

Lomonosov Moscow State University
(MSU)
Peoples' Friendship University of Russia
(RUDN University)

Analytical and Computational Methods in Probability Theory and its Applications



PROCEEDINGS

October 23–27, 2017
Moscow, Russia

Федеральное государственное образовательное учреждение высшего образования
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ М.В. ЛОМОНОСОВА»
Федеральное государственное автономное образовательное
учреждение высшего образования
«РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ»

LOMONOSOV MOSCOW STATE UNIVERSITY
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA (RUDN University)

**АНАЛИТИЧЕСКИЕ
И ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ
В ТЕОРИИ ВЕРОЯТНОСТЕЙ
И ЕЁ ПРИЛОЖЕНИЯХ
(АВМТВ-2017)**

Материалы Международной научной конференции

Россия, Москва, 23–27 октября 2017 г.

*Под общей редакцией
доктора физико-математических наук А.В. Лебедева*

**ANALYTICAL AND COMPUTATIONAL
METHODS IN PROBABILITY THEORY
AND ITS APPLICATIONS
(ACMPT-2017)**

Proceedings of the International Scientific Conference

23–27 October 2017, Moscow, Russia

Under the general editorship of D.Sc. A.V. Lebedev

**Москва
2017**

УДК 004.7:004.4.001:621.391:007(063)

ББК 32.073.202:32.968

A64

A63 Аналитические и вычислительные методы в теории вероятностей и её приложениях (АВМТВ-2017) = Analytical and Computational Methods in Probability Theory and its Applications (АСМРТ-2017) : материалы Международной научной конференции. Россия, Москва, 23–27 октября 2017 г. / под общ. ред. А. В. Лебедева. – Москва : РУДН, 2017. – 743 с. : ил.

В научном издании представлены материалы международной научной конференции «Аналитические и вычислительные методы в теории вероятностей и её приложениях» по следующим основным направлениям:

- Аналитические методы в теории вероятностей и её приложениях;
- Вычислительные методы в теории вероятностей и её приложениях;
- Асимптотические методы в анализе;
- История математики.

Сборник материалов конференции предназначен для научных работников и специалистов в области теории вероятностей и её приложений.

Текст воспроизводится в том виде, в котором представлен авторами.

Утверждено к печати Программным комитетом конференции

ISBN 978-5-209-08291-0

© Коллектив авторов, 2017

© Российский университет
дружбы народов, 2017

Sakhobov O. Tw
operation of the sy

Zaytseva O.B., F
semi-Markov queui

Kashtanov V.A.
queuing, reliability

Pavlov A.V. Ide
Laplace

Kanzitdinov S. K
networks with an in

Nazin A.V. Algo
Convex Optimizatic

Dimitrov M. Fluid

Efimov V.V., S
tomated operation,
network

Malyshkin Y. The
ential attachment m

Kalimulina E. Yu
the Unreliable Queu

Shukhman E. V.,
G.W. Leibniz Associ

Alyabieva V.G. T
torics in the 17th-19

Slepov N.A. Gene
tions.

Grishunina Yu., M
Based on Regenerati

Tsitovich I.I. On F

Fedotkin M.A., I
stationary distributic
conflict flows.

Limnios N. Diffusio

Contents

Sakhobov O. Two-sided bounds of the probability of failure-free operation of the system under the discipline of Soloviev	12
Zaytseva O.B., Kondrashova E.V. Priority management in a semi-Markov queuing model	21
Kashtanov V.A. Controlled stochastic processes and control in queuing, reliability and safety	26
Pavlov A. V. Identical service and the odd or even transform of Laplace	31
Kanzitdinov S. K., Vasilyev S. A. Analysis of random neural networks with an infinite number of cells	36
Nazin A. V. Algorithms of Inertial Mirror Descent in Stochastic Convex Optimization Problems	41
Dimitrov M. Fluid model with jumps in heavy traffic	46
Efimov V. V., Shchemelinin D. A. Control system for automated operation of a globally distributed telecommunications network	51
Malyshkin Y. The number of vertices of fixed degree in the preferential attachment model with choice	55
Kalimulina E. Yu. On the Rate of Convergence to stationarity of the Unreliable Queueing Network with Dynamic Routing	60
Shukhman E. V., Shukhman A. E. Unpublished Manuscripts by G.W. Leibniz Associated with Nondecimal Systems.	65
Alyabieva V. G. The idea of order in geometry, algebra, combinatorics in the 17th-19th centuries.	70
Slepov N. A. Generalized Stein equation on extended class of functions.	75
Grishunina Yu., Manita L. A Model for Network Virus Protection Based on Regenerative Process	80
Tsitovich I. I. On Robust Sequential Parameters Estimating	85
Fedotkin M. A., Kudryavtsev E. V. Necessary conditions for stationary distribution existence in the adaptive control system of conflict flows.	90
Limnios N. Diffusion Approximation of Branching Processes	95

