 patients is not monitored at home and they are provided with medical care within the system according to the current health care model in Poland.

2 SYSTEM ARCHITECTURE AND FUNCTIONAL EXPERIENCE

Key aspects in the design of a modern system of telecare is the integration of technological solutions, and the existing information systems, as well as the applicable procedures of patient care with mobile technologies in telecare. These issues are the subject of numerous works on computer systems and their clinical effects [Guy Pare and all, 2010, Guy Pare and all 2010, Spyros Kitsiou and all 2015]. These actions should find a way to show the benefits and how to teach the end user, that is a patient, how to operate mobile measuring devices at home. Therefore, the platform building design involves the integration of three sub platforms:

- Educational - Information Platform (Social)
- Integration Platform (Service Buss)
- Monitoring Platform

As Lower Silesia currently does not have many IT systems implemented to support the delivery of care or share information, both CareWell pathways will be significantly improved with the proposed ICT-enabled services and functionality.

The development of a platform, presented in Figure 1, is to provide interoperability between different IT systems used in primary and secondary care. It will enable information to be shared by various care practitioners and patients within new functionalities:

1. Registration of patient referrals for home care and telemedicine. This is the first task in the LSV Telecare platform.
2. Logged user access to the Information - Education Portal and to Integration Platform.
3. Patients Registry Update Service in HIS by Integration Platform.
4. Service of research results transfer by HIS Patient Portal to Integration Platform.
5. Registration of the performed patient results in HIS Portal.
6. GPs access to EHR and their tasks supporting LSV Telecare procedure.
7. Nurses access to EHR, and their tasks supporting LSV Telecare procedure.
8. Patients access to their own EHR and their tasks supporting the process of LSV Telecare procedure.
9. Implementation of developed services at the country level, like e-Prescription (P1 Project) within LSV Telecare procedure.
10. Call Centre staff access to their own tasks supporting LSV Telecare procedure. - receiving e-mail and SMS alerts.
11. Doctor, nurse and patient access to the Education – Information Portal.
12. Call Centre staff access to the and Education-Information Portal.
13. Some of the developments and changes which will revolve around the new interoperability of Integra TIS system.

Each of the above mentioned systems functionally meets the requirements of the identified key aspects. The Monitoring Platform is responsible for operation of measurement devices. The most important issue is the reliability of measurements in the context of user authentication. It is unacceptable to assign mistaken measurements to a patient. On the one hand, suitable authentication, authentication and data security, and on the other hand, greater ease of use and reliability. To meet these requirements it is necessary to take these constraints into account at the stage of designing a subsystem, which is supposed to manage its tasks.

The next subsystem is an Integration Platform whose main task is to integrate all the subsystems and enable their use to cover specific requirements which are put before the health care system in the region (country). At the stage of preparation for implementation there are identified requirements, which are then transferred into the BPM process model.

This model is consulted with specialists and then their approval is followed by the implementation of telecare process. The Integrated Platforms mainly task is to take care of the implementation of telecare in accordance with the modelled procedures and allow for an adequate response in any situation. Another task of the Integration Platform is the storing, processing and sharing of EHR.

