

Apostolos Dimitromanolakis

Last updated: October, 2018

Contact Information

Joseph and Wolf Lebovic Centre.
5th floor, Room 5-201,
60 Murray Street, Toronto, Ontario, Canada M5T 3L9.
Tel: 647-862-1901
email: apostol@cs.toronto.edu

Education

- 2010** University of Toronto.
M.Sc. Biostatistics, Dala Lana School of Public Health. Supervisor: Lei Sun.
- 2008** University of Toronto.
M.Sc. Computer Science, Department of Computer Science. Thesis title:
"Complexity of distance labeling schemes in planar graphs.". Supervisor: A. Borodin.
- 2002** Technical University of Crete, Greece.
Diploma of Engineering (MS.c Equivalent). Department of Electronic and
Computer Engineering. GPA: 9.2/10 . Thesis title: "Analysis of the Golomb ruler and
Sidon set problems, and determination of large near-optimal Golomb rulers" (received
10/10 mark). Supervisor: Dollas A.

Areas of Expertise

Computer Science: 10+ years of work experience

Biostatistics & Statistical Genetics: 8 years of work experience

Specialization

Computer Science:

Algorithm design and analysis,
Combinatorial Optimization,
Graph Theory,
Low latency web application development.
C++ / R / nodejs integration

Biostatistics:

High dimensional data analysis,
Genome-wide association studies,
Linkage studies,
Statistical analysis automation,
Normalization and organization of large-scale microarray databases,
SRM normalization methods.

Computing/Programming Skills

Programming Languages:

C (expert, 10+ years experience),
C++ (expert, 10+ years experience),
Node.js / Javascript (expert, 5 years experience),
Java (5 years experience),
Python (2 years experience),
Javascript (2 years experience)
Ruby, Shell scripting, MATLAB, Mathematica

Statistical Environments:

R (expert, 8 years experience),
SAS (3 years experience)

Databases:

mongodb (2 year of experience)
mysql, hadoop

Genetic Analysis Software: PLINK, sam/bcftools, bwa, tophat, GATK, QTD, PREST, MERLIN

Visualization: d3.js, R/ggplot2, shiny

Cluster environments: Sun Grid Engine (4+ years of experience)

System Administration: Linux (5+ years of related work experience)

Recent Work Experience

2017 - ongoing

Research Assistant - University of Toronto, Department of Statistical Sciences

Shared segment detection in large scale sequencing studies.

2016 - ongoing

Project Manager II - Statistical Genetics. Mt. Sinai Hospital, Toronto.

Implementation of efficient genetic simulation software and methods. Analyzing exome sequencing data.

2014 - 2016

Chief Technology Officer - Manager of operations. Algocian Inc, Toronto.

Machine learning models for real-time object recognition.

2010 - 2016

Biostatistics and Bioinformatics Specialist. ACDC Laboratory, Mt. Sinai Hospital, Toronto (Supervisor: Dr. E.P. Diamandis).

Statistical analysis of large-scale microarray and proteomic experiments. Automation of statistical methods and creation of billion point microarray databases.

Recent Software and Projects

TRUFFLE: Shared segment detection in genomics data (C++):

23andMe like software for estimating genetic relatedness. Algorithmic optimizations down to bare metal, led to a new genetics software that can estimate related individuals in 30000 x 400000 sized datasets, with just a laptop computer. Release available at <https://adimitromanolakis.github.io/truffle-website/index.html>
Very user friendly with just the minimum of options for a more streamlined experience. Next version will be released soon. (more details available on request)

sim1000G: Simulation of rare and common variants based on 1000 genomes data (R).

Genetic simulator for thousands of markers in families or unrelated individuals. Easy to use and fast. Languages: R. Available at CRAN and <https://github.com/adimitromanolakis/sim1000G>

Real-time Machine Learning models for person detection: Algocian Inc. (commercial)

Development of machine learning models for person detection. Performance optimization of algorithms & code for real-time use. Creation of a real-time RESTful

API in node.js and integration with the C++ base code. Creation of a load-balancing cluster and API to support thousands of cameras. Reduced model evaluation latency to < 50ms and API latency to < 30ms. Languages: C++ / node.js / Javascript / mongodb. Role in project: lead code developer.

prest-plus: PREST-plus: Detection pedigree errors and cryptic relationships.

