Sergey Voronin, Ph.D. | sergey.voronin@outlook.com | svoronin.neocities.org | www.linkedin.com/in/sergey-voronin-36566a211/ | U.S. Citizen

Background in applied and computational mathematics, large-scale data analysis, parallel computing, and machine learning. M.A., Ph.D., and 12+ years of work in academia and industry. Experience as a researcher & developer and leading small teams. Developed multiple computational solutions with a mix of statistical and machine learning based approaches for different applications. Successfully prepared grant proposals and led several funded project efforts. Experience with C/C++/Python/R and Linux-based tools. Multiple publications (www.researchgate.net/profile/Sergey-Voronin-4) and open source software (www.github.com/sergeyvoronin).

EDUCATION

Princeton University, Princeton, NJ. **M.A.** & **Ph.D.**, Applied and Computational Mathematics, Oct. 2009, Nov. 2012. Developed and mathematically analyzed novel iterative algorithms for solving large matrix-based optimization problems with sparsity constraints, applying to Compressive Sensing, Imaging and Geophysical inverse problems.

Fu Foundation School of Engineering and Applied Science, Columbia University, New York, NY. **B.S.**, Applied Mathematics major with Computer Science minor, May 2007, GPA: 3.67.

International STEM focused education.

SOFTWARE DEVELOPMENT & DATA ANALYSIS EXPERIENCE

- Significant development experience using C / C++, Python, R, Java, Perl/bash scripting, SQL and Linux dev tools.
- Statistical data analysis, anomaly detection and visualization experience on many different data types, including time series, text documents, imagery, audio/video, and various scientific data. Work with different statistical and signal processing methods and numerical algorithms (integration, optimization, Fourier/Wavelet transforms, time series analysis and forecasting, etc.).
- AI/ML development with Scikit-learn, Keras, PyTorch in Python, Java Weka, and other major libraries. NLP for text analysis and classification tasks. Outlier discovery, missing data imputation, and anomaly detection (e.g. Gaussian Process Regression, Isolation Forest). Implementations with ARIMAX, GPR, LSTM, Transformers for time series analysis and prediction problems.
- Developed several open-source software packages, including RSVDPACK (randomized algorithms for matrix factorization and dimensionality reduction), Sparse Optimization Pack (iterative algorithms for sparsity-inducing optimization), Multi-Resolution Audio Classification (wavelet-based decomposition and ensemble scheme for audio classification), PBWTb_compressor (new prototype parallel lossless compressor).
- P-threads/OpenMP/MPI shared/distributed memory parallel programming experience and GPU computing with CUDA.

PROFESSIONAL EXPERIENCE (ACADEMIA & INDUSTRY)

PERSONAL PROJECTS, FAIRFAX, VA. NOV. 2024 – PRESENT.

- Implemented multivariate time series predictors using VAR, ARIMAX, MLP, LightGBM, and LSTM methods with parameter optimization and technical indicator computations and analysis for financial and weather data using Python and R libraries.
- Worked on classification and regression challenges, implementing a multi-label per-instance text document classification engine with NLP using NLTK and Transformer models with Scikit-learn and PyTorch libraries in Python and a regression engine for initializing weather prediction models using Gaussian Process Regression (GPR).
- Performed model parameter tuning and evaluated various performance metrics for different ensemble models.
- Researched and implemented image segmentation and OCR-based text extraction and analysis from food product label images with Python computer vision libraries.

INTEL CORPORATION, FAIRFAX, VA.

RESEARCH SCIENTIST / SOFTWARE ENGINEER: SOFTWARE AND ADVANCED TECHNOLOGY GROUP, APRIL 2022 – NOV 2024.

