SPMS 2018

Monday 18 June 2018 - Friday 22 June 2018

Sokol Dobřichovice

Book of Abstracts

Dear Friends and Colleagues,
it is our pleasure to welcome you to the 9^{th} International Conference
STOCHASTIC & PHYSICAL MONITORING SYSTEMS.

The SPMS 2018 conference is held with the aim to bring together students and researchers with areas of interest related to the following topics:

- Analysis of microscopical structure of vehicular traffic streams and traffic modeling,
- Monitoring and classification of acoustic signals in material defectoscopy,
- New statistical distances and informational divergences with applications in acoustic emission,
- Small area estimation of geographical characteristics gained from data sets,
- and Data processing in high energy particle physics,

which are all together briefly called Stochastic and Physical Monitoring Systems. The conference links both the informal character of lively student meeting and the unique platform for the research presentations and discussions of the conference participants.

The meeting is organized by the Group of Applied Mathematics and Stochastics (GAMS), Department of Mathematics, FNSPE Czech Technical University in Prague and this year's venue is for the third time in Dobřichovice, just a stone's throw from Prague, close to one of the most famous Czech castle, Karlštejn, and has many sports facilities.

Local Organizing Committee:
Václav Kůs (Chair),
Jiří Franc,
Jana Vacková,
Petr Bouř.

Scientific and Program Committee:
Milan Krbálek,
Tomáš Hobza,
Václav Kůs,
Jiří Franc.

We gratefully acknowledge the institutional support provided by the funds of specific research, namely by the SGS grant SVK 24/18/F4.

`The Organizers`

Contents

Traffic and agent monitoring systems / **16**

3S-Unification for Vehicular Headway Modeling

Milan Krbálek[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): milan.krbalek@fjfi.cvut.cz

We explain why a sampling (division of data into homogenous sub-samples), segmentation (selection of sub-samples belonging to a small sub-area in ID plane), and scaling (a linear transformation of random variables representing a standard sub-routine in a general scheme of an unfolding procedure) are necessary parts of any vehicular data investigations. We demonstrate how representative traffic micro-quantities (in an unified representation) are changing with a location of a segmentation zone. It is shown that these changes are non-trivial and correspond fully to some previously-published results. Furthermore, we present a simple mathematical technique for the unification of GIG-distributed random variables.

Traffic and agent monitoring systems / **7**

A novel approach to interaction detection in vehicular traffic

Zuzana Szabová[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): szabozuz@fjfi.cvut.cz

Knowledge of an interaction range in particle systems, especially in vehicular traffic could significantly contribute to modeling of traffic flow. Combination of simulation methods, analytical predictions of headway distribution, and correlation analysis led to several remarkable observations. We observe, that interaction range depends on both resistivity and type of repulsive potential. Moreover we introduce a novel method for detection of number of actively followed vehicles based on perturbation function. Beside that, significant progress has been made in theory of balanced density functions, which can be of help in derivation of distribution of clearances in vehicular systems.

Stochastic monitoring systems / **22**

A practical look on financial mathematics: price, returns and challenges

Petr Veverka[1]

[1] *FinMason Europe*

Corresponding Author(s): panveverka@seznam.cz

In the talk, we start with some basic concepts of finance such as price and returns. Then we mention briefly how these variables can be modelled and what is the economical mechanism behind them. Futher, several classical trading problems and strategies will be introduced. Special attention will be payed to some statistical issues on real data (change point detection, data clustering, free online data sources etc.).

Small area estimation / **31**

An application of gamma mixed models to small area estimation

Author(s): Tomáš Hobza[1]

Co-author(s): Domingo Morales [2] ; Yolanda Marhuenda [2]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

[2] *UMH Elche, Spain*

Corresponding Author(s): tomas.hobza@fjfi.cvut.cz

Average incomes and poverty proportions are additive parameters obtained as averages of a given function of an income variable. As the variable income has an asymmetric distribution, it is not properly modeled via normal distributions. When dealing with this type of variables, a first option is to apply transformations that approach normality. A second option is to use non-symmetric distributions like the gamma distribution.
This contribution proposes unit-level gamma mixed models for modeling asymmetric positive variables and for deriving three types of predictors of small area additive parameters, called empirical best, marginal and plug-in. The parameters of the introduced model are estimated by applying the maximum likelihood method to the Laplace approximation of the likelihood. For estimating the prediction error an adaptation of parametric bootstrap is proposed.
Some simulation experiments are carried out to study the behavior of the fitting algorithm, the small area predictors and the estimator of the mean squared errors. By using data of the Spanish living condition survey of 2013, an application to the estimation of average incomes and poverty proportions in counties of the region of Valencia is given.

