Pavel Senin
Data Science Software Engineer

PERSONAL DETAILS
3, Allée des Oliviers,
Castanet-Tolosan,
31320, France.
Cell. +33 6 42 51 18 94
senin@hawaii.edu
http://seninp.github.io

AREAS OF EXPERTISE
Data analytics

Bioinformatics

Temporal data mining

Machine learning, NLP

Software engineering

Software metrology

Software process analysis

TECHNICAL SKILLS
Full SDLC, legacy wrapping

Java, R, Python, C/C++

SQL, XML, Perl, bash

DBA, ORM & persistence

deeplearning4j, Spark, BUPC

SOA, REST, MVC, n-tier

HTML, CSS, JavaScript, D3.js

Arduino & sensors

Kanban & Lean

SOFTWARE PROJECTS
GrammarViz 3.0
- *time series pattern mining*
- *discr. param. optimization*

SAX-VSM
- *time series classification*
- *behaviors discovery*

JMotif-R
- *time series mining toolkit*

JMotif-SAX
- *time series discretization*

JMotif-GI
- *grammatical inference*

Hackystat
- *software metrology*

PERSONAL SUMMARY
Looking for a research-oriented position with a practical bent. A goal-oriented, forward thinking researcher with excellent organizational and interpersonal skills.

EXPERIENCE

Advanced Engineering Consultant — current
Altran Sud-Ouest — Toulouse, France

Working on spatio-temporal mining and root cause analysis for Airbus; participating in predictive maintenance capability development at Liebherr Aerospace.

Research Associate, 2016–2017 (Research Technologist, 2007–2010)
Los Alamos National Laboratory — PO Box 1663

Worked on high-throughput comparative analyses of microbiome for biosurveillance. Contributed a metagenome comparative toolkit to R&D 100 award winning EDGE platform. Participated in DOE Exascale Computing Project. Developed a fragment recruiting technique for pioneering single-cell genome sequencing project.

Graduate student/Research assistant — 2005 – 2015
University of Hawai'i at Mānoa *ICS CSDL & RCUH* — Honolulu, HI

Proposed a novel technique for characteristic recurrent behaviors discovery and behavior-based software process analysis using software telemetry. Proposed a novel technique for spatio-temporal anomaly discovery based on algorithmic (Kolmogorov) complexity. Developed a pioneering technique for SAX discretization parameters optimization based on grammatical compression. Assembled the first transgenic plant genome (Carica Papaya) and optimized the cost of its finishing.

Bioinformatics Software Engineer — 2010 – 2015
INRIA/INRA — Rennes/Toulouse, France

Took a part in Life Sciences research projects conducted by French premier research institutions. Designed and implemented high-throughput pipelines for sequencing data acquisition and analysis including genomic and transcriptomic data assembly, annotation, variant and ncRNA discovery and pathway analysis. Worked on structural motif discovery in GPCRs, mating disruption, and vector control.

EDUCATION

PhD, Computer Science (Software process analysis) — 2015
University of Hawai'i at Mānoa — Honolulu, HI, USA

MS, Computer Science (Capstone on MCMC simulation) — 2007
University of Hawai'i at Mānoa — Honolulu, HI, USA

MS, ABD, Appl. Mathematics (Num. optimization) — 2002
Rostov State University (SFedU) — Rostov-na-Donu, Russia

SELECTED PUBLICATIONS — scholar profile

GrammarViz 3.0: Interactive Discovery of Variable-length Time Series Patterns. Senin, P., Lin, J., Wang, X., Oates, T., Gandhi, S., Boedihardjo, A.P., Chen, C., Frankenstein, S. **ACM Trans. Knowl. Discov. Data**, *2018*.

STAVIS 2.0: Mining Spatial Trajectories via Motifs. Chen, C., Boedihardjo, A.P., Jenkins, B.S., Ellison, C.L., Lin, J., Senin, P., Oates, T. **International Symposium on Spatial and Temporal Databases**, *2017*

Time series anomaly discovery with grammar-based compression. Senin, P., Lin, J., Wang, X., Oates, T., Gandhi, S., Boedihardjo, A.P., Chen, C., Frankenstein, S. *In Proc.* **EDBT**, *2015*.