Co-Chairmans: Dr. Saulius Maskeliūnas (Lithuanian Computer Society)
Prof. Gintautas Dzemyda (Vilnius University, Lithuanian Academy of Sciences)

Programme Committee:
Prof. Janis Grundspenkis (Latvia)
Prof. Hele-Maja Haav (Estonia)
Prof. Albertas Čaplinskas (Lithuania)
Prof. Yuriy Kharin (Belarus)

Organizing Committee:
Dr. Jolita Bernatavičienė
Dr. Olga Kurasova
Aidas Žandaris

Contacts:
Dr. Jolita Bernatavičienė
jolita.bernataviciene@mii.vu.lt
Dr. Olga Kurasova
olga.kurasova@mii.vu.lt
Tel. (8~5) 2109 312

Supported by:
NRD CS
www.nrdcs.lt

Ernst & Young Baltic
www.ey.com
Algoritmų sistemos
www.algoritmusistemos.lt

IO projects
www.ioprojects.lt
Computer Science Center
www.inscience.lt

Goals of the Event

1. Introduce computer science and IT business community with research undertaken at Lithuanian and foreign universities in the fields of software engineering and data mining (the number of scientists presenting their projects at the conference grows every year).

2. Foster relations between business entities and research community. Business representatives will have an opportunity to introduce their problems that need scientific solutions.

3. Facilitate scientific networking that will lead to joint research projects.

4. Initiate connections with foreign research institutions and scientists.

Topics
- Data Mining; Software Engineering
- Visualization Methods of Multidimensional Data
- Medical Informatics
- Ontological Engineering
- Business Rules

Publications
We will recommend the authors of talks to submit their papers for publication in the special issue of Baltic Journal of Modern Computing, www.bjmc.lu.lv.
Schedule of events

Thursday, December 4

11^{00} – 13^{00} – Registration/Info Desk Open

13^{00} – 13^{15} – Opening. Welcome and introductory remarks

**Gintautas Dzemyda** “Data Analysis Methods for Software Systems” (Vilnius University, Institute of Mathematics and Informatics)

13^{15} – 14^{15} – Plenary session

**Yurij Kharin, Zhurak M.** “Statistical Analysis of Spatio-temporal Discrete Data by Conditional Autoregressive Models” (Belarusian State University)

**Dmitry Podkopaev** “Research Challenges of Interactive Multiobjective Optimization” (System Research Institute, Polish Academy of Science)

14^{15} – 15^{00} – Coffee, Poster Session and discussions


2. **E. Radvilė, A. Ėnys, S. Ramanauskaitė** “Fitts Law Curve Screen Usability Testing with Children”


4. **P. Tamošiūnas** “Knowledge Acquisition Method for Virtual Learning Environment”
5. O. Kurasova, J. Tichonov “Classification of Large Images before Applying Compression Algorithm”

6. D. Eringis, G. Tamulevičius “Modified Filterbank Analysis Features for Speech Recognition”


8. M. Bilinskas, G. Dzemyda, M. Trakymas “An Approach to Tumor Location and Treatment Results Evaluation by CT Images”


15<sup>00</sup> – 16<sup>00</sup> – Plenary session

Elias Vyzas “Employing Analytics to Address Customer Needs” (Center of Excellence for Enterprise Intelligence, Central & Southeast Europe, Ernst&Young)


16<sup>00</sup> – 17<sup>00</sup> – Coffee, snacks, Poster Session and discussions

1. A. Mackutė-Varoneckienė, T. Krilavičius “Informative Feature Selection for Document Clustering”

2. Ž. Ledas, R. Baronas, R. Šimkus “Three-Dimensional Modeling and Visualization of the Bacterial Self-Organization in a Circular Container”

3. S. Gudas, A. Lopata “Towards Definition of Management Function in Information Systems Engineering”

5. **T. Liogenė, G. Tamulevičius** “Low-order Multi-level Features for Speech Emotions Recognition“


7. **V. Marcinkevičius, P. Treigys, J. Bernatavičienė, O. Kurasova, L. Chudzij, V. Medvedev** “Discover DAMIS – the Tool for Data Mining”

8. **V. Veikutis, A. Puodžiukynas, A. Sakalauskaitė, K. Stasiukynaitė** “IRT in Endocardial Transcatheter Laser Ablation Safety and Effectiveness Control”
Friday, December 5

9:00–10:00 – Plenary presentation

Janis Grundspenkis “Integration of Multi-Agent and Ontology Technologies for Intelligent Tutoring System Development” (Riga Technical University)


10:00–11:00 – Section session

Section 1 Chairman Dr. O. Kurasova

1. T. Krilavičius, R. Užupytė “Application of Clustering for Electricity Usage Patterns Detection”


3. E. Filatovas, O. Kurasova “Extensions of the Preference-based Multi-objective Evolutionary Algorithm R-NSGA-II”


Section 2 Chairman Dr. A. Lančinskas

1. V. Čyras, F. Lachmayer, K. Lapin “Visualising Legal Meaning in Legal Informatic”

2. V. Kaminskas, A. Vidugirienė, E. Vaškevičius “A Comparison of Identification Models for Human Emotions as Reactions to a Dynamical Virtual Face”

3. J. Gordevičius, K. Koncevičius, T. Grigalis, A. Smoliakovas “Identification of Epigenetic Markers for Early Diagnosis of Colorectal Cancer”

4. O. Vasilecas, D. Kalibatienė “Rule Based Dynamic Business Process Modelling and Simulation”
11^{00}–11^{20} – Coffee
11^{20}–12^{20} – Section session

Section 3 Chairman Dr. J. Gordevičius

1. **G. Tamulevičius**, A. Serackis, T. Sledevič, D. Navakauskas
   “Reference Template Update Technique for Isolated Word Recognition”

2. **O. Kurasova**, V. Marcinkevičius, **V. Medvedev**, A. Rapečka, P. Stefanovič
   “Data Mining in Big Data Era: Methods and Technologies”

3. **A. Lančinskas**, P. Fernandez, B. Pelegrin, J. Žilinskas
   “Solution of Discrete Competitive Facility Location Problem for an Entering Firm”

4. **T. Meškauskas**, F. Ivanauskas, V. Laurinavičius
   “Modeling of Biosensor, Taking into Account Substrate, Product and Enzyme Kinetics”

Section 4 Chairman Prof. V. Čyras

1. **M. Sangiovanni**, L. Maddalena, M. Guarracino
   “Following the Changes: HeLa Cells Lineage from Phase Contrast Microscopy Time-Lapse Data”

2. **D. Čalnerytė**, R. Barauskas
   “Modeling Physical Behavior of the Unidirectional Composite Materials with FEM Using Reduced Data”

3. **J. Miliauskaitė**, A. Čaplinskas, A. Lupeikienė
   “View Balancing Problem in the Modelling of QoS in SoES”

4. **S. Minkevičius**
   “An Application of Recurrent Method for the Analysis of Computer Network”
12\textsuperscript{20}–14\textsuperscript{00} – Lunch

14\textsuperscript{00}–15\textsuperscript{00} – Plenary presentation

\textbf{A. Jakaitienė, M. Avino, M. R. Guarracino} “Beta-Binomial Model for the Detection of Rare Mutations in Pooled Next-Generation Sequencing Experiments” (High Performance Computing and Networking Institute, National Research Council)


15\textsuperscript{00}–17\textsuperscript{00} – Coffee, Poster Session and discussions

1. \textbf{G. Stabingis, L. Stabingienė} “The Analysis of Influence of Noise Level to SCRD and Other Common Supervised Classification Methods”

2. \textbf{R. Alzbutas, V. Janilionis, M. Kavaliauskas, L. Bikulčienė} “Application of Data Mining Methods in Smart Systems Based on Wearable Sensors for Health Monitoring”

3. \textbf{A. Varoneckas, A. Žilinskas} “Visualization of Pareto Front Approximations in the Space of Decisions”

4. \textbf{M. Sabaliauskas, V. Marcinkevičius} “Comparison of Mods of Shoetrees Obtained by Theoretical and Experimental Methods”


6. \textbf{J. Zubova, O. Kurasova} “Challenges of Big Data Visualization”

8. **R. Baronas, J. Kulys** “Computational Modelling and Optimization of Biosensors: Status Quo”


10. **A. Gimbutas** “One-step Worst-case Optimal Bivariate Algorithm for Bi-objective Optimization”


15. **V. Veikutis, S. Valiukevičienė, V. Kučinskienė** “New Possibilities in Psoriatic Arthritis Diagnosis Using IRT”


19:00–21:00 – Dinner
Saturday, December 6

9^{30}–11^{30} – General discussion.

11^{30}–12^{00} – Coffee break, Poster Session and discussions

12^{15}–14^{00} – General discussion, concluding remarks

14^{00}–14^{30} – Closing
Application of Data Mining Methods in Smart Systems Based on Wearable Sensors for Health Monitoring

R. Alzbutas, V. Janilionis, M. Kavaliauskas, L. Bikulčienė

Kaunas University of Technology
Department of Applied Mathematics
liepa.bikulciene@ktu.lt

With the increase of healthcare services in non-clinical environments using vital signs provided by wearable sensors, the need to process and analyse the physiological measurements is growing significantly. In the presentation of recent KTU (Kaunas University or Technology) research the information about ITEA2 project CareWare - “Electronic Wearable Sport and Health Solutions” will be introduced. The aim of this project (which involves more than 20 partners) is to develop and leverage novel unobtrusive cyber physical systems for monitoring and advancing personal health and wellbeing. Lithuanian partners: KTU, Lithuanian Sport University, Audimas and Optitecha are involved in Health promotion use case and will develop functional state monitoring system, based on use of new generation sensors integrated for best wearing comfort and best signal quality during the motion (dynamic mode). Integrated solution of different sensors, smart interfaces, modelling and data analysis techniques should warrant that the created system will be comfortable and effective for assessment individuality and dynamics of functional state during daily-life activities as well as for exercise dosage control.