Stochastic monitoring systems / **10**

Analysis of Egyptological Data - Title Classification

Author(s): Marek Bukáček[1]

Co-author(s): Jana Vacková [1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): marek.bukacek@seznam.cz

The scientists from the Czech Institute of Egyptology, Charles University have uncovered, collected and analysed historic materials for decades. Their databases contain millions of records giving interpretation to both, material objects and real people.

In the present project, we focus on the period of Old Kingdom (2700 - 2180 BC), especially to the development and the fall of social infrastructure. Using titles depicted on the facades of tombs, we get the personal details of several thousands citizens, from the pharaohs to the last scribes. Although there is a certain uncertainty in this dataset, we can categorize and rank professions, analyse family relations or study religious aspects.

The first task (and this paper) handles with categorizing titles to more general groups representing the pillars of administration, religion or honorific maters. Bayesian analysis based on the distribution of titles among the population is used to show the social distance of the groups and the hierarchy of this system. The time scale is incorporated by selecting the part of population living only in specific period, e.g. during the rule of some dynasty.

The preliminary results show a grouping tendency, the titles related to the king are often accompanied by various secular or priestly titles, on the other hand the craft group frequently related to a location tag, as ‘two palaces’ or ‘great temple’. Many correlations are unstable in time which implies changes in the society. And exactly such dynamics is the global object of this project.

Small area estimation / 19

Area-Level Gamma Mixed Model

Ondřej Faltys[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): ondrej.faltys@fjfi.cvut.cz

Area-level model with responses having conditionally the gamma distribution is introduced. It is a special kind of a generalized linear mixed effects model that can be useful in some applications involving only positive responses (e.g. in a financial sector). To obtain estimates of the regression parameters, the penalized quasi-likelihood (PQL) algorithm and ML Laplace approximation algorithm are employed. Three simulation experiments are performed to evaluate the quality of the estimates of the regression parameters. Based on results of the experiments, the PQL algorithm is not convenient for this task. The ML Laplace algorithm seems to work properly.

Small area estimation / 28

Asymptotic properties of the modified median estimator

Jana Novotná[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): novotj88@fjfi.cvut.cz

A theorem about asymptotic distribution of the modified median estimator for logistic regression models was published last year. One of the assumptions of this theorem is that the modified median estimator is the consistent estimator. Because this assumption has not been verified yet, two simulation experiments were carried out. In the first experiment consistency of the modified median estimator was studied and in the second experiment the statement of the theorem was verified. The main aim of this presentation is to present the results of the experiments.

Stochastic monitoring systems / 14

Bayesian Approach to Source Term Estimation

Karel Hybner[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): hybnekar@fjfi.cvut.cz

When a radioactive substance is detected in the air, the primary task is to determine the location of the release source as well as its magnitude and temporal variation (i.e. to determine the source term). This task can be,under some assumptions, formulated as a linear inverse problem. The vector of measured values (concentrations) can be in this case written as a product of the source-receptor sensitivity (SRS) matrix (obtained from an atmospheric transport model) and the unknown source term vector. As this problem is typically ill-posed, the classic method of solution is a regularization of this problem (Tikhonov regularization, LASSO, etc.) and subsequent optimization. A drawback of these methods is the sensitivity to the choice of the so-called regularization parameter. Hence, as a possible alternative, a probabilistic Bayesian model is formulated and subsequently estimated using technique of variational Bayes approximation. The main advantage of the resulting algorithm

compared to conventional techniques is the estimate of the regularization parameter directly from data, along with the estimate of the source term.