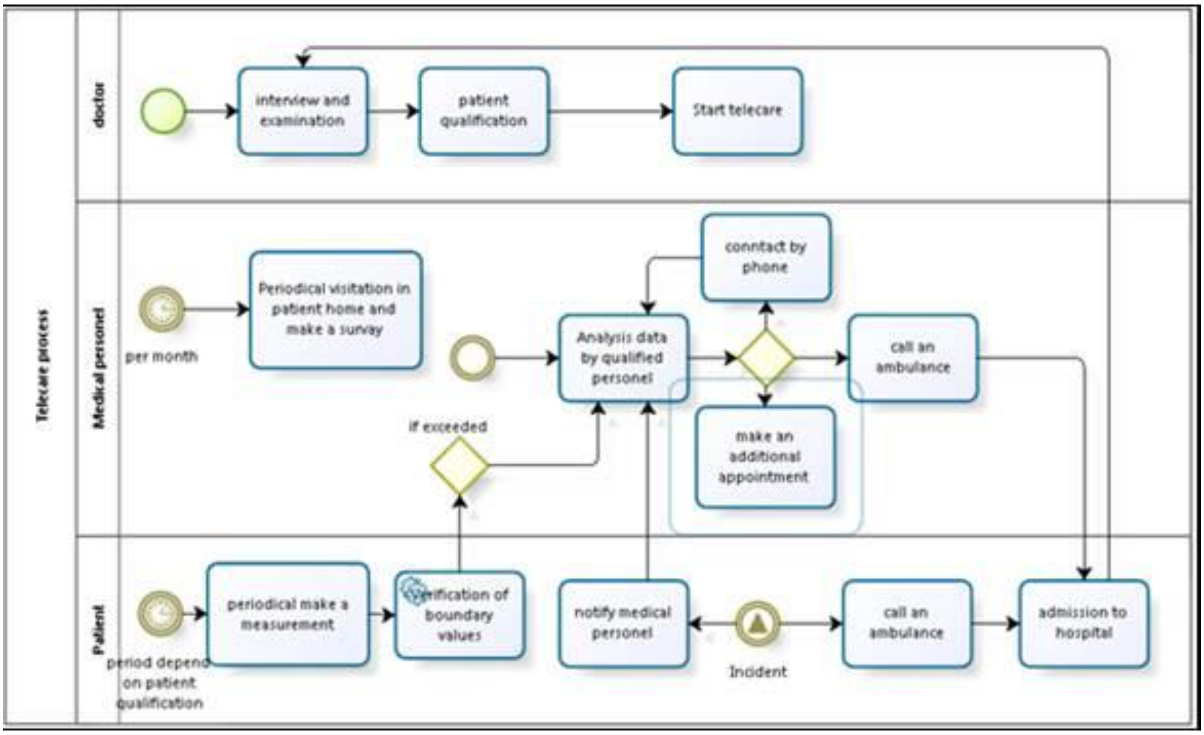

Figure 2: CareWell Homecare process model

Another aspect, which is equally significant as the previous ones, is to enable patients to benefit from telecare services in a safe way that they can understand. Facing the problem of an aging population and the fight against social exclusion, it becomes increasingly important to educate the public, create the opportunity for people to learn about and understand a model of telecare and the benefits it brings. The most important task, as well as most difficult one to be completed by that sub platform is to educate patients make them aware that the use of telecare increases their safety and a quality of life. Confronted with the standard model of health care, telecare give you more benefits. Social portal functionality also means to patients an easy access to their care history (of the disease), the possibility of being kept informed with their results and the feeling of having more control over the process of health care.

2.1. BUSINESS PROCESS MODEL FOR LSV - TELECARE

The first step in implementation of *LSV - TELECARE is suitable qualification of patients and, then dep*ending on its outcome configuration of the appropriate telecare procedure. This is important because the process of telecare which is implemented in the system, described crucial flow of information and tasks, but does not define how various steps have to be performed by individual patient

The telecare process of the Lower Silesia CareWell System assumes that at fixed intervals a patient will perform life parameters measurements at home and the results will be transferred to a healthcare unit. In contrast to the old style home care, the telecare results have to be checked by a doctor who has to determine what specific tests and at what intervals the patient should pursue. Of course, during the process there may be a need to change some details such as measurements distance.

While the results of the patients' measurements flow into the central system, algorithms analyze the results and examine whether they exceed the thresholds, and check if their behaviour is similar to the expected one. If there is a record of a departure from the norm, in the system appears a task of analyzing these results by hospital staff - in our case, by a nurse.

Her task is to verify whether the test was carried out in a correct way, whether the patient at that time may have taken any medicine responsible for the distortion of the results or if his behaviour did not affect their values (e.g. increased physical activity). When the observed anomaly is an erroneous measurement or it is caused by a human error, the patient is recommended to repeat the test. If it is a worrying signal which may endangers the patient's health a nurse can contact a doctor or intervene immediately by calling an ambulance to the patient.

And here we meet another phenomenon described in telecare procedure that is intervention, which we understand as the situation, that is caused by an undesired phenomenon (e.g. accident) or it is a significant deviation from the standard implementation of the procedure. The incident may be reported by the patient in two ways. In the first the patient using the supplied phone numbers call the Call Center (in hospital conducting this procedure), where he can obtain help from a nurse. In some situations, a nurse may consult it with the doctor. She can also arrange a home visit earlier, or in special situations call an ambulance to the patient. The situation in which a patient calls the emergency room directly is considered to be an incident. Then he is admitted to hospital according to standard procedures that he undergoes, and after their completion (after begin discharge) the patient record is supplemented with an extract from hospital.