During this project KTU researchers will develop data mining methodologies and algorithms (anomaly detection, prediction, decision making, data integration and etc.) for physiological monitoring of vital signs in healthcare service. They will also be involved in analysis of stochastic processes and health characteristics, information from wearable sensors as well as analysis and integration of information in order to provide the algorithms for effective and robust decision making and creation of low-footprint low-power software modules for data processing.

Acknowledgment
This research was funded by a grant (No. 31V-88) from the Agency for Science, Innovation and Technology (MITA) regarding Eureka project "Electronic wearable sport and health solutions" (ITEA2 CareWare 13034).
Computational Modelling and Optimization of Biosensors: Status Quo

R. Baronas\(^1\), J. Kulys\(^2\)

\(^1\) Vilnius University
Faculty of Mathematics and Informatics
\(^2\) Vilnius Gediminas Technical University
Faculty of Fundamental Sciences
romas.baronas@mif.vu.lt

The computational modelling of biosensors permits to simulate the response of biosensors to a full concentration interval of analyte to be determined at steady state and transient conditions. The simulation covers biosensors with multiplex geometry and complex kinetics schemes of bio-catalysers action (Baronas et al., 2010).

For the efficient simulation of practical biosensors, multi scale numerical approaches and adaptive numerical algorithms have been developed (Britz et al., 2009). A mathematical optimization model has been developed and applied to a quantitative analysis of mixtures of compounds using the response of a biocatalytical amperometric biosensor (Baronas et al., 2013). The influence of the white noise and the temperature induced trend to the precision of the multianalyte concentration determination was investigated (Baronas et al., 2013).

For increasing the sensitivity, selectivity and stability of analytical systems, specialized computer simulators with powerful graphical user interfaces and tools conjugating practical biosensors with methods of chemometrics such as artificial neural networks are under development.

Acknowledgment
The research is funded by the European Social Fund under the Global Grant measure, Project No. VP1-3.1-ŠMM-07-K-01-073/MTDS-110000-583.

References
Visual Method of Streaming Data Similarity Estimation

J. Bernatavičienė, G. Dzemyda, G. Bazilevičius, V. Medvedev, V. Marcinkevičius, P. Treigys

Vilnius University
Institute of Mathematics and Informatics
jolita.bernataviciene@mii.vu.lt

Time series observations are frequently encountered in many domains such as business, economics, medicine, industry, engineering and science, etc. A time series is a sequential set of data points, measured typically over successive times. Recently there has been much interest in the problem of similarity search in time series. The proposed research deals with the physiological multivariate time series analysis seeking to compare the current situation with that in chronologically collected historical data and to find the subsequences of the multivariate time series most similar to the sample. A visual method for finding the best subsequences matching to the sample is proposed. Using the proposed method, an investigator can consider the results of comparison of the sample and some subsequence of the series from the standpoint of several measures that may be supplementary to one another or may be contradictory among themselves.

In the research, we suggest to integrate the specific visualization tools and methods for multidimensional data visualization that are effective and do not require intensive training of the users. Moreover, we show a possibility of making the decision, based on five criteria of similarity of the sample with the subsequence of real-time data stream by representing the similarity as a point on a plane. Five similarity measures were integrated in this method: Matrix Correlation Coefficient, Frobenius Norm, Extended Frobenius Norm, Multidimensional Dynamic Time Warping, Principal Component Analysis Similarity Factor. The method proposed in this research is universal and can be used in the analysis of streaming data of various natures.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.

15
An Approach to Tumor Location and Treatment Results Evaluation by CT Images

M. Bilinskas\textsuperscript{1}, G. Dzemyda\textsuperscript{1}, M. Trakymas\textsuperscript{2}

\textsuperscript{1}Vilnius University
Institute of Mathematics and Informatics
\textsuperscript{2} National Cancer Institute
mykolas.bilinskas@mii.vu.lt

One of the liver tumor treatment methods is radio-frequency ablation (RFA). The minimal ablative margin of RFA of a tumor needed to prevent local tumor progression is at least 5-10 mm. Tumor and ablative zone are evaluated using pre- and post-treatment Computed Tomography (CT), but a tumor is not visible in post-treatment CT image. Liver is a soft tissue, so a tumor can move relatively to other parts and it is not possible to measure a distance from tumor to ablated zone boundary – ablative margin. Therefore, special titanium markers are inserted around tumor – then pre- and post-treatment CT images can be registered and ablative margin measured.

Nowadays radiologist must somehow register pre- and post-treatment CT images and evaluate doubtful points. Usually the registration is done by hand, but it takes a lot of time and is inaccurate. This needs software to aid registering CT images and show doubtful points.

In this work, we present a semiautomatic method for tumor locating in CT images. It consists of two steps. In the first step, the liver is located, and in the second one, the tumor area is defined. For the tumor location, we suggest two items procedure: linearly normalize voxels intensity histogram and then locate a tumor or ablated zone. A Graph Cut framework is used for both liver and tumor location. It is adopted and implemented with directional graphs and a single 3D object restricted. The linear normalization is necessary because a region function of Graph Cut (needed for primary voxel classification) does not support such narrow domain – voxel intensities of tumoral and healthy tissues differ in less than 10 out of 256. All computations are performed in using the voxel 6-neighborhood adjacency system.
Composite materials are the materials with complicated internal structure composed of the two or more materials with the significantly different properties. The typical unidirectional composite is composed of the epoxy reinforced with the fibers which have higher stiffness. Due to the fact that composite materials are artificial and the desirable parameters can be selected by varying the material structure or the composing materials, the unidirectional composites are widely used to manufacture products for sports, aircraft, medical use and etc. However, if the internal structure of the composite is considered, the modeling of the physical behavior requires a lot of computational resources and time. The aim of the research is to maintain the same behavior with the reduced amount of the degrees of freedom (unknowns) in the system of the equations of the motion which results to less computations. In order to reduce the resources, the homogenization is applied. The representative element with all the properties of the structure is considered as homogeneous element with the linear material properties evaluated using pure strain assumptions in the implicit analysis. If the non-linear behavior or failure is considered, the properties are evaluated using the dynamical explicit analysis. In addition, the degrees of freedom can be reduced by varying the type of the element. The finite element program LS-DYNA is used for simulation.
Numerical Modelling of Heat Transfer in Underground Electrical Cables
R. Čiegis, V. Starikovičius, O. Suboč
Vilnius Gediminas Technical University
Faculty of Mathematics and Informatics
Department of Mathematical Modelling
rc@vgtu.lt

The aim of this project is to develop a virtual modelling tool which can be used to construct optimal design of power transmission lines and cables. They should meet the latest power transmission network technical and economical requirements. The two dimensional mathematical model is based on a general heat conduction equation.

The velocities are obtained by solving a coupled thermoconvection problem including the heat conduction problem and a standard Navier-Stokes model of the flow in air. The changes of material coefficients in soil due to influence of heating are taken by solving a simplified Richard's model of flows in porous media. The finite volume method is used to solve the obtained system of differential equations. Discretization of the domain is done by applying a Cute mesh generation tool, which is modification of the well-known Triangle mesh generator.

The discrete schemes are implemented by using the OpenFOAM tool. Parallel versions of basic algorithms are also investigated.

Visualizing Legal Meaning in Legal Informatics
V. Čyras, F. Lachmayer, K. Lapin
Vilnius University
Faculty of Mathematics and Informatics
vytautas.cyras@mif.vu.lt

This presentation is about visualization in the domain of legal informatics. We propose an approach to visualization which is called sequential legal visualization (SLV). It is about a sequence of images which facilitates the comprehension of the meaning of legal contents. Visualising legal meaning differs from data visualization and information visualization and is relevant to knowledge visualization. A sequence of images in SLV can be compared with a narrative. There is a link to focus+context techniques in information visualization.
Pictures in law bring a risk of drawbacks, such as redundancy, a low level of abstraction, trivialization, and emotions. However, the use of logical pictures can bring advantages because metaphors and symbols can be employed.

Visualization is important as tertium comparationis. This is important more generally – for multilingual scientific communities – and not only in legal informatics. Tertium comparationis (Latin – the third [part] of the comparison) is the quality that two things that are being compared have in common. Different languages can be used in a multilingual scientific discourse. On the one hand, English, a global language, can be used. Let us call this top-down approach. On the other hand, regional languages, such as German of French, can also be used. Let us call this usage bottom-up approach. The language of a big region (not of a province) allows a scientist to unfold his ideas more naturally. However this brings translation problems. Thus dictionaries and translation tools emerge, e.g. Inter-Active Terminology for Europe (IATE), the inter-institutional terminology database of the European Union (http://iate.europa.eu). Visualization can supplement translation. Examples are books for visualized learning, where illustrations complement translation. Thus visualization brings an additional syntactic dimension to languages.

Next we introduce another concept – tertium communicationis – to denote the third part of the communication between an agent who speaks language A to an agent who speaks B. Besides visualization other intermediate formats such as XML schema can be employed. We aim at a tertium communicationis as a document type definition that improves communication between human beings or machines. The question “What are the formats which contribute to better communication?” depends on various factors, such as document type and the task of communication. Therefore we speak about transitions text-visualization-model-metamodel.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.
The main issue discussed in this presentation is the development of a versatile information system framework for inventory purpose. The Lithuanian National Forest Inventory (NFI) is presented as case of information system framework development. Statistical sampling method is used for NFI, regular (every 5 years) inventory of forest land using permanent and temporary plots is proceeded. Over 16000 sampling plots and over 180000 trees are inventoried and derivative parameters are calculated with high 1.0% accuracy. The main objectives for information system have been defined: the data aggregation and filtering in information system must be configurable, information system logic and data structure must be easy changeable. The information system framework based on business rules and ontology conception has been developed. The proposed information system framework enhancing the comfort and work productivity when implementing new or existing inventory and it ensures efficient information system maintenance in a dynamic and consistently fluctuating business environment.
Modified Filterbank Analysis Features for Speech Recognition
D. Eringis, G. Tamulevičius
Vilnius University
Institute of Mathematics and Informatics
deividas.eringis@mii.vu.lt

Commonly-used speech signal analysis techniques for automatic speech recognition are based on filterbank analysis. The set of band-pass filters generalizes and simulates the speech signal processing in human auditory system and thus enhances the automatic speech recognition (ASR). For instance, filterbank based mel-scale cepstral (MFCC) analysis is used in almost all worldwide known and exploited ASR systems.