Acoustic emission / **39**

Classification of Acoustic Emission Signals and Application of Phi-Divergence in the Field of Non-Destructive Testing

Zuzana Dvořáková[1] ; Václav Kůs[2] ; Zdeněk Převorovský[1] ; Serge Dos Santos[3]

[1] *Institute of Thermomechanics CAS, Prague, Czech Republic,*

[2] *Department of Mathematics, FNSPE Czech Technical University in Prague*

[3] *INSA Centre Val de Loire, Blois, France*

Corresponding Author(s): zuzana.dvorakova@fjfi.cvut.cz

The acoustic emission (AE) belongs among methods of nondestructive testing serving to investigate materials. One of the most important task is to clasify measured source of AE, because not only cracks and other damages produce AE, but also external forces and stress in the materials produce signals of AE. Our aim is to classify measured AE signals. We compare several classification methods as Fuzzzy Classification (FC), Support Vector Machines (SVM), Model Based Clustering (FBC) and new designed Divergece Decision Tree (DDT). The last one method uses the phi-divergence as a decision criterion for the classification. Application of the phi-divergence as a parameter in the field of non-destructive testing is our next step, because it is necessary to compare measured signals by means of some parameters. The phi-divergence seems to be good parameter for comparing spectra from various sources. Different methods and classification parameters are compared by means of several experiment in the field of non-destructive testing.

Stochastic monitoring systems / **37**

Cyber-Egyptology: Cybernetic methods applied on data from Ancient Egypt

Radek Mařík[1] ; Veronika Dulíková[2]

[1] *Department of Telecommunications Engineering, FEE Czech Technical University in Prague*

[2] *Czech Institute of Egyptology, Faculty of Arts, Charles University*

Corresponding Author(s): veronika.dulikova@ff.cuni.cz

The sophisticated administration in the age of the pyramid builders offers a remarkable time span for research and a unique opportunity to analyse the dynamics of a complex society in a diachronic perspective. Although scholarly interest in the Old Kingdom administration has always been relevant, the grasping of its complexity and the tracing of the particular processes which led to changes and innovations of the system have been missing. Their study is crucial, because it adds a valuable insight to our knowledge of the varying Old Kingdom social and administrative structure.
Contrary to traditional approaches relying on statistics and logic, we will present an overview of our achievements in society development reconstruction covering both structural and dynamic aspects using a number of methods of cybernetic and artificial intelligence that provide and ensure: automated grouping people into families (uncertainty handling and logic), automated family tree building, layout, and visualization (various possibilities of visualization were implemented), automated detection of families with a significant level of nepotism (techniques of social network analysis and data mining), detection of strategic titles and powerful officials (information theory), development of administration during the Old Kingdom (hidden Markov models), assessment of society stratification from an individual perspective (community detection, Fruchterman-Reingold method), etc.
On the ground of this approach, we defined cyber-Egyptology as a process of interpretation based

on (semi-)automated processing of data volume gained and used in Egyptology using a number of cybernetic techniques, aimed on the one hand to grasp mechanisms of society development and transformation, on the other hand to predict and to influence positively the future development of humankind.
Our interdisciplinary team is successful at better understanding of the ancient Egypt state and its legacy for future of our society, since then mechanisms worked which are effective also for our modern time. Our Egyptological-cybernetic cooperation excited interest as well and showed that big discoveries do not emerge only from the sand.

Data processing in HEP / **5**

Deep Learning at NOvA

Petr Bouř[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): petr.bour@fjfi.cvut.cz

With the raise of modern computing capabilities and new approaches in deep learning, we are able to design convolutional neural networks suitable for purposes of particle identification at NOvA Experiment in Fermilab. Utilizing deep learning techniques leads to the signifcant increase in signal effciency classifcation. We provide an overview of the experiment setup, raw data measurements and application of convolutional visual network.

Data processing in HEP / **20**

Deep Learning in High Energy Physics

Miroslav Kubů[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): kubumiro@fjfi.cvut.cz

Data analysis in high energy physics includes solving difficult classification tasks; hence the deep learning approaches such as deep neural networks and convolutional neural networks (CNN) are often used. The core problems of particle identification share many similarities with the problems faced in computer vision. We describe the benefits of CNN in the area of image recognition tasks originating from its ability to learn features from raw image pixels. Following a summary of the core properties of CNN with experiments demonstrating the effectiveness of the approach, we discuss the possible application of CNN to the NOvA neutrino experiment in Fermilab.