Sorokovikova O.S., Dzama D.V., Asfandiyarov D.G., Blagodatskikh D.V. Probabilistic models of contamination in large water areas. Statistics and stochastic algorithms	100	Veretennikov A.Yu. C istence and uniqueness
Chinenova V. Goryachkin and His Idea in the Agricultural mechanics: Human - Machine - Environment	105	Smirnova G.S. Mos Conference
Bashtova E.E., Chernavskaya E.A. Limit theorems for infinite-channel queueing systems with heavy-tailed service times	110	Kondrat'yev A.I. Co mpression
Shklyaev A.V. Large Deviation Probabilities for the Branching Process in Random Environment'	113	Molchanov S., Zhang Turing's Formula
Grishunina S.A. Limit theorems for queueing systems with different service disciplines	119	Konakov V., Menozz ion processes on solval quasi-local limit theorer
Kozlov V.V. On limit theorems of A. D. Soloviev for regenerative processes	124	Shevtsova I., Gabdul gence rate estimates in
Kolesnikov S.N. Different formalism of classical mechanics	129	Kolnogorov A.V. An Two-Armed Bandit
Simonov R.A. Kirik the Novgorodian – a mathematician of the Early Renaissance.	134	Gribkova N.V. On p for L -statistics: a survey
Butko Ya.A. Chernoff Approximation of transition kernels of Markov processes	139	Glazunov N.M. Ari lands correspondence.
Bulinskaya E.VI. The Particles Population Propagation in Catalytic Branching Random Walk	144	Farkhadov M.P., Po comparative analysis ar for managing speech dis
Foss Sergey Non-standard randomized multiple access transmission protocols: stability and optimization	147	Ivnitskiy V.A. Abou fect and different intera
Vatutin V.A. Branching processes in random environment	148	Pastukhova Yu.I. A) constructing asymptotic functionals from the reg
Abramov O.V. The reliability for gradual failure: functional-parametrical approach	150	Burnaev E.V., Golu Signal Detection Proble
Grusho A.A., Timonina E.E., Shorgin S.Ya. Probabilistic Models of Meta Data.	156	Savelov M.P. Limit c homogeneous polynomi
Rusakov O.V. From the Pseudo-Poisson Processes with the Random Intensity to the Fractional Brownian Motion	161	Bochkov A.V., Ya Creation and Developm
Basharin G.P., Naumov V.A. Markovian Modelling of Arrival Processes	166	Tikhomirov V.M. G and 60th years).
Soloviev A.D., Sobolev V.N. One server queue with bulk arrivals	171	Demidov S.S., Pet torrant of the Soviet hi
Kolesnikov S.N. Scientific revolution of 16–17 centuries and period of “mechanism” in science	176	

Veretennikov A. Yu. On mean-field $GI/GI/1$ queueing model: existence and uniqueness	182
Smirnova G. S. Moscow 1935: the First International Topology Conference	187
Kondrat'yev A. I. Computational models of diamond anvil cell compression	192
Molchanov S., Zhang Zh., Zheng L. Central Limit Theorem of Turing's Formula	198
Konakov V., Menozzi S., Molchanov S. Approximation of diffusion processes on solvable Lie groups by random walks. Local and quasi-local limit theorems	202
Shevtsova I., Gabdullin R., Makarenko V. On natural convergence rate estimates in the Lindeberg theorem	207
Kolmogorov A. V. An Asymptotic Minimax Theorem for Gaussian Two-Armed Bandit	210
Gribkova N. V. On probabilities of large and moderate deviations for L -statistics: a survey of some recent developments	215
Glazunov N. M. Arithmetic Statistics, Probabilities and Langlands correspondence.	220
Farkhadov M. P., Petukhova N. V. Probabilistic methods for comparative analysis and optimization of scenarios and algorithms for managing speech dialogue.	226
Ivnitskiy V. A. About the thinning of a flow with limited aftereffect and different interarrivals distributions.	234
Pastukhova Yu. I. Applying the maximum likelihood method for constructing asymptotically effective nonparametrical estimators of functionals from the regression function	239
Burnaev E. V., Golubev G. K. Bayesian Test for Multi-Channel Signal Detection Problem	244
Savelov M. P. Limit distributions of the Pearson statistics for non-homogeneous polynomial scheme	249
Bochkov A. V., Yastrebenetsky M. A. The History of the Creation and Development of the e-Forum Gnedenko	252
Tikhomirov V. M. Golden periods of Moscow mathematics (30th and 60th years).	258
Demidov S. S., Petrova S. S. Mathematics in Moscow in the torrent of the Soviet history	260