Results of various experimental studies show that mel-scale arrangement of filters is not the most effective one and the modification of the filterbank can provide additional accuracy and noise robustness to the ASR system. Therefore we modified the initial mel-scale filterbank and applied it to Lithuanian speech recognition task. We varied different filterbank parameters (central frequencies, filter bandwidth values, etc.) in search of the most efficient filterbank configuration. Results show that we should talk about adaptive filterbank analysis which deals with the individual characteristics of the speaker thus enhancing speech recognition accuracy.

Extensions of the Preference-based Multi-Objective Evolutionary Algorithm R-NSGA-II
E. Filatovas, O. Kurasova
Vilnius University
Institute of Mathematics and Informatics
ernest.filatov@gmail.com

Evolutionary multi-objective optimization algorithms have been applied in various real-world applications. However, most of the Pareto domination based multi-objective optimization evolutionary algorithms are computationally expensive for many-objective problems. Incorporation of the decision maker’s preferences into multi-objective evolutionary strategies have become the relevant trend during the last decade, and several preference-based evolutionary algorithms have been proposed in the literature. In our research we focus on improvement of one well-known preference-based evolutionary algorithm R-NSGA-II.
Several enhancements of the algorithm has been developed and evaluated by solving a set of well-known multi-objective optimization test instances of different scope. The experimental investigation has shown that proposed improvements has positive impact to the quality of the algorithm in the sense of speed a convergence to the Pareto front.

Acknowledgement
Postdoctoral fellowship of E. Filatovas is being funded by European Union Structural Funds project Postdoctoral Fellowship Implementation in Lithuania.

One-step Worst-case Optimal Bivariate Algorithm for Bi-objective Optimization
A. Gimbutas
Vilnius University
Institute of Mathematics and Informatics
albertas.gimbutas@mii.vu.lt

Bivariate one-step worst-case optimal algorithm for bi-objective Lipschitz optimization problems is presented. Lipschitz constants with respect to Euclidean metric are used. The feasible region is partitioned by simplexes. The bivariate algorithm is used to compute discreate Pareto front representation.

Identification of Epigenetic Markers for Early Diagnosis of Colorectal Cancer
J. Gordevičius, K. Koncevičius, T. Grigalis, A. Smoliakovas
Vilnius University
Institute of Mathematics and Informatics
juozas.gordevicius@mii.vu.lt

Colorectal cancer (CRC) is one of the most prevalent forms of malignancy worldwide, which every year claims about half a million lives. Detection and treatment at a pre-metastatic stage is the most effective way to reduce mortality from CRC. Five-year survival for CRC patients diagnosed when their tumours are localized exceeds 90%, but drops below 12% when the cancer has metastasized. Use of a simple and effective biomarker for early colon cancer will prevent more deaths than any pharmacologic approach to advanced disease. There is a critical
need for highly specific and sensitive non-invasive biomarkers for CRC diagnosis that could be used routinely in clinical laboratories.

One promising line of research in early diagnostics of cancer is the epigenomics of blood circulating DNA. During the growth of a tumour some malignant cells die, and the contents of such cells are released into the blood stream. It has been known for some time that cancer cells exhibit numerous genetic and epigenetic changes. These latter are chemical modifications to the structure of genes that normally regulate the process of turning them on or off. The observation that DNA carrying aberrant epigenetic signatures appears in the blood stream of cancer patients can be used to develop epigenetic markers for early detection of cancer.

In this talk I will present ongoing data analysis of epigenetic profiles in blood circulating DNA to detect epigenetic biomarkers for early CRC diagnosis. Our goal is to apply novel algorithmic methods that will reveal epigenetic aberrations in extended regions when comparing 200 cancer patients to 200 matched controls. The detected regions will be used to train machine learning algorithms and to build and evaluate an early diagnostic CRC classifier. The bioinformatics analysis conducted in this project may provide novel, non-invasive markers, with the potential for translation into clinical practice.

Integration of Multi-Agent and Ontology Technologies for Intelligent Tutoring System Development

J. Grundspenkis
Department of Artificial Intelligence and Systems Engineering
Institute of Applied Computer Systems
Faculty of Computer Science and Computer Systems
Riga Technical University
Latvia
janis.grundspenkis@rtu.lv

The internet has become a major source of information in many knowledge domains including education domain. The issues associated with the current use and management of information on the Web as well as two major technologies for solutions to these problems, namely, agent and ontology technologies are discussed. As quantity of available information increases every day, it is becoming more and more difficult for teachers and learners not only to find specific information, but also to choose highly relevant information in case if many
sources are available. Designing software agents which are able to search and select information in compliance with specific needs of each user is a solution of this problem. Intelligent agents which form a multi-agent system must be able to locate meaningful information and understand the meaning of it, to advertise their capabilities and to know the capabilities of other agents. All this can be achieved through ontology sharing between agents and through the annotation of information resources.

The design of agent-based systems requires a methodology to guide designers through the design process. Regardless of the fact that there are many different multi-agent system design methodologies, none is suitable for design of intelligent tutoring systems (ITSs) based on agent paradigm. For that reason the multi-agent systems based ITS development methodology MASITS and the corresponding tool have been developed at the Department of AI&SE. The MASITS supports the development of ITS MIPITS which offers learning materials and tasks adapting them to the current level of each learner’s knowledge and his/her preferences. The architecture and the tutoring scenario of MIPITS are presented in details.

The concept map (CM) based intelligent knowledge assessment system IKAS has also been developed. The core of IKAS is the intelligent knowledge assessment agent (IKAA) which essentially is a multi-agent system consisting of the communication agent, agent-expert, knowledge evaluation agent, and interaction registering agent. The IKAA assesses each learner’s knowledge by comparing his/her constructed CM with the teacher’s CM. The IKAS has high capacity for adaptation to each learner’s current knowledge level having possibility to change the degree of task difficulty and to choose the form of feedback. To make the CM construction task for teachers easier, the teacher’s module includes the repository of ontologies. The role of ontologies is twofold. First, the specific ontology may be transformed into teacher’s CM or, if the ontology is not available, the teacher’s CM may be transformed into the ontology. Second, in the IKAS the CM merging based on methods and tools developed for the same task in the domain of ontologies is used. Thus the functionality of the IKAS is extended allowing the reuse of captured CMs. The extension of the functionality is also carried out by implementation of three algorithms which determine the content of learning objects for correction of learner’s knowledge (based on analysis of his/her errors) and generation of personalized learning path. In result learners receive more support from the IKAS and the adaptability of it is increased.
Towards Definition of Management Function in Information Systems Engineering

S. Gudas\textsuperscript{1}, A. Lopata\textsuperscript{2}
\textsuperscript{1}Vilnius University
Institute of Mathematics and Informatics
\textsuperscript{2}Vilnius University
Kaunas Faculty of Humanities
saulius.gudas@khf.vu.lt
audrius.lopata@khf.vu.lt

Model-driven IS engineering methods invoke development of relevant IS application domain modeling methods. IS application domain and business domain are not isolated and could be investigated using the same paradigm of modeling – internal modeling approach. The understanding of IS application domain as a self-managed system requires to redefine such concepts as management function and management transaction in the context of information systems engineering. The definitions and substantiation of these and related concepts from internal viewpoint of modeling are discussed and illustrated.

Beta-Binomial Model for the Detection of Rare Mutations in Pooled Next-Generation Sequencing Experiments

A. Jakaitienė, M. Avino, M. R. Guarracino
High Performance Computing and Networking Institute
National Research Council, Italy
audrone.jakaitiene@na.icar.cnr.it

With next-generation DNA sequencing technologies, one can interrogate a specific genomic region of interest at very high depth of coverage and identify less prevalent rare mutations in heterogeneous clinical samples. However, while this technology has the capacity to generate billions of nucleotides of DNA sequences in a single experiment, the error rate of around 1\% yields in millions of sequencing mistakes making detection of rare mutation computationally very challenging.

Currently, sequencing costs are on downward trend, although they still remain expensive. The use of untagged pooled next-generation sequencing (NGS) data might reduce these costs and theoretically might be more effective in mutation discovery, even though one would need a model to
predict whether the single nucleotide polymorphisms (SNPs) are sufficiently covered and still detectable in the experiment. So far, only in few publications the experiments with pooled data were analysed. At present, available standard software for the detection of rare mutations are less accurate when pooled NGS experiments are considered.

Flaherty et al. 2012 proposed the method for the detection of rare single nucleotide mutations at the 0.1% level for synthetic DNA samples. Following the latter, we adapt hierarchical Beta-Binomial model for the detection of rare mutations in pooled experiments on 98 genes of 128 patients with a clinical diagnosis of muscular disease. Our aim is to show that the proposed statistical approach is outperforming currently available variant calling software (such as GATK, SNVer, FreeBayes).