Stochastic monitoring systems / **35**

Domination relations between distances on probability spaces, their properties and implications to L_1 consistency of MDE

Jitka Hrabáková[1] ; Václav Kůs[2]

[1] *Department of Mathematics, FIT Czech Technical University in Prague*

[2] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): jitka.hrabakova@fjfi.cvut.cz

This contribution deals with domination relations between distances on probability spaces. Implications of domination relation between Kolmogorov and total variation distance to L_1 consistency of Kolmogorov MDE is known. We will extend this result to another distances (Lévy, discrepancy, bounded divergences), for this purpose we focus especially on relation of particular distances to total variation distance and Kolmogorov distance, which enable us to prove L_1 consistency of respective minimum distance estimators using formerly published results. Various assumptions leading to consistency in L_1 norm and expected L_1 norm of particular MDE are studied and compared. Further, generalization of domination relation, so called asymptotic domination relation is introduced and relation to original domination relation is studied. We will show that asymptotic domination relation suffices to ensure L_1 consistency of MDE, thus we prove the same results for Kolmogorov and other distances under general assumptions.

Data processing in HEP / **6**

Estimates of the learning set size for k-NN and IINC separating methods in the high energy physics.

Frantisek Hakl[1]

[1] *Institute of Computer Science CAS, Prague, Czech Republic*

Corresponding Author(s): hakl@cs.cas.cz

Reliability of separation methods based on learning with the teacher (supervised learning) is primarily studied by verifying the independence of the achieved results on selected parts of data sets used. For this purpose, both data exploited in the process of separator parameters settings as well as independent test data are used. For example, the first one data are frequently used in so called cross-validation and the second one in the test of over-learning.

More sophisticated methods of verifying the reliability of learning methods with a teacher exploits additional knowledge of statistical characteristics of the data processed. These expected statistical characteristics are tested by standard statistical procedures (as an example can serve the statistical distribution of the mutual energy of $M_{b\bar{b}}$ pair in the decaying tree of $p\bar{p}$ collision).

All of these methods are based on the properties of the processed data only and do not give any assessment of the suitability of the method used to process the particular data. At the same time these methods do not provide any information regard to appropriate amount of separated data or convenient complexity of the separating model.

Concurrently an exact theory of PAC-learning has been developed for supervised learning methods. PAC-theory provides quantitative relationship between the number and dimension of the processed data on the one hand and the appropriate size of the parametric space of the separation methods on the other hand, under predefined conditions on expected quality and reliability of separation.

In our contribution we will show the application of the PAC-theory to separation methods of k-NN and IINC type. We will obtain necessary characteristics of these methods in the terms of PAC-learning and indicate application of derived data size estimates in the field of analyzing data produced by elementary particle detectors.

Traffic and agent monitoring systems / **30**

Flow aproximation in cellular models of pedestrian dynamics

František Gašpar[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): gaspafra@fjfi.cvut.cz

This work presents the approximation of stationary flow through a bottleneck in the cellular model of pedestrian dynamics. The method consists of the use of an approximation model of reduced bottleneck neighborhood and application of the Markov chain theory. The formulated procedures are used to approximate flow in the floor-field model in the entire parametric space of the model. The same approximations are also used for the model using heterogeneous agents. In all cases, the approximations are compared with the values yielded by Monte-Carlo simulations. The results show that the approximations correspond well with the observed values. On the basis of approximations and known flow value a procedure for initial calibration of the model is formulated.

Traffic and agent monitoring systems / 12

Follower-Leader Concept in Microscopic Analysis of Pedestrian Movement in a Crowd

Author(s): Jana Vacková[1]

Co-author(s): Marek Bukáček [1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): janca.vackova@fjfi.cvut.cz

This paper presents a microscopic analysis of factors influencing pedestrian movement and interactions with their surroundings for two considered modes: independent movement influenced only by the surrounding conditions and synchronized movement based on following another pedestrian.

This study analyses which of these effects prevail in different phases of the movement. The results show that the significant value of correlation between pedestrian velocity and corresponding individual density is observed mainly during approaching the crowd. Contrarily, in the segment of pedestrian trajectory which corresponds to movement inside the crowd, correlation between the velocity of follower and leader is more important.

