

---

# **SPIROS PAPADIMITRIOU**

## **Curriculum Vitae**

---

### **Personal information**

Spyridon (Spiros) Papadimitriou  
Rutgers Business School,  
100 Rockafeller Rd #2133,  
Piscataway, NJ 08854

E-Mail: [spapadim@business.rutgers.edu](mailto:spapadim@business.rutgers.edu), [spapadim@gmail.com](mailto:spapadim@gmail.com)

---

### **EMPLOYMENT AND EDUCATION**

---

#### **Education**

- 2005            Ph.D. Degree, Computer Science Department, Carnegie Mellon University.  
2001            M.Sc. Degree, Computer Science, Carnegie Mellon University.  
1998            B.Sc. Degree, Computer Science, University of Crete, Greece.

#### **Employment history**

- 7/2016–present    Management Science and Information Systems (MSIS) Department,  
Business School, Rutgers University.  
Associate Professor
- 9/2012–6/2016    Management Science and Information Systems (MSIS) Department,  
Business School, Rutgers University.  
Assistant Professor
- 8/2010–8/2012    Google Inc., Mountain View, CA, USA.  
Research Scientist
- 10/2005–7/2010   IBM Thomas J. Watson Research Center, Hawthorne, NY, USA.  
Research Staff Member
- 6/2000–12/2000   IBM Almaden Research Center, San Jose, CA, USA.

#### **Other affiliations**

- 12/2012–present   Graduate faculty (courtesy appointment),  
Computer Science Department, Rutgers University.
- 1/2013–present    Center for Dynamic Data Analytics, Rutgers University.

#### **Visiting positions**

- 8/2004–9/2004    Basic Research Unit (BRU), Helsinki Institute for Information Technology (HIIT), Finland.
- 6/2004            Department of Computer Science, Hong Kong University of Science and Technology.

---

## AWARDS AND HONORS

---

PAKDD 2014 Best Student Paper Runner-up Award.  
WebSci 2014 Best Poster Award.  
KDD 2012 Outstanding Reviewer Award.  
SDM 2008 Best Research Paper Award.  
IBM Second Plateau Invention Achievement Award (2008).  
IBM Research Division Award (2008).  
Siebel scholarship recipient (2005).  
Class valedictorian (1998), School of Sciences, University of Crete, Greece.  
Bronze medal, International Olympiad in Informatics V (1993), Mendoza, Argentina.

---

## PUBLICATIONS

---

(H-INDEX: 29)

### Journal publications

1. Miguel Araujo, Stephan Günnemann, Spiros Papadimitriou, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos E. Papalexakis, Danai Koutra, *Discovery of ‘Comet’ Communities in Temporal and Labeled Graphs (Com<sup>2</sup>)*, KAIS 46(3): 657-677, 2016.
2. Bin Liu, Hui Xiong, Spiros Papadimitriou, Yanjie Fu, Zijun Yao, *A General Geographical Probabilistic Factor Model for Point of Interest Recommendation*, IEEE TKDE 27(3): 755-768, 2015.
3. Senjuti Roy Basu, Tina Eliassi-Rad, Spiros Papadimitriou, *Fast Best-Effort Search on Graphs with Multiple Attributes*, IEEE TKDE 27(3): 755–768, 2015.
4. Hanghang Tong, Spiros Papadimitriou, Christos Faloutsos, Philip S. Yu, Tina Eliassi-Rad, *Gateway Finder in Large Graphs: Problem Definitions and Fast Solutions*, Inf. Retr. 15(3–4): 391–411, 2012.
5. Cui Zhu, Hiroyuki Kitagawa, Spiros Papadimitriou, Christos Faloutsos, *Outlier Detection by Example*, J. Intell. Inf. Syst. 36(2): 217–247, 2011.
6. Yasushi Sakurai, Christos Faloutsos, Spiros Papadimitriou, *Fast Discovery of Group Lag Correlations in Streams*, ACM TKDD 5(1), 2010.
7. Hanghang Tong, Spiros Papadimitriou, Philip S. Yu, Christos Faloutsos, *Fast Monitoring of Proximity and Centrality on Time-evolving Bipartite Graphs*, SAM 1(3): 142–156, 2008.
8. Jimeng Sun, Dacheng Tao, Spiros Papadimitriou, Philip S. Yu, Christos Faloutsos, *Incremental Tensor Analysis: Theory and Applications*, ACM TKDD 2(3), 2008.
9. Kyriakos Mouratidis, Dimitris Papadias, Spiros Papadimitriou, *Tree-based Partition Querying: A Methodology for Computing Medoids in Large Spatial Databases*, VLDB Journal 17(4):923–945, 2008.
10. Spiros Papadimitriou, Anthony Brockwell, Christos Faloutsos, *Adaptive, Unsupervised Stream Mining*, VLDB Journal 13(3):222–239, 2004.