Reference

Research on Identification of Defects' Contours of Road Surface in Raster Images
V. Jakštys, V. Marcinkevičius, P. Treigys
Vilnius University
Institute of Mathematics and Informatics
vitaakas@gmail.com

This article analyses issues related to detection and evaluation of road surface defects that were formatted by smart device raster image as well as assessment of measurements of particular objects and related bias. These certain problems arise in practice when solving real life issues such as detecting pits, calculating distance to vehicle ahead, improving navigation systems, recognizing signs and road marking as well as eliminating irrelevant objects. This article aims to asses defects of road surface, in particular, road pits. Most common methods of object’s parameterization that are mentioned in academic literature include employing cameras and various kinds of sensors such as sound, visual, infrared or ultrasound ones. Images fixed by one came a are used in this article in order to identify contours of road pits and parameterize the dimensions. This method allows recognizing object in terms of their color, shape and structure. The methods assessed in the article are applied to identify and outline contours of road pits, and their accuracy is evaluated through empirical research.
Some Applications of Numerical Models of Temperature Distribution in Electrical Power Cables and it's Surrounding Area

G. Jankevičiūtė, N. Tumanova, M. Meilūnas, A. Bugajev
Vilnius Gediminas Technical University
Faculty of Fundamental Sciences
Department of Mathematical Modelling

This research is aimed to develop design rules for power transmission lines and cables, which have to meet the latest power transmission network technical and economical requirements. At present the power lines are over-dimensioned by up to 60% in terms of transmitted power. However, today, as the new distributed generating capacities are installed e.g. large wind farms, bio-gas plants or waist-to-energy plants, the infrastructure of power grid must be re-designed or new optimization strategies for the available grid developed. Power cables for power distribution applications are still rated according to IEC 287 and IEC 853 standards. However, today there are many applications where analytical and heuristic formulas cannot describe precisely enough the conditions under which the cables are installed. An example could be an underground cable route, where the installation conditions for a cable are different only for a short distance (crossing the road). The present standards require that the cable's current-carrying capacity must be reduced only due the fact that the cables route crosses the road. To be on the safe side this rule is acceptable, but today the cost effective designing of cable installations comes first as the copper price level has reached its maximum since decades. The knowledge of dynamics (in time) of heat distribution in/around electrical cables is necessary to optimize the usage of electricity transferring infrastructure. The mathematical modelling can reduce cable usage restrictions, by taking into account different circumstances precisely. We consider some heat transfer models in a cross-section of power cable and its surrounding area, when cables are be directly buried or laid out in buried ducts, and present numerical results of steady state and transient simulation under various geometrical and physical conditions.
A Comparison of Identification Models for Human Emotions as Reactions to a Dynamical Virtual Face

V. Kaminskas, A. Vidugirienė, E. Vaškevičius

Vytautas Magnus University
Department of Systems' Analysis
v.kaminskas@if.vdu.lt

The report presents the research of dependencies between human emotional signals (excitement and frustration) and virtual dynamical stimulus (virtual 3D face with changing distance between eyes). The dependencies are described by linear “input-output” type models.

Emotional signals are received using EEG measuring device Emotiv Epoc with 14 channels. Recorded signals are pre-processed automatically using the device and specialized software and relative values of emotional signals are used for modelling. Two model identification methods are proposed – modified correlation-based and the least squares method with projection to a stability domain what ensures building of stable models. Identification was performed using two types of stimuli for nine volunteers. Validation results of the models showed, that they predict emotional signals with a relatively high accuracy – absolute relative error of excitement prediction does not exceed 9% and absolute relative error of frustration prediction does not exceed 3%.

Techniques for Estimating the Intrinsic Dimensionality of High-Dimensional Data

R. Karbauskaitė, G. Dzemyda

Vilnius University
Institute of Mathematics and Informatics
rasa.karbauskaite@mii.vu.lt

In the exploratory data analysis, we often confront with real-life data that are of a very high-dimensionality. However, these data are usually not truly high-dimensional, i.e. they are only embedded in a high-dimensional space, but can be efficiently summarized in a space of much lower dimensionality, such as a nonlinear manifold. Therefore, the problem is to find out the exact dimensionality of that manifold, i.e. the intrinsic dimensionality of the analysed data. The intrinsic dimensionality of a data set is usually defined as the minimal number of parameters or latent variables necessary to describe the
data. When the intrinsic dimensionality of a data set is known, it is possible to reduce the dimensionality of data without losing much information.

The intrinsic dimension may be considered as either a topological dimension or a fractal dimension. While a topological dimension always yields an integer value, a fractal dimension does not have to be an integer and it is often a real number. The topological dimension is the basis dimension of the local linear approximation of the hypersurface on which the data resides, i.e. the tangent space. A disadvantage of the topological dimension is that it is very difficult to estimate the topological dimension technically, if only a finite set of points is available. Hence, practical methods use various other notions of the intrinsic dimension. The most usual ones are related with the fractal dimension. Fractal dimension is a parameter that characterizes how densely a fractal fills the space. However, the fractal dimension has also a number of different interpretations: the capacity dimension, the correlation dimension, etc. In this research, we review the techniques to estimate the intrinsic dimensionality of a data set. Two classifications of estimators are discussed: a) local/global estimators, b) projection techniques/geometric approaches.

Statistical Analysis of Spatio-temporal Discrete Data by Conditional Autoregressive Models

Y. Kharin, M. Zhurak

Research Institute for Applied Problems of Mathematics and Informatics
Belarusian State University
kharin@bsu.by

Studying the probabilistic models of spatio-temporal data, that arises during the monitoring different processes in a specific geographical area, is a new topical scientific direction. Statistical analysis of spatio-temporal data is an important problem for practice: it allows to model adequately the underlying stochastic phenomenon taking into account both the dependence on time and the dependence on space. Models based on spatio-temporal data become widely used for solving practical problems in meteorology, ecology, economics, medicine and other fields. The Poisson and binominal conditional autoregressive models of spatio-temporal data are developed. These models can be used to describe the discrete spatio-temporal data. It is proved that under the Poisson and binominal conditional autoregressive models the observed process is the nonhomogeneous vector Markov chain with countable and finite space of states respectively. Probabilistic properties of these models are studied: the formulas for calculation of the
one-step transition probability matrix, the current probability distribution, expectation and variance for each site (spatial regions or space locations into which the analyzed spatial area is partitioned). To estimate parameters of the models we use maximum likelihood method. We apply numerical Newton procedure to maximize the log-likelihood function. Asymptotic properties of estimators are studied. The optimal forecasting statistics that minimize the probability of the forecasting error and the mean square risk are built. According to probabilistic properties of the Poisson conditional autoregressive model we propose an iterative algorithm for calculating the forecast values. Risks of forecasts are also calculated. "Plug-in" forecasting statistics are built in the case of unknown parameters. The computer experiments were carried out on simulated and real medical data. Real data describes the incidence rate of children leukemia.

Evaluation Framework for Software Security Requirements Engineering Tools

W. A. Khilji, R. Matulevičius
Tartu University
rma@ut.ee

Security requirements play an important role in secure software development. Understanding them from the early stages, could help to envisage security threats and their consequences and to reason about the necessary countermeasures. In order to support security requirements engineering, different approaches and software tools are suggested. However these tools are still at their infancy and are lacking mass adoption in practice, because they have steep learning curve, which potentially adds-up to development time, and end up pushing more cost to the project. In order to decide which tool to select, engineers need to test them that, in return, will increase the project cost again. Moreover, using unstructured tool selection process could lead to wrong tool selection.

Based on the analysis of nowadays security engineering processes, methods, and tools, in this work we propose a security engineering tool evaluation framework (SETEF). This contributes to a structured process, which potentially helps engineers to select the security engineering tool according to the project needs. We analyse validity of our proposal by comparing it to the approach for the general requirements engineering tool selection.
Hierarchical Fuzzy Rule Models for Project Risk Evaluation
D. Krikščiūnienė, V. Sakalauskas, S. Strigūnaitė
Vilnius University
Kaunas Faculty of Humanities
dalia.kriksciuniene@khf.vu.lt

The article analyses methods for creating fuzzy rules by applying hierarchical models. The method is suggested for evaluating project performance and risk factors, based on fuzzy hierarchical principles. The results of application of the proposed model are experimentally explored in virtual IT project settings.

Application of Clustering for Electricity Usage Patterns Detection
T. Krilavičius, R. Užupytė
Baltic Institute of Advanced Technology
and Vytautas Magnus University
t.krilavicius@if.vdu.lt

Electricity usage patterns are important for suppliers in order to ensure efficient electricity distribution and pricing, and for users to save costs. Smart meters provide high granularity electricity usage data (e.g., hourly), which can be used to extract patterns of daily/weekly/monthly/yearly electricity usage. In this research we propose a method for grouping electricity users based on hourly electricity usage data, based on the k-means clustering and periodicity analysis, with semi-automatic parameters selection based on adequacy measure. We illustrate the method with 1500 electricity users 3 years’ data, discuss pros and cons of the methods, and future plans.

Acknowledgements
We thank Binar Solutions for cooperation and European Social Fund and the Republic of Lithuania (grant number VP1-3.2-ŠMM-01-K-02-002) for funding.
Performance Enhancement in Biological Structures
Visual Examination by Multivariate Analysis Based
Optimized Multimodal Imaging: Visualization of
Malfunctioning Structures in Eye Fundus Imaging
Lithuanian University of Health Sciences
algimantas.krisciukaitis@lsmuni.lt

Visual evaluation of morphological structures in biomedical images is widely used in medical diagnostics. Together with interactive or sometimes even automatic estimation of morphologic parameters it produces features used to diagnose diseases or monitor status of the patient. Significant improvement of the diagnostic performance is reached by merging multimodal images representing the same structures to create new images with enhanced contrast of target structures. Multivariate analysis methods could be successfully used for elaboration of such optimized image transforms. We illustrate the idea by example how Principal Component Analysis is used to increase the contrast of specific structures in eye fundus images used for Age Related Macular Degeneration patients follow up.

Objects of interest - drusen (focal deposits of extracellular debris located between the basal lamina of the retinal pigment epithelium and the inner collagenous layer of Bruch membrane) usually are represented by various intensity but the same unique color in the image. Construction of the optimal color scheme to increase the contrast of drusen is realized by means of Principal Component Analysis, which transforms original RGB color representation into principal components space. The study demonstrates that proposed method can increase contrast-to-noise ratio of the drusen areas 10-fold or more.