Dimitromanolakis A, Sun L (2010).

Available at: <http://www.utstat.toronto.edu/sun/Software/Prest/>

During this project we extended the analysis capabilities of prest from analyzing hundreds of microsat markers to GWAS datasets with millions of SNPs. Collectively through code and numerical optimizations, we made prest-plus more than 100x faster than the previous version. In addition we added numerous new features and significantly improved the user experience. Developed in C++. Role in project: code developer

br2: BR-squared - Bias-Reduced estimates via Bootstrap Resampling for overcoming the winner's curse in genome-wide association studies of either disease status or quantitative traits. Dimitromanolakis A, Faye L, Bull SB, Sun L (2009).

Available at: <http://www.utstat.toronto.edu/sun/Software/BR2>

Applying computationally intensive statistical methodology to GWAS data is not trivial. Under the guidance of S. Bull and L. Sun, I was the developer of br2, a software that implements a two-level bootstrap methodology for bias reduction in genome-wide studies. We implemented a software capable of performing up to 1000x1000 bootstrap sampling on datasets containing millions of SNPs. The software can harness multi-core environment, computational clusters or both to reduce data analysis time, in addition to being highly optimized itself. Being a highly complex software it accepts over 50 different command line options. Developed in C++. Role in project: code developer

Refereed Journal Publications -

Differential diagnosis of azoospermia with proteomic biomarkers ECM1 and TEX101 quantified in seminal plasma. Drabovich AP, Dimitromanolakis A, Saraon P, Soosaipillai A, Batruch I, Mullen B, Jarvi K, Diamandis EP. Sci Transl Med. 2013 Nov 20

Integrating meta-analysis of microarray data and targeted proteomics for biomarker identification: application in breast cancer. Pavlou MP, Dimitromanolakis A, Martinez-Morillo E, Smid M, Foekens JA, Diamandis EP. J Proteome Res. 2014.

Serum LAMC2 enhances the prognostic value of a multi-parametric panel in non-small cell lung cancer. Korbakis D, Dimitromanolakis A, Prassas I, Davis GJ, Barber E, Reckamp KL, Blasutig I, Diamandis EP. Br J Cancer. 2015.

Validation of four candidate pancreatic cancer serological biomarkers that improve the performance of CA19.9. Makawita S, Dimitromanolakis A, Soosaipillai A, Soleas I, Chan A, Gallinger S, Haun RS, Blasutig IM, Diamandis EP. BMC Cancer. 2013 Sep 3.

Laminin, gamma 2 (LAMC2): a promising new putative pancreatic cancer biomarker identified by proteomic analysis of pancreatic adenocarcinoma tissues. Kosanam H, Prassas I, Chrystoja CC, Soleas I, Chan A, Dimitromanolakis A, Blasutig IM, Rückert F, Gruetzmänn R, Pilarsky C, Maekawa M, Brand R, Diamandis EP. Mol Cell Proteomics. 2013 Oct;12.

Quantitative proteomic analysis of amniocytes reveals potentially dysregulated molecular networks in down syndrome. Cho CK, Drabovich AP, Karagiannis GS, Martínez-Morillo E, Dason S, Dimitromanolakis A, Diamandis EP. Clin Proteomics. 2013 Feb 8.

Coupling proteomics and transcriptomics in the quest of subtype-specific proteins in breast cancer. Pavlou MP, Dimitromanolakis A, Diamandis EP. Proteomics. 2013 Feb 5

Glycoproteomic identification of potential glycoprotein biomarkers in ovarian cancer proximal fluids. Kuzmanov U, Musrap N, Kosanam H, Smith CR, Batruch I, Dimitromanolakis A, Diamandis EP. Clin Chem Lab Med. 2012 Dec 8.

Quantitative analysis of energy metabolic pathways in MCF-7 breast cancer cells by selected reaction monitoring assay. Drabovich, A. P., Pavlou, M. P., Dimitromanolakis, A., & Diamandis, E. P. (2012). Molecular & cellular proteomics: MCP.

Proteomic Signatures of the Desmoplastic Invasion Front Reveal Collagen Type XII as a Marker of Myofibroblastic Differentiation During Colorectal Cancer Metastasis. Karagiannis, G. S., Petraki, C., Prassas, I., Saraon, P., Musrap, N., Dimitromanolakis, A., & Diamandis, E. P. (2012). Oncotarget, 3(3), 267–285.