This confirms the pedestrian behaviour in a crowd is such a complex field.

Acoustic emission / 17

Geodesic optimization and acoustic emission localization maps processing

Petr Gális[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): galispet@fjfi.cvut.cz

We furthermore extend numerical model of localization of acoustic emission (AE) sources on real complex solid bodies based on exact geodesic curves on 3D vessels composed of several parametrized surfaces. The numerical computations are provided via Finite difference, Newton–Raphson, and Fixed-point iteration methods applied to geodesic equations. To speed up computations, some technical improvements and optimizations are proposed. The variable propagation velocity and also the case when the geodesic curve has to bypass given obstacles is also included in the model. These techniques are employed in the real experiments, with aluminium watering can and steam reservoir. The resulting localization maps of AE using length (Δl), or time (Δt) differences, are then

processed through the two-dimensional Kernel probability density estimates executed directly on the 3-D surfaces, which give us the most probable areas of the AE source positions. The placement of piezo-ceramic AE sensors is outside the central part of the vessel because it can be inaccessible due to possible high temperature or radioactivity, such as in the case of nuclear power station health monitoring. This outward position of all AE sensors can result in a dispersed or attenuated AE waves because of welded intersections of different surfaces. Thus, the change-point analysis of AE signals is also discussed in order to obtain the most precise arrival times of AE events, which is crucial for Δl or Δt localization.

Stochastic monitoring systems / **4**

Heuristics in blind source separation

Václav Kautský[1] ; Jakub Štěch[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): kautsvac@fjfi.cvut.cz, stechjak@fjfi.cvut.cz

This paper deals with application of heuristic algorithms (DEBR, MCRS) in blind source separation (BSS). BSS methods focus on a separation of the (source) signal from a linear mixture. The idea of using heuristic algorithms is introduced on the independent component extraction (ICE) model. The motivation for considering heuristics is to obtain an initial guess needed by many ICE algorithms. Moreover, the comparison of this initialization, and other algorithms accuracy is performed.

Data processing in HEP / **26**

Homogeneity tests in high energy physics

Adam Novotný[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): novotad2@fjfi.cvut.cz

Homogeneity tests play a significant role in analysis of experimental data in high energy physics. These tests verify whether measured data samples and Monte Carlo simulated samples consistent with Standard Model come from the same distribution. Due to the fact that every simulated sample comes with a corresponding weight, we propose the modification of known homogeneity tests, namely Kolmogorov-Smirnov, Anderson-Darling, Cramér-von Mises and Pearson χ^2.

Modified tests along with their asymptotic behavior need to be firstly verified by generated samples and weights. Because all four tests have different properties, comparison between them is then studied. Finally, tests can be applied to weighted data sets, which were obtained at DØ experiment in Fermilab.

Stochastic monitoring systems / **36**

Identification of malicious Autonomous systems

Dominik Vít[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): vitdomin@fjfi.cvut.cz

Autonomous system (AS) is a group of routers and IP prefixes under the control of single routing policy and administrative. The possibility of acquiring the name of the visited AS only by observing network traffic suggests their exploitation in the domain of web security. In our work we stern from previous efforts in creating a set of detectors and introduce new possibilities of using the reputation of ASes as well as an improved method of computing ROC curves used for evaluating efficiency of these detectors. We also focus more on exploring eligibility of this approach and the time evolution of ASes. From real traffic data we observe the level of their stability and then use the property of Markov chains of first and second order to determine their future state given the knowledge of their history. By using methods of Monte Carlo we then simulated this prediction and evaluated it on a training data set.