- Developed a streaming network data analysis system with an FPGA device with a low level Python and C/C++ implementation coupled with an ML-based solution. Network statistics and select packets are collected with the device, analyzed with statistics and autoencoder-based feature mining, and anomalies reported for measurement segments (e.g., 5 min), based on, e.g., latency changes, DNS connections, or packet content information. Auto generated dynamic HTML/JavaScript environment for reviewing network data statistics over the collection period and identified changepoints and anomalies.
- Developed parallelized software for data compression, multi-variate time series analysis and prediction, and parallel sorting. This included the development of a multi-core Burrows Wheeler transform implementation with Arithmetic coding and byte value histogram clustering for a new parallel lossless compressor prototype, published as open source package, statistical and iterative machine learning based predictors for multi-variate series data (e.g. GPR with R, multi-stage LSTM implementation with Keras in Python) applied to heat profiling problems, and a parallel merge sort implementation in C++ for sorting large samples from the unified host and network data sets with custom implementations and STL functions.

INTELLIGENT AUTOMATION, INC. ROCKVILLE, MD.

SR. SCIENTIST / Project Lead. JUNE 2017 – APRIL 2022.

- Applied signal processing and AI/ML methods to different applications focusing on time series, audio/video analysis, compression, and anomaly detection tasks. Created multiple from scratch implementations for concept demo purposes. Utilized container architectures and deployed models on computational clouds.
- Developed algorithmic and computational solutions (with C++, Python, Java) in data compression, audio and imagery analysis, anomaly detection and localization, position estimation from intermittent navigation and telemetry sensors, and machine learning based classification and regression implementations with different data types. Performed audio processing with denoising schemes and utilized diffusion models for audio speaker segmentation and multi-resolution methods for audio classification. Worked with hydrophone array audio data and acoustic simulations to design a similarity based compression scheme for satellite transmission. Utilized multi-core and GPU acceleration (e.g. parallel kernel evaluations, parameter optimization, and model training).
- Authored multiple technical proposals and won several SBIR and STTR grants on data compression, acoustics, antenna array systems, electrical fault analysis from DOE (FY '21 SBIR award), MDA (17-004), ARMY (18B-T009), NAVY (181-067, 192097, 211-003), Air Force (AF19C-T009) and others (please see https://www.sbir.gov/awards). Led awarded projects as Principal Investigator with small teams and prepared technical reports and meeting materials. Presented work updates and deliverables to the contracting office and prepared working software demos for onsite presentations.

TUFTS UNIVERSITY, MEDFORD, MA; **UNIVERSITY OF COLORADO BOULDER**, BOULDER, CO; **UNIVERSITY OF NICE SOPHIA ANTIPOLIS**, FRANCE.

NORBERT WIENER ASSISTANT PROFESSOR; POSTDOCTORAL RESEARCH ASSOCIATE / INSTRUCTOR, OCTOBER 2012 – JUNE 2017.

- Designed and analyzed novel numerical linear algebra algorithms with randomization for efficiently computing large-scale matrix factorizations (low rank ID/CUR/SVD) for data mining, feature extraction, and classification. Created RSVDPACK high-performance matrix factorization package written in C with shared memory CPU and GPU parallelism.
- Taught University courses in statistics, high-performance scientific computing, differential equations, and linear algebra.
- Developed compression and outlier detection techniques for large-scale Geophysical data and used cluster-based visualization.
- Created tools using MPI & OpenMP with C and C++, with developed blocking and projection techniques, for performing large-scale optimization-based seismic inversion using complex constraints. Implemented Wavelet transforms on a cubed sphere grid for Earth models. Constructed high resolution slowness (density) models for Earth's interior structure.

UNIVERSITY TEACHING / PRESENTATION EXPERIENCE

Multiple courses and seminars given at Universities. Full semester courses taught to undergraduate and graduate students, including Intro Statistics with R (2 semesters), Differential Equations and Linear Algebra (3 semesters), Calculus 2 and 3 (2 semesters), High Performance Scientific Computing (1 semester). Designed and led seminars on the use of cluster computing systems, data analysis, and machine learning tools. Prepared and oversaw multiple technical presentations and software demos to project sponsors for SBIR/STTR projects.