A genome-wide linkage study of mammographic density, a risk factor for breast cancer. Greenwood, C. M., Paterson, A. D., Linton, L., Andrulis, I. L., Apicella, C., Dimitromanolakis, A., Kriukov, V., et al. (2011). Breast Cancer Research : BCR, 13 (6), R132.

Digitoxin-induced cytotoxicity in cancer cells is mediated through distinct kinase and interferon signaling networks. Prassas, I., Karagiannis, G. S., Batruch, I., Dimitromanolakis, A., Datti, A., & Diamandis, E. P. (2011). *Molecular cancer therapeutics*, 10(11), 2083–2093.

Combined analysis of three genome-wide association studies on vWF and FVIII plasma levels. Antoni, G., Oudot-Mellakh, T., Dimitromanolakis, A., Germain, M., Cohen, W., Wells, P., Lathrop, M., et al. (2011). *BMC Medical Genetics*, 12, 102.

A flexible genome-wide bootstrap method that accounts for ranking and threshold-selection bias in GWAS interpretation and replication study design. Faye, L. L., Sun, L., Dimitromanolakis, A., & Bull, S. B. (2011). *Statistics in medicine*, 30(15), 1898–1912.

BR-squared: a practical solution to the winner's curse in genome-wide scans. Sun, L., Dimitromanolakis, A., Faye, L. L., Paterson, A. D., Waggott, D., & Bull, S. B. (2011). *Human genetics*, 129(5), 545–552.

A spatially explicit learning model of migratory fish and fishers for evaluating closed areas. Moustakas A, Silvert W, Dimitromanolakis A. (2006). *Ecological Modelling* 2006. 192: 245-258.

Refereed book chapters

Identifying cryptic relationships. Sun, L., & Dimitromanolakis, A. (2012). *Methods in molecular biology* (Clifton, N.J.), 850, 47–57.

Theses

Distance labeling schemes in planar and other classes of graphs. Department of Computer Science, University of Toronto, 2008

Analysis of the Golomb ruler and Sidon set problems, and determination of large near-optimal Golomb rulers. Dimitromanolakis A. Diploma thesis, Technical University of Crete, June 2002 (<http://www.cs.toronto.edu/~apostol/golomb/>). Supervisor: Dollas A.

Conference presentations

Differential expression of kallikrein genes in normal and cancerous tissues.

A. Dimitromanolakis, V. Milou. 4th International Symposium on Kallikreins and Kallikrein-related peptidases

BR-squared: a Practical Solution to the Winner's Curse in Genome-Wide Scans.

L. Sun, S.B. Bull, L. Faye, A. Dimitromanolakis, D. Waggott, A.D. Paterson, The DCCT/EDIC Research Group. Platform presentation at the American Society of Human Genetics 59th Annual Meeting (2009).

Conference proceedings

Prioritized-GWAS based on linkage information identifies novel putative loci influencing Factor XII plasma levels.

Gagnon F, Dimitromanolakis A, Antoni G, Martinez A, Greliche N, Buil A, Soria JM, Morange P, Wells PS, Tregouet D, Sun L (2011).

XXIII International Society on Thrombosis & Haemostasis (ISTH) Congress with 57th Annual SSC Meetings

Genome-wide bootstrap bias reduction for point and interval estimation that accounts for ranking- and threshold- selection bias in discovery GWAS, with implications for replication study sample size.

Faye L, Sun L, Dimitromanolakis A, Bull SB (2010).

Abstract #214 presented at the IGES annual meeting. Genetic Epidemiology 34:977.

A flexible genome-wide bootstrap method that accounts for ranking- and threshold-selection bias in GWAS interpretation and replication study design.

Faye L, Sun L, Dimitromanolakis A, Bull SB (2010).

The 5th Canadian Genetic Epidemiology and Statistical Genetics Workshop.

Accurate IBD inference identifies cryptic relatedness in 9 HapMap populations. Dimitromanolakis A, Paterson AD, Sun L.

The American Society of Human Genetics 59th Annual Meeting (2009).

A comprehensive look at the likelihood and bootstrap approaches to overcome the winner's curse in GWAS.