Veretennikov A. Yu. On mean-field $GI/GI/1$ queueing model: existence and uniqueness	182
Smirnova G. S. Moscow 1935: the First International Topology Conference	187
Kondrat'yev A. I. Computational models of diamond anvil cell compression	192
Molchanov S., Zhang Zh., Zheng L. Central Limit Theorem of Turing's Formula	198
Konakov V., Menozzi S., Molchanov S. Approximation of diffusion processes on solvable Lie groups by random walks. Local and quasi-local limit theorems	202
Shevtsova I., Gabdullin R., Makarenko V. On natural convergence rate estimates in the Lindeberg theorem	207
Kolnogorov A. V. An Asymptotic Minimax Theorem for Gaussian Two-Armed Bandit	210
Gribkova N. V. On probabilities of large and moderate deviations for L -statistics: a survey of some recent developments	215
Glazunov N. M. Arithmetic Statistics, Probabilities and Langlands correspondence.	220
Farkhadov M. P., Petukhova N. V. Probabilistic methods for comparative analysis and optimization of scenarios and algorithms for managing speech dialogue.	226
Ivnitskiy V. A. About the thinning of a flow with limited aftereffect and different interarrivals distributions.	234
Pastukhova Yu. I. Applying the maximum likelihood method for constructing asymptotically effective nonparametrical estimators of functionals from the regression function	239
Burnaev E. V., Golubev G. K. Bayesian Test for Multi-Channel Signal Detection Problem	244
Savelov M. P. Limit distributions of the Pearson statistics for non-homogeneous polynomial scheme	249
Bochkov A. V., Yastrebenetsky M. A. The History of the Creation and Development of the e-Forum Gnedenko	252
Tikhomirov V. M. Golden periods of Moscow mathematics (30th and 60th years).	258
Demidov S. S., Petrova S. S. Mathematics in Moscow in the torrent of the Soviet history	260

- Turzynsky M. K.** On the properties of the system of gas dynamics equations in Lagrangian coordinates with Coriolis force 265
- Timonin V. I., Tyannikova N. D.** The Kolmogorov – Smirnov type tests for the Lehmann – Cox hypotheses in the case of progressively censored samples — about the possibility to use the Kaplan – Meier estimates in test statistics 270
- Chernousova E., Molchanov S.** Steady state for the critical branching random walk with the general number of offsprings 276
- Han D., Makarova Yu., Molchanov S., Yarovaya E.** Branching Random Walks with Immigration 281
- Savinkina E. N., Sakhanenko A. I.** Asymptotic Explicit Optimal Estimators of an Unknown Parameter in One Power Regression Problem 286
- Grebeshkov A. Yu., Zaripova E. R., Pshenichnikov A. P.** Vertical Handover Time Estimation Method 290
- Martynov G. V.** Anderson-Darling and New Weighted Cramér-von Mises Statistics 295
- Mishkoy Gh. K.** Priority systems with orientation. Analytical and numerical results 301
- Kozlov M. V.** On the asymptotics of large deviation probabilities for some partly or asymptotically homogeneous Markov chains 306
- Boldin M. V., Petriev M. N.** Robust GM-estimators in the autoregression and Pearson's chi-square tests 309
- Moiseev A. N., Shklennik M. A.** Heterogeneous infinite-server queueing tandem with customers' type defined by state of Markovian arrival process 314
- Melezhik V. S.** Quantum-semiclassical approach for treating excitation and ionization of hydrogen-like atoms and ions in collisions with heavy particles 319
- Senatov V. V.** On numerical studies of the accuracy of approximations in the central limit theorem 324
- Alawadhi S. A.** A probability model for Assessments of system loads 331
- Kolev N., Ngo N., Ju Y. T.** Bivariate Teissier Distribution 336
- Antonyuk P. N.** A brief history of the study of the fragment size distribution 341
- Khmourkin G. G.** W. tions on the introductory A.D.)
- Nazarov A. A., Brone** with On/Off control and values
- Krishnamoorthy A., S** multi-server priority queu
- Krishnamoorthy A.,** queueing-inventory syste customers.
- Kovalenko I. N.** ACMI
- Koroliouk D., Koroliu** with random change of ti
- Vvedenskaya N. D., L** bartsev A. A. Local la Markov processes
- Veretennikova M., Sik** dicting neuro-developmer to cerebral malaria
- Zaryadov I. S., Kradev** ysis of Cloud Computing Servers and a Single Buff
- Thorisson H.** On Coup Distribution
- Lemeshko B. Yu., Le** Novikova A. Yu. Appli Solution
- Lotov V. I.** Factorizati for random walks
- Kelbert M., Suhov Y.,** properties.
- Zverkina G. A.** Industr
- Malyshev V. A., Zan** on Z^d
- Kuzicheva Z. A.** Implic of the XX century.