In the course of the procedure there are also anticipated periodic visits by a nurse in the patient home. Normally this is done once a month. Although in case of incidents appearance, their frequency can be increased.

Once the telecare goal is reached, a patient visits a doctor, who may decide to continue the treatment or end the procedure. In the case of telecare procedure termination, there is generated an automatically record of results and doctor prepares a detailed report for the whole period covered by telecare.

2.2 Stable Patients – out of hospital care

The implementation of the LSV teleCare integrated pathway will enable the following developments to the service model:

Better understanding of the roles and responsibilities of the different care practitioners involved in delivering services and interventions within the care pathway.

Integrating the hospitalisation of those patients who require it as part of the care pathway to provide better patient care transition experiences across the different sectors and professionals.

Introduction of telemonitoring for patients who require this service.

Easier access to healthcare response service for patients through the platform.

Electronic Case Record (ECR) will provide an improved communication mechanism through on email box, and thus enhance the co-ordination of a patient’s care.

The platform will provide a directory of services for patients, family members and informal care givers, as well as professionals, to search for appropriate quality of the assured health and wellbeing services that are available.

Patients will be able to access the e-Prescription and choose their dispensing pharmacy.

2.2.1 Unstable Patients – out of hospital care

The above enhancement for the ‘stable’ patient will also be relevant for the ‘unstable’ patient. In addition, virtual consultations will be able to be activated, if necessary, among hospital specialists, nurses and GPs via email box when a patient’s health and wellbeing deteriorates.

2.2.2 Inpatient - hospital care

The hospital information system (HIS) should be integrated by the ECR; healthcare professionals will have access to the information (anonymised) in the Platform if a patient gets admitted. Selected doctors involved in CareWell will have access not only to the information in HIS, but also to LSV CareWell Platform. If a doctor is interested in the information uploaded by the patient, they will ask permission from the patient to look at this data. This should provide improved information on the patient’s medical history and the events leading up to hospital admission.

The educational platform in this phase of the project is not targeted at hospital doctors, but they will be able to access the information in the platform if they are interested in it.

2.2.3 Inpatient – hospital discharge preparation

The hospital will be able to refer the patient for telemonitoring if they are not already receiving the intervention according to the defined CareWell criteria, and determine their physiological parameters and frequency accordingly. In addition, patients will be signposted to appropriate patient empowerment services and educational content through the platform.

For patients who were receiving telemonitoring prior to their admission, it is expected that they will return to receive the telemonitoring service upon being discharge from hospital.

3 MATERIAL AND METHOD

The main Project action was preparation of the material for the Ethics Committee, which was supposed to agree on a research project. At the meeting of the Ethics Committee, in accordance with the applicable legislation, it was requested to prepare a proposal to the Bioethical Committee containing:

1. CV of the principal researcher
2. A detailed description of the project
3. A written acceptance of managers of the centers where the examinations are performed
4. Consent of a trustee of archival material to its use (each center)
5. An assessment card of assessment of a service beneficiary directed to a care / staying in a care unit (rating performed by service provider according to the Barthl scale)
6. The information model for the participants of the examination-doctor
7. The information model for the participants of the examination-patient
8. Specimen of the informed consent of participants or legal representatives to participation in the examination and data processing related to this participation (in the study).
9. Statement of the applicant about the knowledge of principles of medical confidentiality
10. Submit a policy of obligatory liability insurance of the entity engaged in medical activities.

Another element of the work in the project design was based on modelled telecare processes presented in Figure 2 and then selection of the suppliers of technical solutions, ie. Platform, which is shown in Figure 1. The integration and implementation work followed by training of medical staff - doctors and nurses, lasted until 15.06.2015r. Individual technical means for patients include a set of 50 mobile phones, smartphones L65 LG (LG-D280n) for each patient. And in addition:

A Diabetic set - glucometer- ProfiLine Blutzucker-Messsystem - 20 pcs.
B. COPD set - pulse oximeter PC-60NW - 5 pcs. and peakflow meter - Asma-1 Vitalograph company
C. Hipertension set - Blood pressure meter - SeniorLine BT model TD-3128- 15 pcs.
D. Set heart failure – pulse oximeter PC-60NW, weight scales - 10 pcs.

- Patients of Group A (diabetes) receive the glucometer.
- Patients of Group B (POHP) will receive peakflowmeter and pulse oximeter.
- Patients of Group C (hypertension) will receive blood pressure meter.
- Patients of Group D (heart failure) receive pulse oximeter and a weight scales.