The principle of creation of optimized diagnostic features can be extended including images obtained by using of various sources of excitation light or by means of other modalities or imaging technics.
Nowadays, we often deal with big data, new technologies are able to store and process a larger and larger amount of data. Big data can be collected not only from computers, but also from billions of mobile phones, social media posts, different sensors and many other sources. Big data bring new challenges to data mining therefore large volumes and different varieties must be taken into account. The common methods and tools for data processing and analysis are unable to manage such amounts of data, even if powerful computer clusters are used. To analyse big data, new data mining and machine learning algorithms as well as technologies have been developed. Big data do not only yield new data types and storage mechanisms, but also new methods of analysis. Various data mining techniques are implemented for clustering, classification, association rule learning and other data mining problems. If the data sets to be analysed are small enough, the well-known data mining systems can be used, eg. Weka, Knime, RapidMiner, Orange. Grids and computer clusters can be used to solve large data mining problems. New Grid-based systems are implemented, eg. Taverna, Dame, Knime Cluster Execution, etc. When we are dealing with big data, we are dealing with known big data challenges such as data velocity, data volume and data variety. Enterprise data warehouses and Hadoop technologies can be helpful to manage these challenges. The Internet is full of information on big data, however, there is a lack of systematized information about which methods and techniques can be used for a big data analysis. The aim of the research is to propose some recommendations for selection of big data mining methods and technologies according to volume of the data.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.

Data Mining in Big Data Era: Methods and Technologies
O. Kurasova, V. Marcinkevičius, V. Medvedev, A. Rapečka, P. Stefanovič
Vilnius University
Institute of Mathematics and Informatics
olga.kurasova@mii.vu.lt
Classification of Large Images before Applying Compression Algorithm

O. Kurasova, J. Tichonov

Vilnius University
Institute of Mathematics and Informatics
jevgenij.tichonov@mii.vu.lt

The article proposes a model for managing large-scale digital image groups by using remotely operations and storing. This model provides image management using web services, which divide images into different classes, and for each class the best image compression algorithm is applied. The article also provides an overview of the possible solutions for the implementation of the individual model parts and describes an experimental investigation, which shows that classification of large images before applying compression algorithms can be useful.

Solution of Discrete Competitive Facility Location Problem for an Entering Firm

A. Lančinskas¹, P. Fernandez², B. Pelegrin², J. Žilinskas¹

¹Vilnius University
Institute of Mathematics and Informatics
²University of Murcia, Spain
algirdas.lancinskas@mii.vu.lt

In general the location of facilities deals with the determination of the optimal locations for facilities and is important for the firms providing services or goods to customers in a certain geographical area.

There are a lot of models of facility location proposed in the literature, which differ on their properties such as location space, describing possible locations for the facilities being located, attractiveness of the facilities, or behavior of customers when choosing the most attractive facility to get a service. However, the determination of optimal locations for the new facilities usually leads to solution of a complex global optimization problem with various properties and constraints.

In particular we deal with the solution of the discrete competitive facility location problem for an entering firm, where a new firm wants to enter the market by establishing a set of facilities. The entering firm is
interested in maximization of the market share of the new facilities by selecting their locations from a given set of candidate locations thus encountering a discrete global optimization problem. We developed a heuristic population-based algorithm for discrete global optimization which is specially adopted to solve the competitive facility location problem for firm expansion. The algorithm has been experimentally investigated by solving different instances of the facility location problem of different scope reasonable for real-world facility location problems. The real data (the geographical coordinates and the population) of several thousands of demand points and different models of behavior of customers when choosing the most attractive facility have been used to investigate performance of the developed algorithm.

Acknowledgment
This research was funded by a grant (No. MIP-051/2014) from the Research Council of Lithuania.

Three-Dimensional Modeling and Visualization of the Bacterial Self-Organization in a Circular Container
Ž. Ledas, R. Baronas, R. Šimkus
Vilnius University
Faculty of Mathematics and Informatics
zilvinas.ledas@dict.lt

Microorganisms and bacteria move toward food sources and away from various dangerous chemical gradients. Such directed movement is called chemotaxis and it plays an important role in a wide range of biological processes.

Recently, pattern formation of a luminous Escherichia coli colony in a circular tube-like container was modeled by Keller-Segel type reaction-diffusion-chemotaxis equations and various model modifications were studied to determine the suitable minimal model in quasi one dimension. Then it was extended for two dimensions by adding the third – oxygen – equation. In this work, we investigate the suitability of the proposed minimal model by modeling it in three dimensions.

The 3D output of the numerical simulation was visualized using video game engine to enable convenient real time investigation of pattern evolution during time as well as space dimensions. The
mathematical model and the numerical solution were validated by experimental data. The three-dimensional computational model simulates the pattern formation more precisely than the corresponding one-dimensional model.

Acknowledgment
This work was supported by the Research Council of Lithuania (LMT grant no. MIP-001/2014).

Low-order Multi-level Features for Speech Emotions Recognition
T. Liogienė, G. Tamulevičius
Vilnius University
Institute of Mathematics and Informatics
tatjana.liogiene@mii.vu.lt

Vast majority of speech emotion recognition researches tend to explore huge feature sets up to a few thousand different features. This causes the “curse of dimensionality” problem which can be solved by increasing speech data amount or reducing selected feature sets. Various techniques are used for reduction of feature sets – PCA analysis, multidimensional scaling analysis, and various feature selections schemes like sequential forward selection, sequential floating forward selection, and sequential backward selection. Some original techniques like hierarchical features, multi-stage classification, and separated analysis of male and female speech records are proposed. The aim of our research is the use of low-order multi-level features for speech emotion recognition. We use two-stage classification scheme of 4 speech emotions – anger, sadness, joy and neutral state. During the first stage low pitch and high pitch speech emotions are separated using fundamental frequency based features. The second stage is intended for identification of speech emotion using particular features. This organization of recognition process enables us to apply different features for particular emotions and to improve recognition accuracy of the selected emotion without affecting the other emotion recognition process. The proposed classification technique is experimentally compared to “classical” speech emotion organization using Berlin database of emotional speech.
Informative Feature Selection for Document Clustering

A. Mackutė-Varoneckienė\textsuperscript{1}, T. Krilavičius \textsuperscript{1,2}

\textsuperscript{1}Vytautas Magnus University, Kaunas, Lithuania  
\textsuperscript{2}Baltic Institute of Advanced Technology, Vilnius, Lithuania  
ausra.mackute@gmail.com

Document clustering incorporates a number of data mining techniques, and to achieve good clustering results, all of them should be well attuned. Full text document clustering problem is mostly challenging due to the data dimensionality. The number of terms in corpora may constitute hundreds of thousands and a particular document may contain only hundreds of terms. Usually, documents – terms matrices are very sparse due to high diversity of terms in different documents. One of the important tasks is to find the feature subset which includes the most informative features. Efficient feature selection helps to cope with the data dimensionality. We investigate a number of thresholding feature selection methods based on the document frequency (DF), term frequency based document frequency (TFDF), term frequency-inverse document frequency (TFIDF), term strength (TS) and term contribution (TC), and their performance for Lithuanian, Russian and Azeri (Azerbaijani) languages. F-score, purity and entropy are used to evaluate performance of different feature selection methods. Results show that feature selection based on TFDF performs better than other feature selection methods for all languages and it is enough to select up to 7\% informative features of all feature set to obtain best clustering results. In future we will investigate effectiveness of Principal Component Analysis (PCA) technique for feature selection and possible combinations of feature selection methods commonly used in document clustering with other feature reduction techniques.

Acknowledgement
This research was funded by a grant (No. VP1-3.1-ŠMM-10-V-02-025) from the ESFA.
Discover DAMIS – the Tool for Data Mining

V. Marcinkevičius, P. Treigys, J. Bernatavičienė, O. Kurasova, L. Chudzij, V. Medvedev

Informatics Science Center
virginijus.marcinkevicius@inscience.lt

Recently, web-based applications as tools for data mining have been rapidly developed. They replace intensely the previously popular desktop applications such as SPSS, Weka, Knime, RapidMiner. The most important advantage of web-based application over desktop is that only a web browser is needed. This brings to the users out of the box functionality that the users use in everyday tasks. A newly developed web-based data mining system DAMIS (http://damis.lt) implements user-friendly interface and allows to carry out the data analysis such as clustering, classification, and dimensionality reduction. It provides means for multidimensional data investigation: data visualization, data projection on a plane, and estimation of the similarities between the data items. DAMIS can explore the influence of individual features and their relationships by visual analysis techniques. It implements the paradigm of scientific workflows. Researchers or other users can specify the order in which the data pre-processing and/or mining algorithms should be executed. There is a possibility to change and save the workflows composed, as well as to reuse workflows for a newly introduced data. The data analysis results can be saved to a user’s computer or other cloud storage such as national open access scientific information archive (MIDAS). Intense calculus and big data analysis are achieved by using the high performance computing resources: computer clusters or supercomputers.
Mathematical model for simulation of multilayer electrochemical biosensor has been elaborated. The model, describing kinetics of substrate, product and enzyme, consists of nonlinear system of partial differential equations (of diffusion-reaction type) with initial values and boundary conditions. Two cases: when product diffuses out of the biosensor and when the outer membrane is impermeable for product (product is trapped inside the biosensor) have been dealt with by adjusting boundary conditions in the mathematical model. An algorithm (based on Crank–Nicolson finite difference scheme) for computing the numerical solution of the mathematical model has been applied. The algorithm has been implemented in C programming language, employing Message Passing Interface (MPI) communications protocol. Beside other computing platforms, the software has been deployed and tested on supercomputer at Vilnius University, Faculty of Mathematics and Informatics, Information Technology Open Access Centre. Numerical results have been presented. Computed curves of current density at biosensor electrode have demonstrated analogous profiles to those observed in real experiments. The software can be employed for simulation of the biosensor response and evaluation of other biosensor characteristics.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.
View Balancing Problem in the Modelling of QoS in SoES

J. Miliauskaitė, A. Čaplinskas, A. Lupeikienė

Vilnius University
Institute of Mathematics and Informatics
jolanta.miliauskaite@mii.vu.lt

The quality of service (QoS) modelling in a service-oriented enterprise system (SoES) is exceedingly complex because of the various understandings of stakeholders about what QoS is and how it should be defined and specified. The development of new services raises a number of QoS related problems. A methodology and its supporting tool are required to capture, balance and integrate the different views on QoS. The main aim is to predict, at least approximately, the required investments (cost, time, etc.). In the presentation, we will discuss the problems which should be solved in order to develop such the methodology and the current project state.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.