265	Khmourkin G.G. Why do Indians need mathematics? (Reflections on the introductory chapter of Mahavira's treatise, 9th century A.D.)	347
270	Nazarov A.A., Broner V.I. Modified Kramer-Lundberg model with On/Off control and arbitrary distribution of demand purchases values	350
276	Krishnamoorthy A., Shajin Dhanya, Manjunath A.S. On a multi-server priority queue with preemption in crowdsourcing	355
281	Krishnamoorthy A., Shajin Dhanya MAP/PH/1 retrial queueing-inventory system with orbital search and renegeing of customers	360
286	Kovalenko I.N. ACMPT-2017 Conference Thesis Template	365
290	Koroliouk D., Koroliuk V.S. Adapted statistical experiments with random change of time	370
295	Vvedenskaya N.D., Logachev A.V., Suhov Yu.M., Yambartsev A.A. Local large deviation principle for inhomogeneous Markov processes	375
301	Veretennikova M., Sikorskii A., Boivin M. Data mining in predicting neuro-developmental scores from EEG data during coma due to cerebral malaria	380
306	Zaryadov I.S., Kradenyh A.A., Gorbunova A.V. The Analysis of Cloud Computing System as a Queueing System with Several Servers and a Single Buffer	384
309	Thorisson H. On Coupling and Convergence in Density and in Distribution	389
314	Lemeshko B.Yu., Lemeshko S.B., Veretel'nikova I.V., Novikova A.Yu. Application of Homogeneity Tests: Problems and Solution	395
319	Lotov V.I. Factorization method in boundary crossing problems for random walks	400
324	Kelbert M., Suhov Y., Stuhl I. Weighted entropy: basic facts and properties	406
331	Zverkina G.A. Industrial revolution and reform of mathematics	411
336	Malyshev V.A., Zamyatin A.A. One-particle quantum walk on Z^d	416
341	Kuzicheva Z.A. Implication in the works of logics of the first half of the XX century	420

Lyuter I.O. Geometrization of the doctrines of place and continuum in medieval Arabic scholasticism.	425
Andronov A.M., Rykov V.V., Vishnevsky V.M. On Reliability Function of a Parallel System with Three Renewable Components.	430
Bulinskaya E.V. Stability problems in modern actuarial sciences.	435
Singpurwalla Nozer D. Subjective Probability: Its Axioms and Operationalization.	440
Konstantinides D.G., Li J.Z. Asymptotic Ruin Probabilities for a Multidimensional Renewal Risk Model with Multivariate Regularly Varying Claims.	446
Ludkovsky S.V. Feynman-type local integration of stochastic PDE.	453
Zolotukhin I.V. On Multivariate Geometric Random Sums.	456
Zeifman A., Satin Ya., Kiseleva K., Korolev V. Two-sided Bounds for the Convergence Rate of Markov Chains.	462
Pakhteev A.I. Some Algorithms of Record Generation.	467
Sokoł M. Markov stochastic processes in biology — almost the same than in mathematics but a bit different.	471
Zaytsev E. Mathematical Methods in Practical Mechanics: From Heron of Alexandria to Galileo.	475
Belopolskaya Ya.I., Stepanova A.O. Probabilistic algorithms of constructing numerical solution of the Cauchy problem for systems of nonlinear parabolic equations.	480
Belkina T.A., Konyukhova N.B., Slavko B.V. Analytic-Numerical Investigation of the Dual Risk Model with Investments: Survival Probability Functions as the Solutions of Singular Problems for Integro-Differential Equations.	485
Zubkov A.M., Filina M.V. Algorithm of exact computation of divisible statistics distributions and its applications.	490
Mogulskii A.A. Integro-local limit theorems for multidimensional compound renewal processes.	495
Koldanov P.A., Koldanov A.P., Kalyagin V.A., Pardalos P.M. Optimal test of conditional independence testing in multivariate normal distribution.	500
Farkas Yu., Hashorva E., Piterbarg V.I. Asymptotic behavior of reliability function for multidimensional aggregated Weibull type reliability indices.	506