Small area estimation / **13**

Iterative Methods for Fast Reconstruction of Undersampled Dynamic Contrast-Enhanced MRI Data

Author(s): Hynek Walner[1]

Co-author(s): Michal Bartoš [2] ; Marie Mangová [3] ; Olivier Keunen [4] ; Rolf Bjerkvig [4] ; Radovan Jiřík [2] ; Michal Šorel [5]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

[2] *The Czech Academy of Sciences*

[3] *Brno University of Technology*

[4] *Luxembourg Institute of Health*

[5] *Institute of Information Theory and Automation CAS, Prague, Czech Republic*

Corresponding Author(s): walner@utia.cas.cz

This paper introduces new variational formulation for reconstruction from subsampled dynamic contrast-enhanced DCE-MRI data, that combines a data-driven approach using estimated temporal basis and total variation regularization (PCA TV). We also experimentally compare the performance of such model with two other state-of-the-art formulations. One models the shape of perfusion curves in time as a sum of a curve belonging to a low-dimensional space and a function sparse in a suitable domain (L + S model). The other possibility is to regularize both spatial and time domains (ICTGV). We are dealing with the specific situation of the DCE-MRI acquisition with a 9.4T small animal scanner, working with noisier signals than human scanners and with a smaller number of coil elements that can be used for parallel acquisition and small voxels. Evaluation of the selected methods is done through subsampled reconstruction of radially-sampled DCE-MRI data. Our analysis shows that compressed sensed MRI in the form of regularization can be used to increase the temporal resolution of acquisition while keeping a sufficient signal-to-noise ratio.

Stochastic monitoring systems / **38**

Little data analysis of bone marrow transplant patients

Kateřina Henclová[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): katerina.henclova@fjfi.cvut.cz

Sinusoidal obstruction syndrome is a serious complication of chemotherapy for bone marrow transplant patients. We are provided with a series of blood samples of patients, who did or did not experience this syndrome during their treatment. In the dataset of twenty patients, each represented by

hundreds of measurements, we strive to find the biomarkers defining the syndrome and to predict which patients are likely to experience it. Working with extremely limited and inaccurate data, we attempt to tailor tools of machine learning and big data analysis to this practical problem.

Small area estimation / **8**

Mixed effect model for SAE

Tomáš Košlab[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): koslatom@fjfi.cvut.cz

One of the crucial parts of all models used in Small Area Estimation are the effects of small areas. In general, these are modelled either by fixed or random parameters. As the majority of research questions is posed on area level - such as prediction of area means or finding a percentage of the population over/under a specific threshold in every small area - the handling of area effects plays a key role in determining the overall quality of the model. In this work, a new model containing both fixed and random area effects is presented. Formulas for area means predictions using the plug-in predictor and the Empirical Best Predictor are given. A simulation experiment illustrating the gain in precision obtained by the use of the proposed model is presented.

Data processing in HEP / **21**

Multivariate Data Analysis and Machine Learning in HEP

Jiří Franc[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): jiri.franc@fjfi.cvut.cz

Multivariate Analysis (MVA) techniques using machine learning (ML) algorithms play an important role in many High Energy Physics (HEP) data analyses. In the last decades, the development of ML in HEP has lived its own life, but it changes. In the last couple of years, HEP community has discovered very powerful tools and methods from industry and adapt them to their unique and interesting problems. This contribution is intended to sketch and introduce how this progression took place and what is the state of the art. Some MVA and ML techniques are presented together with applications we met on different experiments.

Data processing in HEP / **27**

Neural networks for text classification

Marko Sahan[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): sahanmar@fjfi.cvut.cz

The aim of this work is to propose a semi-supervised algorithm that provides multi-class classification of documents with as little amount of training documents as possible. The basis of the algorithm is a neural network classifier that has a vectorized representation of a document on its input and class of the document on its output. Conventional classification approaches train the network using point estimates and thus require large amounts of data. We aim to approach the problem in a semi-supervised manner where we can ask for correct document class in the training phase. To select such documents that bring as much information about the classifier parameters, we aim to estimate the parameters in a Bayesian way and compute their expected information gain. Different types and modifications of Neural Networks algorithm are compared with Naive Bayes and Support Vectors Machine methods.

Acoustic emission / **33**

Noninvasive testing of human skin properties in-vivo

Julie Kuklewicz[1] ; Zdeněk Převorovský[2]

[1] *INSA Centre Val de Loire, Blois, France*

[2] *Institute of Thermomechanics CAS, Prague, Czech Republic,*

Corresponding Author(s): julie.kuklewicz@insa-cvl.fr

Human skin is a complex stratified biomaterial with nonlinear viscoelastic mechanical behavior. Its mechanical properties in-vivo are of great interest in dermatology, plastic surgery and cosmetics. A special small skin loading device with embedded ultrasonic probes is used for noninvasive in-vivo testing of the skin in NDT laboratory of the Institute of Thermomechanics. It enables instantaneously determine elastic modules and viscoelastic model parameters of the skin, and also to study nonlinear properties under various conditions and influences. The device is used for the skin studies already many years, but new electronic equipment and signal processing techniques allows for realize testing more precisely and reliably. Measured parameters and their evaluation are described in this contribution, and examples of some tests results are presented and discussed.