Ultimately, the care will be provided for 100 people. The criteria for inclusion of patients into the follow-up observation are age 65-85; combination of not fewer than 2 types of diseases: hypertension (ICD I10), diabetes (ICD E11), chronic obstructive pulmonary disease (ICD J44), heart failure (ICD I50). Another required condition is to obtain at least 60 points according to Barthl scale.

Among the exclusion criteria involved in the project there were established: age below 65 and above 85 years, obtaining fewer than 60 points on a Barthl scale, previous myocardial infection or stroke, ischemic or haemorrhagic, in the last three months, an active process of cancer of any location, mental illnesses and unintentional loss of body weight: BMI <19, or weight loss ascertained by a doctor.

Qualifying took place on the basis of analysis of the information card of hospital stay. Patients qualified for observation were divided into two groups. The target size of both groups is 50 people. In Group I there were qualified persons covered by telemonitoring, who were provided with measuring devices depending on the disease entity. In Group II there were enrolled patients who were not covered by telemonitoring and who received no measuring devices.

Both in Group I and Group II there are patients with a similar disease profile, age group and degree of disability. In the next stage, persons qualified for the observation will be evaluated by a nurse (assessment of vital signs, including efficiency by Barthl scale) and a Primary Care physician during scheduled visits in the clinic. The estimated time of follow-up for individual patients is 18 months.

Until 06.30.2015 in the Project there were enrolled 75 people, including 39 women representing 52% of the respondents, and 36 men respectively, 48% of respondents. The average age being 73.96 years. The most commonly diagnose disease entity is hypertension (ICD I10) 73 persons (97.3% of the respondents). 45 patients (60%) were diagnosed with diabetes (ICD E11), and respectively in 15% (12 individuals) – with COPD chronic obstructive pulmonary disease, patients and 9 patients (12%) - chronic heart failure.

4 CONCLUSIONS

Key experiences and lessons that we gained at this stage of the project include:

1. Difficulty in understanding and reaching consensus on telecare model, which was then mapped in the implemented user interface
2. Integration of telemedicine devices to be available to patients with the platform (two providers of technical components Germany and Poland) to solve the problems of interpretation.
3. Overcoming the resistance in terms of a new type of telecare service - bath on the side of the organization (main beneficiary) and the method of patient enrolment for the project.
4. Difficulties at the level of patient care in a hospital - lack of willingness to understand the scale of the problem, potential benefits for the geriatric patient resulting from the "no standard" way of medical care
5. Difficulties in convincing patients to the unknown and so fare not processed of providing medical assistances to geriatric patients in Polish conditions
6. We observe and share the opinions of other researchers that it is necessary to make further to make the technology and service of measuring devices in patients home easier, more intuitive and requiring minimal action on the part of the patient.

REFERENCES

Alan Wailer, Tony Maitby.: *Actice again: A strategic policy solution to demographic ageing in the European Union*. International Journal of Social Wekfare. Volume 21, Issue Supplement s.1. pages s117-s130, October 2012.

http://stat.gov.pl/obszary-tematyczne/ludnosc/ludnosc/sytuacja-demograficzna-osob-starszych-i-konsekwencje-starzenia-sie-ludnosci-polski-w-swietle-prognozy-na-lata-2014-2050,18,1.html

Population Projection 2014-2050. Central Statistical Office, Warsaw 2014.

Guy Pare, Mirrou Jaana, Claude Scott., *Systematic Revew of Home Teleomitorin for Chronoc Diswases*: The Evidence Base. J Am Med. Inform Assoc. 2010 May-Jun : 14 (3) : 269-277

Guy Pare, Khali Mogadem, Gilles Pineau, Carole St-Hilare.: *Clinical Effects of Home Telemonitoring in the Context of Diabestes, Astma, Heart Failure and Hypertensinon: A Systematic Review. J* <ed International Res. 2010. Apr-Jun; 12(2)

Spyros Kitsiou, Guy Pare, Mirou Jann.: *Effect of Home Telemonitoring Intervention on Patient witch Cronic Heart Failure: An Overview of Systematic Reviews.,* Jurnal of Medical Internet Research, vol 17.no 3 (2015)

AUTHOR INDEX

AUTHOR INDEX

ICTRS 2015

Proceedings of ICTRS 2015
Fourth International Conference on Telecommunications and Remote Sensing
ISBN: 978-989-758-152-6
http://www.ictrs.org