Isguder H.Okan, K Server Queueing System.
Formanov Sh.K. On theorem.
Yakymiv A.L. On so and their applications t.
Zverkina G.A. On a.
Kobanenko G.K. I cesses.
Borodina A.B., Ef Splitting Method for I tion System.
Sopin E.S. On the in with resources station: workload volume.
Anulova S.V. Fluid I eral Multi-Servers.
Kharin Yu.S. Statis nious models of high-o.
Gadolina I.V., Li building confidence int quality control applicat.
Fedorova E.A., Na: probability distribution.
Afanaseva L.G. Asy on Synchronization Me.
Muromskaya A.A. with several risks with.
Bulinski A.V. Asym.
Sakhanenko A.I. N general random walks.
Zorine A.V. A stop process of Poisson inpu.
Mastikhin A.V. Inte for Weiss multidimentic.

425	Isguder H. Okan, Kocer U. U. Analysis of Two-Heterogeneous Server Queueing System	512
430	Formanov Sh. K. On nonclassical versions of the Lindeberg-Feller theorem	517
435	Yakymiv A. L. On some recent investigations on Tauberian theory and their applications to probability theory	522
440	Zverkina G. A. On a generalization of the Leibniz theorem	528
446	Kobanenko G. K. Limit theorems for bounded branching processes	533
453	Borodina A. B., Efrosinin D. V., Morozov E. V. Adaptive Splitting Method for Failure Estimation in Controllable Degradation System	538
456	Sopin E. S. On the insensitivity conditions of the queuing system with resources stationary distribution to distribution of customer workload volume	541
462	Anulova S. V. Fluid Limit for Closed Queueing Network with Several Multi-Servers	547
467	Khariu Yu. S. Statistical analysis of big data based on parsimonious models of high-order Markov chains	552
471	Gadolina I. V., Lisachenko N. G. Bootstrap-modelling for building confidence intervals for percentiles: high-tech production quality control application	555
475	Fedorova E. A., Nazarov A. A., Paul S. V. Discrete gamma probability distribution approximation in retrial queues	560
480	Afanaseva L. G. Asymptotic Analysis of Queueing Models based on Synchronization Method	566
485	Muromskaya A. A. Optimal reinsurance strategy in the model with several risks within one insurance policy	571
490	Bulinski A. V. Asymptotic Methods and Limit Theorems	576
495	Sakhanenko A. I. Non-classical boundary crossing problems for general random walks	581
500	Zorine A. V. A stopped random walk and stability of a service process of Poisson input flows by a loop algorithm	586
506	Mastikhin A. V. Integral representation of transition probabilities for Weiss multidimensional epidemic process	591

Tarasenko A. S. Moments of the sojourn time of random walk above a certain boundary	595
Zubkov A. M., Orlov O. P. Limit distributions of extreme distances to the nearest neighbor	598
Ermishkina E. M. Simulation of Branching Random Walks on Multidimensional Lattices	603
Yarovaya E. Survival Analysis and Recurrence Criteria for Branching Random Walks	608
Zverkina G. A. Simple bounds for the convergence rate of $M G _\infty$ queueing system	613
Belyaev Yu. K., Källberg D., Rydén P. Statistical Analysis of Data Generated by a Mixture of Two Parametric Distributions	619
Popov G., Yarovaya E. Non-Markovian Models of Branching Random Walks	624
Rachinskaya M. A., Fedotkin M. A. Stationarity conditions for the control systems that provide service to the conflicting non-ordinary Poisson flows	629
Mashnikov O. V., Zamyatin A. A. An ergodicity criterion for Markov chains describing the dynamics of random words.	634
Berezin S., Zayats O. Skew Brownian motion with dry friction: The Pugachev–Sveshnikov equation approach	639
Manita A. Joint distributions of synchronization models	644
Manita L., Manita A. Agreement algorithms for synchronization of clocks in nodes of stochastic networks.	649
Zhdanok A. I. Dimension of the set of invariant finite additive measures of Markov chains in an arbitrary phase space and ergodic consequences.	654
Oprisan A. Limit theorems for Additive Functionals of Semi-Markov Processes	659
Lukashenko O. V., Morozov E. V., Khokhlov Y. S. On the asymptotic bound for the overflow probability of fluid queue with heterogenous input	665
Knobloch E. Leibniz's Contributions to Financial and Insurance Mathematics.	669
Balashova D. Numerical analysis of phase transitions in supercritical branching random walks	674