Acoustic emission / **9**

Preisach-Mayergoyz space in elasticity and damage assessment of hysteretic materials

Author(s): Colette Kožená[1]

Co-author(s): Václav Kůs [1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): kozencol@fjfi.cvut.cz

Preisach-Mayergoyz model (PM model) of hysteresis can be used to evaluate the mechanical properties of hysteresis materials. Program called PM identifikator with it's main function, automatic identification of the PM space probability density within a class of distribution mixtures by means of optimization stochastic algorithms and ϕ-divergence measures, is presented. An extension of Simulated Annealing was proposed and implemented methods were tested with respect to time, reliability and accuracy of the solution found. Influence of parameters of each algorithms and divergence criterions are also tested. Both, new index of elasticity and index of damage based on identified PM space of hysterons of hysteretic material by means of nonparametric methods were defined. Experimentally measured data were completely processed. The increase of damage of the material under the cyclic loading was evaluated.

Stochastic monitoring systems / **18**

Procedural modeling of buildings

Světlana Smrčková[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): smrcksve@fjfi.cvut.cz

Presentation of CGA shape grammar that is the grammar use to the procedural modeling of CG architecture. It can produce extensive architectural models for computer games, movies and etc. Description rules of grammar and contex sensitive rules with examples of their application. It will be showed basic construction of several building with details of theirs facades. The project deals with generation building in context with location in city in wievs of density of population, shape of the ground and etc. It will be showed problem of the generation in context especcialy problem with occlusion of object and its solving.

Data processing in HEP / **24**

Reconstruction of open charm mesons in relativistic heavy-ion collisions

Lukáš Kramárik[1]

[1] *Department of Physics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): kramaluk@fjfi.cvut.cz

ROOT is a framework for large-scale data analysis that provides basic and advanced statistical methods used by the high-energy physics experiments. This framework includes machine learning algorithms from the ROOT-integrated Toolkit for Multivariate Analysis (TMVA). TMVA package is becoming widely used for data reconstruction at STAR experiment in Brookhaven National Laboratory.

Especially, significance of open charm meson reconstruction could increase importantly with TMVA Boosted Decision Trees and Rectangular Cuts methods. These mesons are reconstructed via their hadronic decay channels, where the daughter particles can be tracked and identified with excellent precision by the STAR experiment at RHIC. Topological variables, such as decay length and distance of closest approach, of these mesons are used in TMVA training and classification of signal and background candidates.

Measurements of open charm meson production could give us information on quark-gluon plasma, hot and dense nuclear matter, created in ultrarelativistic heavy-ion collisions and expected to present in the Big Bang. These collisions are accessed at Relativistic Heavy Ion Collider in BNL and Large Hadron Collider in CERN.

Acoustic emission / **11**

Spectral analysis of rifle barrel harmonics

Milan Chlada[1]

[1] *Institute of Thermomechanics CAS, Prague, Czech Republic,*

Corresponding Author(s): chlada@it.cas.cz

Standard approaches to spectral analysis of discrete signals as Fast Fourier Transform (FFT) can turn out insufficient due to their limited frequency resolution caused by improper sampling period or the length of measuring time interval. In such cases, it is needed to compute the power spectral density function estimate for frequencies not available in standard Fourier Basis as well. Similarly to Fourier Transform, one possible way is to correlate measured data with a complex harmonic signal of arbitrarily set frequency. Generally, an increasing of frequency resolution can be smartly implemented by Continuous Wavelet Transform (CWT) enabling also the localization of various signal harmonic shapes in time. By means of above mentioned approaches it was possible to analyse the time envelopes of the composite barrel deformations that significantly affect the rifle accuracy. Such analysis of the barrel vibrations obtained during experimental measurements enables the selection of more suitable materials for further investigation of barrel dampening possibilities or frequency tuning.