An Application of Recurrent Method for the Analysis of Computer Network

S. Minkevičius

Vilnius University
Institute of Mathematics and Informatics
minkevičius.saulius@gmail.com

In this paper we present an example of application of recurrent method for the analysis of computer network. An open queueing network model in light traffic has been developed. The probability limit theorem for the idle time process of customers in heavy traffic in open queueing networks has been presented. Finally, we present an application of the theorem - an idle time model from computer network practice.

Acknowledgement
This research was funded by project: VP1-3.1-ŠMM-08-K-01-010 “Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms”.
Treatment (diagnosing and tracking of therapy process) for epilepsy and some other central nervous system diseases (CNS) generally involves processing and analysis of electroencephalograms (EEG). To aid medical personnel, working with EEG data, we have studied and implemented (in both stand-alone personal computers and parallel computing platforms) the automatic EEG spikes detection algorithm. The algorithm is implemented in Python programming language, employing Message Passing Interface (MPI) communications protocol. Beside other computing platforms, the software has been deployed and tested on supercomputer at Vilnius University, Faculty of Mathematics and Informatics, Information Technology Open Access Centre. The details of implementation, testing, performance and EEG processing examples are presented. In the first stage of the algorithm, EEG spikes are detected by applying morphological filters and Fourier transform. Additionally, we employ blind source separation method to improve signal quality and to reduce fake spikes detection rate. The EEG data is further processed to distinguish between sleep and awake EEG for even more value in diagnostic process by providing user with possibility to evaluate drowsiness in Karolinska scale. Power ratios of all brain rhythms are calculated as well (such data are useful in diagnosing other CNS diseases). The presented computer tool not only speeds up analysis of EEG data dramatically (as manual analysis is very time consuming task for medical personnel), but also enables to calculate parameters, which would be extremely hard to determine manually.
The Prototype of Lithuanian Speech Isolated Words Hardwired Recognizer

D. Navakauskas\textsuperscript{1}, A. Serackis\textsuperscript{1}, G. Tamulevičius\textsuperscript{2}, E. Ivanovas\textsuperscript{1}, T. Sledevič\textsuperscript{1}, L. Stašionis\textsuperscript{1}

\textsuperscript{1}Vilnius Gediminas Technical University  
Faculty of Electronics  
\textsuperscript{2}Vilnius University  
Institute of Mathematics and Informatics  
dalius.navakauskas@vgtu.lt

The prototype of Lithuanian speech isolated words hardwired recognizer for real-time tasks has been developed. The prototype is implemented on developed and assembled custom board utilizing Zynq-7000 chip that contains Artix-7 FPGA and Dual Core ARM Cortex A9. For the feature extraction purpose the programmatically selectable 12th order speech analysis techniques, such as Linear Frequency Cepstrum Coding, Linear Prediction Coding, Linear Prediction Cepstral Coding and Mel-scale Frequency Cepstral Coding, are implemented. 7800 words/s isolated word recognition speed is achieved using single dynamic time warping IP core. The recognizer is tested on male and female speech records under 30 dB and 15 dB signal to noise ratios. Isolated words recognition accuracy of 97 % is achieved. The hardware part of the recognizer is clocked at 50 MHz, while the software part – at 667 MHz frequencies. The power consumption during real-time operation is up to 3.6 W. The prototype is able to control home appliances and devices through standard IR, ZigBee, UART and USB interfaces. The user friendly interface is implemented for the prototype setup. Through voice commands and audible feedback as also as through minimal number of push buttons and light emitting diodes, modifications of existing device commands or input of new ones for a new device control can be easily achieved. The total FPGA slice utilization is not high – 21 %. Thus additional language or even specific audio signals could be used for the control. Therefore, the recognizer can be preconfigured and applied for smart applications in house, office, hospital, industry or nature environments.
Fitts Law is a background for every end user system. The software design made by thousands of rules is orientated into task execution time. But in our days more and more people are using two monitors for daily work, sometime curve monitors and rules of software design for that is not the same as for the flat screen. In article we are analysing data that was taken researching 2 - 4 years children by electroencephalography (EEG) signals to investigate the following issues, using the Time – Frequency Representations method, for each epoch corresponding to one presentation. Study experimental paradigms to understand simplesness of Fitts law for interpretable readable signals: mapping the task to the brain state of the user (child).
Research Challenges of Interactive Multiobjective Optimization

D. Podkopaev
System Research Institute
Polish Academy of Science
Poland
podkop@ibspan.waw.pl

Most of real-world optimization problems are multiobjective in their nature. Multiobjective optimization helps exploiting this nature by using preference information of the decision maker (DM) for making better decisions. The research topic of multiobjective optimization grows in popularity and quickly spreads among new application areas. Interactive methods are believed to be most promising practical methods of multiobjective optimization for their numerous advantages such as simple preference models, relatively small number of optimization problems to be solved, as well as the possibility of learning about the problem provided to the DM. The attempts of applying interactive methods of multiobjective optimization for solving computationally and cognitively complex problems create new challenges related to long DM's waiting times, high cognitive load and lack of transparency of the solution process. We describe some of these challenges and outline new research directions aimed at their addressing.

Comparison of Mods of Shoetrees Obtained by Theoretical and Experimental Methods

M. Sabaliauskas, V. Marcinkevičius
Vilnius University
Institute of Mathematics and Informatics
martynas.sabaliauskas@mii.vu.lt

The latest innovations in 3D acquisition techniques has enabled a highly digitization of complex 3D objects. Therefore a demand of efficient algorithms for preprocessing these objects has increased. The flattening algorithms are useful for texture mapping and they represent a straightened surface of 3D mesh. Another application of these algorithms could be automatic formation of molds using digitized shoetrees. This poster shows an investigation of evolvents obtained by flattening algorithms: ABF++, LSCM, and ARAP.
Following the Changes: HeLa cells Lineage from Phase Contrast Microscopy Time-Lapse Data

M. Sangiovanni, L. Maddalena, M. Guarracino

Phase contrast time-lapse imaging is crucial in biology for long-term, living cells investigations. In particular, time-lapse data may be used to extract the cell lineage, i.e., the set of mother-daughters relationships among proliferating cells. Lineage information is essential when investigating how phenotypic traits are passed to the progeny along different cell generations.

However, the automatic processing of time-lapse images is a very challenging task. This is due both to the huge amount of data usually produced, and to the nature of the imaged targets: cells may move, entering and exiting the field-of-view; they live, proliferate and die, continuously changing the density of the population; they usually show very complex topologies, because they may overlap and modify their shape when in close contact with each other.

In this work we present the first results obtained with a computational workflow, specifically designed to extract cell lineages from phase contrast time-lapse movies of proliferating HeLa cells. We describe the issues we faced and the solutions we found, starting from the raw movies and addressing the registration, segmentation, and tracking steps that are essential for a reliable lineage construction.

Since to the best of our knowledge there are no publicly available annotated datasets containing segmentation, tracking, and lineage data extracted from phase contrast time-lapse microscopy movies of dividing HeLa cells, we built our annotated dataset. These data have been used to assess the performance and reliability of the proposed workflow.
The Analysis of Influence of Noise Level to SCRD and Other Common Supervised Classification Methods

G. Stabingis, L. Stabingienė
Vilnius University
Klaipėda University
giedrius.stabingis@mii.vu.lt

The new supervised classification method was proposed (Stabingis et al., 2014). This method is called SRCD (Spatial Classification Rule with Distance) and is an improved version of SCR method. The improvement is done by incorporating more spatial information into classification rule, in this case by calculating the class prior probabilities according to the distance from the object to be classified to its neighboring training sample points. The main advantage of SCR and SCRD methods is their ability to model the additive correlated noise and to use this model during the classification. It is important to determine how strongly the image can be corrupted by the correlated noise in order to use such an image for classification. The large empirical analysis of noise level influence was performed and the results are presented here. 100 different BW images were taken and for the every image 4 different spatially correlated random fields were generated each representing different amount of correlation. These spatially correlated random fields were used to corrupt original images in 10 different levels of corruption. The corrupted images were classified with SCRD, SCR and other common supervised classification methods. The analysis of classification accuracy is presented here.

Reference
Knowledge Acquisition Method for Virtual Learning Environment

P. Tamošiūnas¹, S. Gudas²

¹Vilnius University
Kaunas Faculty of Humanities
²Vilnius University
Institute of Mathematics and Informatics
petras.tamosiunas@khf.vu.lt
saulius.gudas@khf.vu.lt

Demand of intelligent knowledge acquisition and processing methods is seen for the last decade. Virtual learning environments (VLE), as an educational business computerizing solutions, have analogous problems. Existing VLEs are rather primitive human-based empirical instruments. The paper is dedicated to intelligent knowledge acquisition in VLE. The drawbacks of empirical VLEs are emphasized. The paper describes modern e-learning mechanisms (methods), reveals main aspects of intelligent tutoring systems (ITSs). Control-view based knowledge acquisition method for ITS is proposed.

Reference Template Update Technique for Isolated Word Recognition

G. Tamulevičius¹, A. Serackis², T. Sledevič², D. Navakauskas²

¹Vilnius University
Institute of Mathematics and Informatics
²Faculty of Electronics
Vilnius Gediminas Technical University
gintautas.tamulevicius@mii.vu.lt

Isolated word (or small vocabulary) automatic speech recognition is implemented using pattern comparison technique usually. The main idea of pattern comparison based recognition is the comparison of unknown speech instance with the set of reference templates. The most similar template is declared as the equivalent to the unknown instance.