Domoradzki S. On ma the 19th century till the
Rusev V. N., Skorikov determination Weibull-G
Soós A., Somogyi I., S iterated function system
Rytova A. I. Harmonic
Zolotukhin A. Ya., N culating the absolute co two-point distributions .
Kumar R., Sharma S. ing Model with Discou Customers
Makarichev A. V., B The System Refuse Inter out Waiting And Eleme Into The System With M
Makarichev A. V., M Complexes with the Rest Elements with Minimal
Makarichev A. V. Ass with Minimum Service
Kolosova I. S.; Vasily Schrödinger Equation w
Kalinkin A. V., Pavl «Fundamentals of the m nical university
Rykov V. V., Kozyr tems by Markovization
Zubkov A. M., Kru marked embeddings of marked vertices.

Author index

595
598
603
608
613
619
624
629
634
639
644
649
654
659
665
669
674

Domoradzki S. On mathematics in Lvov from the second half of the 19th century till the end of WWII 678

Rusev V. N., Skorikov A. V. Analytical and discrete methods for determination Weibull-Gnedenko renewal density 681

Soós A., Somogyi I., Simon L. Interpolation using stochastic local iterated function systems 686

Rytova A. I. Harmonic Analysis of Random Walks on Lattices . . . 689

Zolotukhin A. Ya., Nagaev S. V., Chebotarev V. I. On calculating the absolute constant in the Berry–Esseen inequality for two-point distributions 695

Kumar R., Sharma S. Transient Analysis of a Multi-Server Queuing Model with Discouraged Arrivals and Retention of Reneging Customers 700

Makarichev A. V., Brysina I. V. The Asymptotic Estimation The System Refuse Intensively When Elements Are Repaired Without Waiting And Elements That Have Reach The Repair Return Into The System With Minimal Standby Beginning Off any Quantity 705

Makarichev A. V., Makarichev V. A. Reliability Of Systems in Complexes with the Restoration and the Absolute Priority of Service Elements with Minimal Reserve Systems Under Heavy Load 708

Makarichev A. V. Assessment of the Probability of System Failure with Minimum Service Accumulation Elements 711

Kolosova I. S., Vasilyev S. A. Numerical Solving of Relativistic Schrödinger Equation with Random Quasipotential. 718

Kalinkin A. V., Pavlov I. V., Sidnyaev N. I. Special course «Fundamentals of the mathematical theory of reliability» in the technical university 723

Rykov V. V., Kozyrev D. Analysis of renewable reliability systems by Markovization method 727

Zubkov A. M., Kruglov V. I. Number of pairs of identically marked embeddings of given subtree in q -ary tree with randomly marked vertices. 735

Author index 741

Data mining in predicting neuro-developmental scores from EEG data during coma due to cerebral malaria

M. Veretennikova*, A. Sikorskii†, M. Boivin‡

* *Department of Statistics and Data Analysis, Faculty of Economics, National Research University Higher School of Economics, Ul. Shabolovka, 28, 119049, Moscow, Russia*

† *Department of Statistics and Probability, Michigan State University, 619 Red Cedar Rd C413, East Lansing, MI 48824, USA*

‡ *Department of Neurology and Ophthalmology, Michigan State University, East Fee Hall, 965 Fee Road, East Lansing, MI 48824, USA*

Abstract. In this research we compare the performance of different data mining techniques in the analysis of electroencephalogram (EEG) data. We study the question of predicting post-comatose neuro-developmental scores based mainly on statistical features from the EEG recordings. We compare results from applying different data mining techniques, such as the Elastic Net, Lasso, Gaussian Support Vector Regression and Random Forest Regression. We also compare the results produced with different matrix completion methods.