Traffic and agent monitoring systems / **15**

Statistic resistivity of non-equilibrium states in transport gases

Nikola Groverová[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): grovenik@fjfi.cvut.cz

This presentation will focus on the analysis of non-equilibrium states in short-range transport gas. We will try to demonstrate and test, that a non-equilibrium system will have the same distribution of clearances as an equilibrium state, just with a different parameter. In this case we are working with a short-range transport gas with a logarithmic potential, which corresponds with the gamma distribution of clearances. First we will show, how would the distribution change provided, that our hypothesis is valid. We will discuss the variants of the system with random and equidistant initial distribution of particles. Afterwards we will derive the analytic formulation for the estimate of the time dependence of the parameter β. The last task is to verify the quality of the derived analytic formula using statistical methods. Using the bootstrap method we show, that the level of rejection is sufficiently low, therefore we don't reject the hypothesis mentioned above.

Data processing in HEP / **25**

Study of Weighted Kolmogorov-Smirnov homogeneity test's properties

Jakub Trusina[1]

[1] *Department of Mathematics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): jakub.trusina@fjfi.cvut.cz

In particle physics, homogeneity tests can be used to verify whether measured data's distribution corresponds to the distribution of a simulated Monte Carlo sample. However, Monte Carlo generator produces weighted samples because weights are used to modify the sample in order to take into account various attributes of a detector. Generalized homogeneity tests, such as Kolmogorov–Smirnov test, allow us to test the homogeneity of weighted samples. Several approaches to generalizations of this test are compared with generalized χ^2 test in three numerical experiments, which focus on an influence of weights' distribution on a probability of type-I error which is a key element of test's reliability.

Traffic and agent monitoring systems / **32**

TASEP with generalized update and Matrix Product Ansatz

Pavel Hrabák[1]

[1] *Department of Mathematics, FIT Czech Technical University in Prague*

Corresponding Author(s): pavel.hrabak@fit.cvut.cz

The steady-state distribution of Totally Asymmetric Simple Exclusion Process (TASEP) model has been studied by means of Matrix Product Ansatz (MPA) for variety of updates: random-sequential, parallel, backward-sequential. However, the newly developed generalized update has been studied very poorly. The contribution summarizes the principle of MPA and defines the equations following from the MPA concept. Two different solutions of the matrix algebra are presented: by Hrabák and Krbálek (2014) from unpublished PhD Thesis, and Brankov et. al (2017). Both concepts are compared and (dis)advantages and limitations discussed.

Data processing in HEP / **23**

TMVA optimization of KF Particle Finder

Michal Kocan[1]

[1] *Department of Physics, FNSPE Czech Technical University in Prague*

Corresponding Author(s): kocanmi3@fjfi.cvut.cz

In this presentation, I would like to introduce TMVA optimization of KF Particle Finder. This is a framework we are now using for reconstruction of short lives particles at the STAR experiment. For this, I will briefly describe STAR detector and the measured variables which we need for the particle reconstruction. Also the basics of the KF Particle Finder will be introduced. At the end of the presentation, the usage of the BDT method and the results gained with this method will be presented.

Acoustic emission / **34**

Time Reversal Localization of Continuous Acoustic Emission Sources

Zdeněk Převorovský[1] ; Milan Chlada[1] ; Michal Mracko[1] ; Josef Krofta[1] ; Jan Kober[1]

[1] *Institute of Thermomechanics CAS, Prague, Czech Republic,*

Corresponding Author(s): zp@it.cas.cz

Time reversal (TR) processing of acoustic and ultrasonic signals is effective tool for complicated problems solution in NDT /E and structural health monitoring. TR enables space-time focusing of elastic wave and thereby relatively easy location and partial reconstruction of both burst and continuous acoustic emission (AE) sources. AE source location problems come up in situations with high external noise, wave dispersion, and wave velocity changes in complex structures. Localization problems in such cases mostly require large sensor arrays or sophisticated signal processing and filtration.
A new solution using TR approach is discussed in this contribution. This approach allows planar

location of burst AE sources with only one transducer, and continuous AE sources with two transducers, both under high background noise from other sources. TR procedure is here applied to a long random noise signals.