Dragomir Yankov

Department of Computer Science and
Engineering
University of California, Riverside
Riverside, CA 92521

Cell:+1-951-522-1497
Home:+1-951-323-1273
dyankov@cs.ucr.edu
http://www.cs.ucr.edu/~dyankov/

EDUCATION ◇ 2003-2008 **Ph.D.** in Computer Science, University of California Riverside
Advisor: Prof. Eamonn Keogh
Dissertation Title: Learning from Time Series in the Presence of Noise: Unsupervised and
Semi-Supervised Approaches

◇ 1995-2001 **MS** in Computer Science, Sofia University

REFERENCES ◇ Dr. Eamonn Keogh (eamonn@cs.ucr.edu): Academic research advisor

◇ Dr. Nicolas Mayoraz (nicolasm@yahoo-inc.com): Yahoo! Data Mining and Research (DMR)
group

◇ http://www.cs.ucr.edu/~dyankov/reference.png - A publicly available reference from
my second internship with Yahoo!

RESEARCH
INTERESTS

Data Mining and Machine Learning. More specifically I am interested in their application
for time series analysis. Sub-areas in which I have worked and I am particularly excited
about are:

- Mining motifs and discords from large time series data sets
- Manifold learning from high dimensional data (time series in particular) and in the
presence of noise
- Nonlinear dimensionality reduction techniques, again with focus on time series data
- Nearest neighbor approaches for classification, anomaly detection and forecasting
- Density estimation and density based clustering methods
- Ensemble methods

HONORS
AND
AWARDS

◇ 2007 ICDM, Best Paper Award

◇ 2007-2008 UC Riverside, Chancellor Dissertation Fellowship

◇ 2003-2005 UC Riverside, CS Department Scholarship

◇ 2003-2005 UC Riverside, Dean's Fellowship

◇ 2005 UC Riverside, Outstanding Teaching Assistant Award

◇ Travel awards: ECML 2006 (Berlin, Germany), KDD 2007 (San Jose, USA), ICDM 2007
(Omaha, USA)

LIST OF
PUBLICATIONS

◇ **Best Paper Award:** D. Yankov, E. Keogh, U. Rebbapragada: Disk aware discord dis-
covery: finding unusual time series in terabyte sized datasets. *Proceedings of the 7th IEEE
International Conference on Data Mining (ICDM 2007)*

◇ D. Yankov, E. Keogh, K. Kan: Locally constrained Support Vector Clustering. *Proceedings
of the 7th IEEE International Conference on Data Mining (ICDM 2007)*

◇ D. Yankov, E. Keogh, J. Medina, B. Chiu, V. Zordan: Detecting time series motifs un-
der uniform scaling. *Proceedings of the 13th ACM SIGKDD international conference on
Knowledge Discovery and Data mining (KDD 2007), p. 844-853*

⋄ D. Yankov, E. Keogh, L. Wei, X. Xi: Fast best-match shape searching in rotation invariant metric spaces. *Proceedings of the 7th SIAM International Conference on Data Mining (SDM 2007).* An extended version is also to appear in IEEE Transactions on Multimedia, special issue on Data Mining 2007.

⋄ D. Yankov, E. Keogh: Manifold clustering of shapes. *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), p. 1167-1171*

⋄ D. Yankov, D. DeCoste, E. Keogh: Ensembles of nearest neighbor forecasts. *17th European Conference on Machine Learning (ECML 2006), Proceedings. Lecture Notes in Computer Science, p. 545-556*

⋄ D. Yankov, E. Keogh. S. Lonardi, A. Fu: Dot plots for time series analysis. *17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2005), p. 159-168*

RESEARCH EXPERIENCE	

⋄ 2003-present **Graduate Research Assistant, University of California Riverside**
Advisor: Prof. Eamonn Keogh
Research focus - Pattern recognition and anomaly detection in large scale data sets. Probabilistic and dynamic programming techniques for time series analysis. Manifold methods for unsupervised and semisupervised learning. Ensemble methods for learning from time series data. Nearest neighbor forecasting models.

⋄ Summer 2007 **Research Intern, Yahoo! Data Mining and Research group, Sunnyvale, CA**
Advisor: Dr. Nicolas Eddy Mayoraz
Research project - categorization of mail sending IPs to enhance the spam detection system. Density based clustering for identifying regions of similar mail sending patterns. Semisupervised and multiclass supervised approaches mail senders categorization.

⋄ Summer-Fall 2006 **Research Intern, Yahoo! Data Mining and Research (DMR) group, Sunnyvale, CA**
Advisor: Dr. Nicolas Eddy Mayoraz
Research project - time series analysis and machine learning approaches to the *click through protection* problem. Ensembles of heterogeneous learners for click through protection.

⋄ Summer 2005 **Research Intern, Yahoo!Research, Pasadena, CA**
Advisor: Dr. Dennis DeCoste
Research project - time series techniques for predicting web query volume data. Locally weighted regression and multivariate models for time series prediction. Properties of vantage point (VP) trees for time series indexing.

TEACHING EXPERIENCE

TA for the Courses

- · CS 111, Discrete Math, 1 quarter
- · CS 179K, Project in Software Engineering, 1 quarter
- · CS 100, Software Construction, 2 quarters
- · CS 180, Software Engineering, 3 quarters
- · CS 122b, Advanced Embedded and Real-Time Systems, 1 quarter

TECHNICAL BACKGROUND

⋄ Selected Graduate Courses : Data mining techniques, artificial intelligence, design and analysis of algorithms, database management systems, advanced computer networks, distributed systems, computer security, advanced computer architectures

⋄ Languages : C/C++, Java, Matlab

⋄ Databases : Oracle, MS SQL-server, PostgreSQL

⋄ Web Based : XML, HTML, web services, Java servlets, JSP, EJB

⋄ Security : cryptography, X.509 certificates, intrusion detection, SSL

◇ Others: distributed architectures (CORBA, RMI); software design (UML, design patterns etc.)

WORK EXPERIENCE

◇ 2008-present Software Engineer, Data Mining and Research (DMR) group, Yahoo!

◇ 2002-2003 Software Engineer, Sciant www.sciant.com (now VMware Bulgaria)

· Project: XMS (Xml Message Server). The system provides secure management of web services and transactions based on the WSDL protocol

· Responsible for: design and implementation of the data storing mechanisms, object persistence and multiple database support.

◇ 2000-2002 Senior Software Engineer, Practicorp Bulgaria

· Project: Q-go - a system allowing users of PDAs (Palm, Pocket PC) to locate "points of interest" in selected cities.

· Responsible for: Pocket PC implementation of the system

· Project: EIO Backup - a visual client-server application with capabilities to compress, encrypt and store data on remote servers.

· Responsible for: compression, encryption modules of the system. Modules for management of trusted certificates. Participated in the GUI implementation.

◇ 1998-2000 Senior Software Engineer, Prosyst (now SAP Labs Bulgaria - http://www.sap.com/company/saplabs/bulgaria/)

· Project: Prosyst Application Server, subproject: SSL

· Responsible for: Implementation of the SSL v3.0 client and server. Developed a library for issuing and verification of X509 certificates. Participated in the integration modules between SSL and some of the other protocols of the application server - RMI over SSL, HTTPS, and FTPS, SMTP over SSL.

· Project: Prosyst Application Server, subproject: Java Cryptography Provider Suite

· Responsible for: Implementation of a cryptography provider suite providing factories for some digest and (a)symmetric algorithms (SHA1, 3DES, RSA etc). The suite was later used in the SSL protocol implementation.