SELECT PUBLICATIONS AUTHORED (additional content at https://www.researchgate.net/profile/Sergey-Voronin-4):

Voronin, S. SAR image compression with int-int transforms and dimensionality reduction, J. of Comp. and Comm., 2022.

Voronin, S., Borovikov, E., Hasan, R. Clustering and presorting for parallel Burrows-Wheeler compression, IJMSSC, 2021.

Nigam, N., Mohseni, S., Valverde, J., Voronin, S., Alonso, J. A Toolset For Creation of Multi-Fidelity Probabilistic Aerodynamic Databases. AIAA Scitech Forum, 2021.

Voronin, S. Numerical approaches for heat transfer problems, Intech Pub., 2020.

Koc, B., Arnavut, Z., Voronin, S., Koçak, H. Near-lossless Image Compression with Parity Reduction. IEEE, 2020.

Voronin, S., Multi-channel similarity based compression. J. CIS, 2020.

Erichson, B., Voronin ,S., et al. Randomized matrix decompositions using R. J. Stat Soft, 2019.

Voronin, S., Multi-stage image restoration with high noise/blur. J. CIS, 2019.

Voronin, S., Xiao, L., Mei, G., Xu, R. Multi-resolution classification techniques for PTSD detection, ISSPIT IEEE, 2018.

Voronin, S., Zaroli, C. Survey of computational methods for inverse problems, Intech Pub. (chap.), 2018.

Voronin, S., Zaroli, C., Cuntoor, N. Conjugate gradient based acceleration for inverse problems. Int. J. on Geomath, 2017.

Voronin, S., Martinsson, P.G. Efficient Algorithms for CUR and Interpolative Matrix Decomposition, J. of Applied and Computational Mathematics, 2016.

Voronin, S., Daubechies, I. An Iteratively Reweighted Least Squares Algorithm for Sparse Regularization, Contemporary Math., 2016. Voronin, S., Martinsson, P.G. RSVDPACK: An implementation of randomized algorithms for computing the singular value, interpolative, and CUR decompositions of matrices on multi-core and GPU architectures, 2016.

Martinsson, P.G. and Voronin, S. A randomized blocked algorithm for efficiently computing rank-revealing factorizations of matrices. SIAM Journal on Scientific Computing, 2016.

Lodhi, M., Voronin, S., Bajwa, W. YAMPA: Yet another matching pursuit algorithm for compressive sensing. SPIE, 2016.

Voronin, S., Nolet, G., Mikesell, T. Compression Approaches for the Regularized Solutions of Linear Systems from Large-Scale Inverse Problems. Int. J. on Geomath, 2015.

Voronin, S., Mikesell, T., Slezak, I., Nolet, G. Solving large tomographic linear systems: size reduction and error estimation. Geophysical Journal International, 2014.

Voronin, S., Chartrand, R. A new generalized thresholding algorithm for inverse problems with sparsity constraints. ICASSP, 2013.

Charlety, J., Voronin, S., et al. Seismic tomography with a sparsity constraint: comparison with smoothing and damping regularization. J. Geophys Res., 2013.

Voronin, S., Woerdeman, H. A new iterative firm-thresholding algorithm for inverse problems with sparsity constraints. Applied and Computational Harmonic Analysis, 2013.

Voronin, S. Regularization of Linear Systems with Sparsity Constraints with Applications to Large Scale Inverse Problems. Ph.D. thesis, 2012.

Simons, F., Loris, I., Nolet, G., Daubechies, I., Voronin, S., Judd, S., Vetter, P., Charléty, J., and Vonesch, C. "Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity." Geophysical Journal International 187, no. 2, 2011.

Gerber, E.P., Voronin, S., Polvani, L. Testing the annular mode autocorrelation timescale in simple atmospheric general circulation models, Mon. Weather Review, 2008.

Voronin, S., Matthewman, J., Charlton, A., Polvani, L., Esler, G. A New Web Based Resource for Studying Major Mid-Winter Stratospheric Sudden Warmings, Stratospheric Processes and Their Role, Climate Newsletter, Vol. 27, 2006.