Faye L, Sun L, Dimitromanolakis A, Bull S. B.

Abstract #126 presented at the IGES annual meeting. Genetic Epidemiology 33:788-789. (2009).

PREST-plus: An improved software tool for the detection of cryptic relations and pedigree errors using high-throughput genotype data.

Dimitromanolakis A, Paterson AD, Sun L.

The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop (2009).

Technical Reports

A flexible genome-wide bootstrap method that accounts for ranking- and threshold-selection bias in GWAS interpretation and replication study design.

Laura L. Faye, Lei Sun, Apostolos Dimitromanolakis, Shelley B. Bull. Technical Report No. 1006
Oct 16, 2010. University of Toronto, Department of Statistics

Work Experience (detailed)

2017 - ongoing

Research Assistant - University of Toronto, Department of Statistical Sciences
Shared segment detection in large scale sequencing studies.

2016 - ongoing

Project Manager II - Statistical Genetics. Mt. Sinai Hospital, Toronto.
Implementation of efficient genetic simulation methods.
Analyzing exome sequencing data.

2014 - 2016

Chief Technology Officer - Manager of operations. Algocian Inc, Toronto.
Machine learning models for real-time object recognition.

2010 - 2016

Biostatistics and Bioinformatics Specialist. ACDC Laboratory, Mt. Sinai Hospital, Toronto
(Supervisor: Dr. E.P. Diamandis).

Statistical analysis of large-scale microarray and proteomic data. Automation of statistical methods, development and validation of high-dimensional models.

2009 - 2010

M.Sc. Practicum. on Statistical Genetics. Sick Kids Hospital, Toronto (Supervisors: Dr. J. Rommens / Dr. C. Greenwood)

2009 - 2011

Research Associate in Statistical Genetics, Genetic Epidemiology Lab, 7th floor, Dalla Lana School of Public Health, University of Toronto (Supervisors: Dr. F. Gagnon / Dr. L. Sun)

2008 – 2009

Mt Sinai Hospital / Dalla Lana School of Public Health, Toronto. Research Associate in Statistical Genetics (Supervisors: Dr. L. Sun / Dr. S.B. Bull)

2002 – 2007

Teaching assistant for computer science and computer engineering courses.

2001

Software Engineering at Xatzirakis Ltd. Chania, Crete. Development of a high performance protocol parser for a travel agency

2000 - 2001

Database Research and Development. Distributed Multimedia Information Systems and Applications Lab, Chania Crete (www.music.tuc.gr).

1998 - 2001

System administrator at Software Engineering Lab. Administration of a Linux cluster. Technical University of Crete, Greece

1996 - 2000

Software engineer & System administrator. 2M General Electronics Ltd. Service Provider.

Software Projects**2016-**

octo-DSP: User friendly speaker equalization and software-crossover filters, using FFT convolution algorithms. Target: Ubuntu / Raspberry PI. Languages: C++/ node.js / html.

2014-2016

Machine learning models for image/video object recognition, in real-time. Reducing false-positive rate to make it suitable for business use. Creation of a RESTful API and load-balancing computing cluster.

2008-2011

br2: Bias correction method for genome wide association studies. Download link: <http://fisher.utstat.toronto.edu/sun/Software/BR2/>

2008-2010

prest-plus: Identity by descent estimation, relationship inference and relationship hypothesis testing from high throughput genome-wide datasets. Download link:
<http://fisher.utstat.toronto.edu/sun/Software/Prest/>

2007-2008

Serena: a mobile photorealistic astronomical star chart application for J2ME cellphones.
Coded in Java J2ME / Ruby.

2006

"DNA Global Sequence alignment based on Dijkstra's algorithm."

Design and Development of a DNA global sequence alignment algorithm based on a novel graph theoretic formulation. Development in C. (part of CSC2427: Algorithms in Molecular Biology graduate course project)

2003

Development of a Linux kernel driver for the Guillemote Maxi Radio FM 2000 PCI Radio Card.

2002

"Distributed search for large near optimal Golomb rulers."

Design and development of a distributed search application for searching of large combinatorial structures. Developed in C and used a network of 20 personal computers to complete the search. Part of diploma thesis.

1998-2000

Developer of a distributed user management application for 2M G.E Ltd. in TCL/C.