Keywords: data mining, regression, regularization, random forests, matrix completion, EEG, Daubechies wavelets.

1. Introduction

Modern technologies used for statistical analysis of the brain include EEG, MEG, PET, fMRI and optical imaging. EEG is often used for monitoring patients with seizures and epilepsy. Seizures, along with coma, convulsions, and metabolic disturbances are common complications of cerebral malaria (CM) [1], a tropical disease affecting over half a million people annually, mostly in sub-Saharan Africa. Different sources indicate distinct mortality rates, ranging from 4 to 46 percent. Possible mechanisms of brain injury during CM are studied in [1].

In [3] the authors predict neuro-developmental outcomes for term infants with hypoxic ischaemic encephalopathy (HIE) from EEG features such as flat trace, burst suppression and low voltage. The authors highlight that the reason for such a data mining problem is to accurately identify children in need for neuro-restorative therapy. For instance in [4] the authors present the literature review about the relation of between seizures in neonates and development of neuro-developmental complications such as epilepsy, intellectual impairment and cerebral palsy.

In this research we analyze EEG data children in coma due to cerebral malaria. As it is noted in [2] a large proportion of children who recover

o-developmental scores ue to cerebral malaria

rskii[†], M. Boivin[‡]

*Faculty of Economics, National
ool of Economics,
Moscow, Russia*
*nd Probability,
ersity,
sing, MI 48824, USA*
*d Ophthalmology,
ersity,
ansing, MI 48824, USA*

formance of different data mining
gram (EEG) data. We study the
velopmental scores based mainly
t. We compare results from apply-
the Elastic Net, Lasso, Gaussian
st Regression. We also compare
pletion methods.

ation, random forests, matrix com-

ion

l analysis of the brain includ-
g. EEG is often used for moni-
seizures, along with coma, com-
mon complications of cerebral
g over half a million people ac-
ferent sources indicate distinct
cent. Possible mechanisms of

mental outcomes for term in-
thy (HIE) from EEG features
ow voltage. The authors high-
g problem is to accurately iden-
therapy. For instance in [4] the
the relation of between seizures
elopmental complications such
erebral palsy.

ildren in coma due to cerebral
ortion of children who recover

from cerebral malaria have neurological consequences, some even devel-
oping difficulties in cognition and behaviour [2]. However specific mecha-
nisms leading to neuro-developmental deficits have not been established,
hence there is a need to use data mining techniques to identify impor-
tant features, which can then be further researched by neurologists. We
pose the question of predicting post-comatose neuro-developmental scores
based mainly on features from the EEG recordings. The idea is to identify
the most important features for identifying children with the highest risk
of experiencing neuro-developmental problems and to look into usefulness
of EEG statistical features for this prediction problem.

2. Main section

In this research we compare the prediction results using different data
mining techniques, such as the Elastic Net, Lasso, Gaussian Support Vec-
tor Regression and Random Forest Regression. We also compare the re-
sults produced with different matrix completion methods.

Our dataset comprises of the standard 10-20 EEG recordings for 78
patients with the sampling rate of 500 Hz and the average record duration
of 30 minutes. Artefacts due to breath, muscle movement and heart beat
were removed from the raw data using Persyst software based on a neural
network algorithm. We chose to use Daubechies (Db) wavelets for splitting
the clean signal into frequency bands.

Different research indicates particular suitability of Db4 for statistical
EEG analysis and we checked the relative average MSE errors between the
wavelet signal approximation and the actual signal for different Daubechies
wavelets for the data. Our results indicated that Db4 yields an MSE or
order 10^{-9} , which is reasonably low. Also, Db4 frequency band separation
results in frequency intervals which are very close to the traditional ranges
for delta, theta, beta, alpha and gamma bands.

The features include amplitude variances in delta, theta, alpha, beta
and gamma frequency bands, Shannon entropies, relative frequency band
energy, proportion of flat line EEG, presence of seizures as a binary vari-
able, frequencies of peaks in the original cleaned time series differing from
the nearest measurements from both sides by 1/3, 1, 2 and 3 standard
deviations (we will denote these by FP 1/3, 1, 2 and 3 for simplicity),
complexity and mobility of the time series averaged over all considered
channels. Other features included: height, weight, the Blantyre coma
score, age, hemoglobin base level and economic home scores.