In this study we propose to estimate the quality of reference template using distance matrix with intra-reference and inter-reference distance values. Intra-reference distance shows distance between different templates of the same reference instance. Similarly, inter-reference distance presents the distance between templates of different instances.
Thus, reference templates with big intra-reference and inter-reference distance values are desirable for accurate and speaker-robust recognition of speech. The reference template update technique based on analysis of vocabulary distance matrix was introduced in this study. Reference templates with small intra-reference and inter-reference distance values are substituted with candidate instances giving bigger distance values. The technique was tested for Dynamic Time Warping based isolated word recognizer. The reference substitution using proposed technique increased the intra-reference and inter-reference distances by 6.8 % and 13.6 % respectively.

Heuristic Algorithms for Rescheduling of Production Processes with Advanced Precedence Constraints

V. Tiešis¹, G. Dzemyda¹, T. Kizil²

¹Vilnius University
Institute of Mathematics and Informatics
²Hisbim IT Corp. Atap OSB, 26110 Eskisehir, Turkey
vytautas.tiesis@mii.vu.lt

Production frequently encounters with undesirable business events that disturb running plans and schedules. Advanced enterprise resource planning systems usually have rescheduling modules; however, many quick rescheduling algorithms use simple models of the production at the shop floor such as Job Shop or Flexible Job Shop where precedence constraints of the job's operations are of the chain form. Two heuristic algorithms, namely, Affected Operations Rescheduling (AOR) and Active Wilkerson-Irwin (AWI) algorithms were enhanced by authors to cope with precedence constraints that may be described by any simple directed acyclic graph: each operation may have several precedents and several successors. The model of production may be regarded as a mutual execution of several independent projects. The AOR algorithm shifts to the right exclusively affected operations on the same machines preserving processing order. The aim of the method is to preserve the initial schedule as much as possible, in order to minimize disruptions of the raw materials’ and parts’ delivery schedule. The AWI algorithm dispatches operations on machines heuristically seeking to improve tardiness, make-span and stability (minimal deviation of new operations’ start times from the original start times). The AOR algorithm calculates more quickly and makes more stable schedules than AWI when there is
little percentage of affected operations, i.e. when the initial running schedule is flexible. Algorithms have been implemented in the Advanced Planning and Scheduling PEN system.

Acknowledgment
The research has been supported by the EUROSTARS Project E!6232-PEN "Production effectiveness navigator". The authors cordially thank the Agency for Science, Innovation and Technology for the financial support of the research.

Application of Ultrasound Spectral Analysis for Intraocular Tissues Analysis and Differentiation

J. Trumpaitis, R. Jurkonis, A. Paunksnis

Lithuanian University of Health Sciences
j.trumpaitis@gmail.com

At present time the ultrasound imaging technique is widely used for qualitative evaluation of ocular tissue structure. Differential diagnosis of human eye tumors is one of the most important problems in ophthalmology in dealing with cancer prevention and diagnostics. Simple ultrasonic methods such as A scan and B scan images used in ophthalmology helps to identify intraocular tumors, but does not reflect the microstructure, does not allow them to differentiate. New technique for tumors tissue structure evaluation using ultrasound spectral analysis is presented. Based on the obtained results, it can be said that RF ultrasound signals parameters (amplitude envelope, spectrum slope, spectrum intercept and momentary bandwidth) at the healthy tissue area and the area with the intraocular tumor – melanoma and hemangioma statistically are significantly different. This allows us to distinguish healthy tissues from the abnormal tumor tissues, to identify and differentiate chorioidal melanoma and hemangioma between each other.
Visualization of Pareto Front Approximations in the Space of Decisions

A. Varoneckas, A. Žilinskas
Vilnius University
Institute of Mathematics and Informatics
Audrius.Varoneckas@mii.vu.lt

A visualization of the optimal set of solutions of a multi-objective optimization problem is considered. We focus on discovery the structure of multidimensional decisions space by visualization of Pareto sets as well as on the relationship between objective functions and design variables in optimal set of solutions. A multi-objective visualization method is based on mapping a set of efficient multidimensional points into two-dimensional space. The method is analysed experimentally on several well known test functions. This method can help to explore the relationship between competing objectives and may enable a decision maker better understand the relation between different objectives and their place in the overall design space.

Acknowledgement
Postdoctoral fellowship of A. Varoneckas is being funded by European Union Structural Funds project Postdoctoral Fellowship Implementation in Lithuania.
Rule Based Dynamic Business Process Modelling and Simulation

O. Vasilecas, D. Kalibatienë

Vilnius Gediminas Technical University
Faculty of Mathematics and Informatics
diana.kalibatiene@vgtu.lt

In today’s dynamic business world the success of an enterprise increasingly depends on its ability to react to changes in its environment in a quick and flexible way. However, static business processes no longer fit to the actual needs of the business with dynamically changing environment. A necessity to model and simulate the dynamic business process (BP), which can be changed during BP execution, and adopt them to new requirements arises. Examples of such dynamic changes include regulatory adaptations (like, change of raw material prices), market evolution (like, stock price change), changes in customer behaviour (like, rapid change of customer needs), process improvement, strategic shifts and existence of exceptions. To implement this process agility, enterprises are increasingly interested in aligning their information systems (IS) in a process-centred way offering the right business functions to the right users at the right point in time along with the needed information and application services. Process-aware IS (PAIS) has emerged, which target at flexibility and agility of enterprise BP. Examples of such PAISs are as follows: workflow management systems, case handling tools, service orchestration engines. Separating process logic and application code makes PAISs more flexible in comparison to traditional data- or function-centred IS. However, this is not sufficient to meet today’s need for greater system flexibility. The major reason for this is that most existing PAISs require a complete specification (i.e. processes model) of a BP in advance, which is then used as the schema for process execution. However, dynamic BPs demand a more agile approach recognizing the fact that in dynamic environments process models quickly and often during process execution become outdated and hence require closer interaction of modelling and execution.

There exists the need to support the evolution of business process over time. Thus, modelling and simulation of dynamically changing BP, e.g. dynamic processes, is a topical and challenging task. Current BP modelling tools and most approaches are suitable only for static BP modelling and simulation, e.g. a static BP strictly prescribes, which activities and in which sequence to do. Moreover, a static BP does not allow its change during BP execution. In the best way, nowadays tools and proposed methods allow us changing BP by presenting several
occurrences of the same BP or presenting templates for BP execution. It means that all existing tools and approaches require strict specification or a model of a BP and unexpected sequence of BP activities cannot be included into a BP execution. Unlike, a dynamic BP is not defined strictly at the beginning of its execution and they changes under new conditions in its environment during execution.

Existing approaches on dynamic BP modelling and simulation are not complete before now, e.g. they luck theory or case study, or bought. Before describing the dynamic BP, we use the following BP definition. A business process is the combination of a set of activities within an enterprise with a structure describing their logical order and dependence, whose objective is to produce a desired business result. Hence, in our research, we propose a method for rule based dynamic BP modelling and simulation, which is based on changing of BP rules and accordingly rules actions during process execution to adopt the new environment conditions. It means that we have no fixed model of a BP. During execution of a dynamic BP instance, next activity is selected from set of possible activities according to the conditions, which are described in Event-Condition-Action (ECA) rules. Therefore, there is no strict specification of the BP, e.g. each activity is selected from existing and defined activities according to the condition from ECA rules. Based on the proposed method the prototype of dynamic BP simulation tool was developed. A simulation was carried out using this prototype, and the experiment clearly shows differences between dynamic and existing BP simulation methods, as well as advantages of proposed approach.

V. Veikutis¹, A. Bubulis², K. Morkūnaitė³, R. Unikas³

¹Lithuanian University of Health Sciences, Institute of Cardiology
²Kaunas University of Technology, Institute of Mechatronics
³Lithuanian University of Health Sciences, Department of Cardiology

vincentas.veikutis@lsmuni.lt

Ultrasound (US) in medicine develops mainly in two directions: as a value diagnostic toll and a powerful stimulant of various biological effects which could be useful for the treatment of many types of diseases. However mechanisms for many bio effects of US are unknown or not equal investigated. The goal of the present study was to investigate the response of isolated human arterial samples to low frequency US (4-6 W/cm², 20 kHz and 32,6 kHz). We hypothesized that US of pointed characteristics should modulate the vascular effects related with contraction and relaxation processes associated with modulation of calcium-channel function.

Experiments were carried out on isolated samples of human internal thoracic artery (n-152), obtained during CABG operations. Contraction was recorded using a tissue/organ bath system with iFOT10 force transducer. Calcium-chanel function was modulated using the benzodiazepine-type calcium-channel blocker diltiazem and muscarinic receptor’s agonist carbachol. Low frequency US resulted in a significant increase of isometric contraction force. Response of contraction didn’t depend upon time of exposure and exposition, however depended on the frequency and location (internal or external) of the vessel. Pre-treatment with diltiazem resulted in suppression of the contraction force both in US and in US-free groups by maintaining the contraction ration ~ 2.3:1 instead 2:1 which has been obtained in diltiazem-free setting. Relaxation of smooth muscle as a response to carbachol is independent from the US frequency, pulse quantity and period. Low frequency acoustic cavitations increased intracellular Ca²⁺ content involved in contraction mechanisms of the smooth muscles, surrounding cardiac vessels.
IRT in Endocardial Transcatheter Laser Ablation
Safety and Effectiveness Control

V. Veikutis\textsuperscript{1}, A. Puodžiukynas\textsuperscript{3}, A. Sakalauskaitė\textsuperscript{2},
K. Stasiukynaitė\textsuperscript{2}

\textsuperscript{1}Lithuanian University of Health Sciences, Institute of Cardiology
\textsuperscript{2}Lithuanian University of Health Sciences, Medical Academy
\textsuperscript{3}Lithuanian University of Health Sciences, Department of Cardiology
vincentas.veikutis@lsmuni.lt

Catheter radiofrequency ablation (RFA) has been standard and mostly
developed approach. Unfortunately, despite of increased experience and
using modern RFA technologies postoperative complications remain
high. Recently, most novel techniques are optimized and tested mostly
in an isolated thigh-muscle model. Aim of our study was to test laser
ablation technique doing circular laser ablation procedure in pulmonary
vein area, analyze possible complications and determine optimal
ablation parameters on beating pig heart model.