We have used the Elastic Net and Lasso, combined with PCA methods,
Random Forest Regression (RFR) and regularized Support Vector Regres-
sion. To begin with we considered the feature matrix with 54 non-EEG
features complementing the 362 statistical EEG features. This meant we
had to complete 504 missing matrix entries. We have used different meth-
ods to complete the matrix, including the SoftImpute method based on
Singular Value Decomposition considerations with regularization and a
nonparametric method missForest based on the random forest technique.

We also looked at how changing the structure of the feature matrix affects the prediction results.

We have gone checked a grid of regularization parameter λ values in the minimization problem which is the base of the matrix completion method called SoftImpute. The best results were obtained for $\lambda = 100$ and here we present some details.

Running the Elastic Net yielded 27 non-zero coefficients with $MSE = 0.37934$. All coefficients from the FP 2 and 1 come with the positive sign, so do weight, height and hemoglobin base. The top 13 coefficients in absolute value are all positive and are above 50, the top 12 are from FP 2 and FP 1 groups, weight being the 13th largest coefficient in absolute value.

We have experimented with the structure of the matrix, for example doing random selection of a channel from each block of features of the same nature, removing all medical features and all EEG features in turn, changing some of the feature values to the column means to detect importance of the feature in the prediction. Here we present some of the outcomes.

Removing all EEG or all medical features made the Elastic Net zero every single coefficient, but the intercept, which signals that it is beneficial to use both types of feature for prediction purposes. Removing all patients who were identified as having a seizure during the EEG recording made the prediction by the Elastic Net much worse.

Using random forest regression (RFR) with 1000 trees and bootstrapping yields a very low out-of-bag error of 0.015 and the MSE is 0.1208. We have implemented regularized support vector regression with the RBF kernel, going over a grid of penalty C and kernel γ coefficients, but cross validation showed that this method is inferior to the results produced by Random Forest Regression and the Elastic Net.

Using a non-parametric method for matrix completion yielded results similar and slightly better in accuracy to most of the outcomes using the Elastic Net with SoftImpute, the MSE being 0.55709 with 68 non-zero coefficients. However, the result didn't surpass the prediction quality of the best result with the Elastic Net following SoftImpute matrix completion.

3. Conclusions

We conclude that EEG features bring significant value in prediction of neuro-developmental scores of children after their awaking from coma due to cerebral malaria. In particular we identify several potentially most useful biomarkers of EEG nature for this specific prediction problem. These include the frequency of spikes higher than 1 and 2 standard deviations from their nearest neighbours in time, relative wave energies for different frequency bands and variance in the theta frequency band. Details and relation to other research will be presented at the conference.

We wil
pletion me
marker

The w
No. 17-11

1. Idro R
Strateg
2010. -
2. World
Intern:
3. Awal 1
backgr
poxic i
rophys
4. Pisani
neonat
Neurol
5. Mazur
rithms
Learn
6. Stekha
value
Vol. 2:

We will continue investigating the effect of using different matrix completion methods and in the most effective ways to identify efficient EEG biomarkers for such prediction problems.

Acknowledgments

The work is partially funded by the Russian Science Foundation (project No. 17-11-01098).

References

1. *Idro R. et al.* Cerebral Malaria; Mechanisms of Brain Injury And Strategies For Improved Neuro-Cognitive Outcome // *Pediatr. Res.* — 2010. — Vol. 68, no. 4. — P. 267–274.
2. *World Health Organization.* Severe Malaria // *Tropical Medicine and International Health.* — 2014. — Vol. 19, Suppl.
3. *Awal Md. A., Lai M. M., Azemi G., Boashash B., Colditz P. B.* EEG background features that predict outcome in term neonates with hypoxic ischaemic encephalopathy: A structured review // *Clinical Neurophysiology.* — 2016. — Vol. 125, no. 1. — P. 285–296.
4. *Pisani F., Facini C., Pavlidis E., Spagnoli C., Boylan G.* Epilepsy after neonatal seizures: Literature review // *European Journal of Paediatric Neurology.* — 2015. — Vol. 19. — P. 6–14.
5. *Mazumder R., Hastie T., Tibshirani R.* Spectral Regularization Algorithms for Learning Large Incomplete Matrices // *Journal of Machine Learning Research.* — 2010. — Vol. 11. — P. 2287–2322.
6. *Stekhoven D. J., Buhlmann P.* MissForest - non-parametric missing value imputation for mixed-type data // *Bioinformatics.* — 2012. — Vol. 28, no. 1. — P. 112–118.