4 female pigs (weight ~ 20-25 kg) used in experimental study. Hearts
were exposed via sternotomy and pericardtomy, for visual and catheter
position control. For laser ablation we used loop form 10F electrode-
laser catheter, inserted through the \textit{a. femoralis}, position in the heart
controlled by finger, ablation parameters - 10-35 W, 120 s. Temperature
changes in effected heart epicardial surface were registered with
thermocamera \textit{ThermaCAM P640 (FLIR Systems)}. During the
procedures changes in tissues were observed visually, later destruction
zones were excised, cute and the damage of the myocardium was
evaluated histologic. Using laser ablation destruction zone was total,
even and forming regular oval. Destruction zone was statistically
reliably even in all sites of heart and almost independent from
application time. The laser causes localized heating-up of the heart
tissue, with temperature scale 34.1 - 70.2\textdegree C and enough restricted scar
formation. Using standard recommended parameters 40W/180 sec and
35W/120 sec evoked quick critical scar formation. Temperature in
cardiac surface differs from 34.5 till 70.2 \textdegree C and directly depends from
electrode compression level, and energy power. Optimal detected
ablation parameters – 20-30W/60-120 sec.
The Value of IRT on Integrated Investigation of Patients with the Diabetic Neuro- and Angiopathy

V. Veikutis¹, A. Sakalauskaitė², K. Stasiukynaitė², E. Monstavičius²

¹Lithuanian University of Health Sciences, Institute of Cardiology
²Lithuanian University of Health Sciences, Medical Academy

vincentas.veikutis@lsmuni.lt

Angiopathy, microangiopathy, and neuropathy induced angiopathy play an important role in the pathogenesis of the diabetic foot. They are responsible for subtle skin temperature changes, in these cases IRT is becoming the investigation of choice in the evaluation of the diabetic foot. Aim of study was to assess IRT usage in differentiation between neuro- and angiopathy, evaluating the temperature differences in patient with diabetic foot. We examined 64 patients, divided in to 3 groups: control (n=17) non diabetic volunteers, diabetics without ulceration on anamnesis (n=36), and diabetics (n=11) with ulcers in presence or episodes of ulceration in past. Neuropathy was defined using standard neurologic examination: 10g monofilament touch, vibration, pain, temperature perception disorders, when obtained data were compared with thermography results. Tissues of both feet’s were observed visually, whereas the micro temperature changes were supervised and registered with IRT camera “ThermaCAM P640” (FLIR Systems, USA). Protocol of investigation included standard 10-15 min adaptation in room temperature (20-24 C°), in sitting position with the lower extremities hanging freely. Thermal/digital images were made from external and plantar sides. Thermographic views were analyzed with “FLIR tools” computer software. Differences of investigated parameters were compared by the Student’s test, where p<0.05 was designated as statistically significant.

We founded significant thermal differences between neuro- and angiopathic foot. DM patients with polyneuropathy were accompanied by seriously impaired thermoregulation. We consider use of IRT a valuable diagnostic tool for easy detection of potential DM induced neuropathy. Further studies are required to establish the value of IRT in visualization of early subclinical complications.
New Possibilities in Psoriatic Arthritis Diagnosis Using IRT

V. Veikutis¹, S. Valiukevičienė², V. Kučinskienė²

¹ Lithuanian University of Health Sciences, Institute of Cardiology
² Lithuanian University of Health Sciences, Department of Skin and Venereal Diseases

vincentas.veikutis@lsmuni.lt

Diagnostic criteria of psoriatic arthritis (PsA) include inflammatory arthritis, the presence of psoriasis and absence of seropositivity for rheumatoid factors. The evaluation of the PsA activity usually based on clinical, laboratory (CRP, ESR) and radiology findings. However, in early stage of disease is difficult to diagnose PsA, especially when laboratory findings are normal and X-ray joints examination doesn’t show pathology at yet. Our study was designed to estimate the surface temperature over joints by using computerized digital infrared thermal imaging method (IRT) and to evaluate its comparability with clinical features of arthritis, laboratory findings (CRP, ESR) and DAS 28 scale.

9 male and 7 female psoriasis patients were enrolled into the study. Nine patients had the diagnosis of PsA. We evaluated psoriasis area, severity index (PASI) and arthritis activity score (DAS 28). Control laboratory tests - CRP and ESR levels. The joints were divided into three groups: large, small, and axial. IRT/digital images (59 joints per subject) were taken using ThermaCAM P640 camera. Statistic SPSS 20 program was used to analyze data. Observations were made on a 944 joints. The supreme frequency between temperature and pain was observed in large joints (69%), lower in small joints (43%) and the lowest in axial ones (30%). Higher temperature as marker of direct active inflammation process was more often detected in large joints than in small or axial ones. We didn’t find statistically significant correlations between laboratory test results, DAS 28 and IRT findings in small and large joints. However, increased levels of CRP correlate with the markedly higher surface temperature in axial joints.
The main scope of the article is to present enterprise model based UML dynamic models generation process. In information system engineering most design stage models are implemented on the basis of the expert’s empirical experience and knowledge. Computerized information system in knowledge-based information system engineering, is developed by using stored enterprise knowledge base of the subject area, i.e., enterprise model, the composition of which is defined by formal criteria. Enterprise model contains essential elements of business modeling methodologies and techniques, which ensures a suitable UML dynamic models generation process. Enterprise model as essential organization's knowledge repository allows generate UML dynamic models using the transformation algorithms.
Employing Analytics to Address Customer Needs

Elias Vyzas
Center of Excellence for Enterprise Intelligence
Central & Southeast Europe, Ernst&Young
elias.vyzas@gr.ey.com

How does the progress on data analytics translate into real life industry solutions?

EY is using data analytics, to offer advice to its customers. Indicative use cases of how EY has employed analytics to address customer needs in financial industry, energy and retail will be presented. Cases provided will cover:

How can utility companies manage to improve the effectiveness of their on-site inspections, either for fraud or maintenance? Thus, reducing their losses and inspection costs.

How can financial organizations assess the performance of their credit models?

How can banks decide how much debt reduction to offer to their non-responsive customers in order to incentivize them to pay?

How can companies optimize the footprint of their retail network taking into account geo-marketing data?

How can companies assess a person’s social type and interaction profile using social media, text analytics?
Challenges of Big Data Visualization

J. Zubova, O. Kurasova

Vilnius University
Institute of Mathematics and Informatics
elena.zubova@mii.vu.lt

Recently, huge amounts of data are constantly generated by various devices and stored in data repositories. Usually, not only data volumes are huge, but also the data are renewed and updated very quickly and variety of data types and sources is high. These data are called big data. We confront with difficulties to process and analyse big data in many areas, eg., medicine, finance, social networks, telecommunication, etc. Moreover, big data visualization is challenge for data analysts and researchers, because the methods, ways and tools, applied to visualize ordinary data, become unsuitable for big data visualization. Effective visual representation of data allows detecting and extracting useful information from data in simple and user-friendly ways. So, this our research focuses on big data visualization. The aim is to overview ways and techniques applied for big data visualization and to indicate problems arising in this field. It is purposeful to review visualization tools in big data analytics as well as in open source software, to classify them, and to highlight their advantages and disadvantages. Another problem encountered is big data visualization based on dimensionality reduction, because the methods such as multidimensional scaling, principal component analysis, etc. are unable to manage the challenges of big data. It is necessary to develop new or modify existing methods suitable for big data visualization. Technological issues on these problems are considered in the research, too. Hadoop platform, as well as other parallel and distributed computing resources can be of service to solve big data visualization problems.
On an Asymptotic Property of a Simplicial Statistical Model of Global Optimisation

A. Žilinskas, G. Gimbutienė

Vilnius University
Institute of Mathematics and Informatics
grazina.gimbutiene@mii.vu.lt

A homogeneous isotropic Gaussian random field is accepted as a statistical model of objective functions, aiming to construct global optimisation algorithms. The asymptotic of the conditional mean and variance is considered, assuming that the random field values are known at the vertices of a simplex, and that the latter is contracting. The obtained result theoretically substantiates the construction of the recently proposed bi-variate global optimisation algorithm, which arouses interest due to good performance in testing experiments and the established convergence rate. The obtained result also substantiates the extension of the aforementioned algorithm to higher dimensions.

A New Approach to Rational Search for Global Minimum

A. Žilinskas

Vilnius University
Institute of Mathematics and Informatics
antanas.zilinskas@mii.vu.lt

The problem of approximating the global minimum of a smooth function is considered. The algorithm is motivated by the ideas of optimal search with respect to a statistical model of objective functions. New axiomatic of the rationality of search is proposed. Two implementation schemas are considered. The first implementation uses Delaunay triangulation, and the second is based on hyperrectangular partition of the feasible region. Only the objective function values are required by the optimization algorithm. The asymptotic convergence rate is discussed for a class of smooth functions, for example the attainability of the upper bound, \( \delta_n \leq \exp(-\sqrt{nC}) \), for the approximation error of global minimum. Numerical examples are presented.

Acknowledgement.
This research was funded by a grant (No. MIP-051/2014) from the Research Council of Lithuania.
6th International Workshop
DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS

Druskininkai, Lithuania, Hotel “Europa Royale”
http://www.mii.lt/DatAMSS
December 4–6, 2014