**Best papers of VLDB 2003 issue**

11. Rakesh Agrawal, Roberto Bayardo, Daniel Gruhl, Spiros Papadimitriou, *Vinci: A Service-Oriented Architecture for Rapid Development of Web Applications*, Computer Networks 39(5):523–539, 2002.  
**Best papers of WWW-10 issue**

### Conference publications (refereed)

1. Spiros Papadimitriou, Evangelos E. Papalexakis, Bin Liu, Hui Xiong, *Remix in 3D printing: What your sources say about you*, WWW 2015: 367–368, Florence, Italy.
2. Yanjie Fu, Guannan Liu, Spiros Papadimitriou, Hui Xiong, Yong Ge, Hengshu Zhu, Chen Zhu, *Real Estate Ranking via Mixed Land-use Latent Models*, KDD 2015, Sydney, Australia.
3. Spiros Papadimitriou, Evangelos E. Papalexakis, *Towards Laws of the 3D-printable Design Web*, WebSci 2014: 255–256, Bloomington, USA.  
**Winner of best poster award**
4. Miguel Araujo, Spiros Papadimitriou, Stephan Günnemann, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos E. Papalexakis, Danai Koutra, *Com2: Fast Automatic Discovery of Temporal Communities*, PAKDD 2014: 271–283, Tainan, Taiwan.  
**Winner of best student paper runner-up award**
5. U Kang, Spiros Papadimitriou, Jimeng Sun, Hanghang Tong, *Centralities in Large Networks: Algorithms and Observations*, SDM 2011: 119–130, Mesa, USA.
6. Hanghang Tong, Spiros Papadimitriou, Christos Faloutsos, Philip S. Yu, Tina Eliassi-Rad, *BASSET: Scalable Gateway Finder in Large Graphs*, PAKDD 2010: 449–463, Hyderabad, India.
7. Keith Henderson, Tina Eliassi-Rad, Spiros Papadimitriou, Christos Faloutsos, *HCDF: A Hybrid Community Discovery Algorithm*, SDM 2010: 754–765, Columbus, USA.
8. Jimeng Sun, Spiros Papadimitriou, Ching-Yung Lin, Nan Cao, Shixia Liu, Weihong Qian, *MultiVis: Content-based Social Network Exploration Through Multi-way Visual Analysis*, SDM 2009: 1063–1074, Sparks, USA.
9. Spiros Papadimitriou, Jimeng Sun, *DisCo: Distributed Co-Clustering with Map-Reduce*, ICDM 2008: 512–521, Pisa, Italy.
10. Spiros Papadimitriou, Jimeng Sun, Philip S. Yu, Christos Faloutsos, *Hierarchical Parameter-Free Community Discovery*, PKDD 2008: 170–187, Antwerp, Belgium.
11. Xiaohui Gu, Spiros Papadimitriou, Shu-Ping Chang, Philip S. Yu, *Online Failure Forecast for Fault-Tolerant Data Stream Processing*, ICDE 2008: 1388–1390, Cancun, Mexico.
12. Hanghang Tong, Spiros Papadimitriou, Jimeng Sun, Philip S. Yu, Christos Faloutsos, *Colibri: Fast Mining of Large Static and Dynamic Graphs*, KDD 2008: 686–694, Las Vegas, USA.
13. Hanghang Tong, Spiros Papadimitriou, Philip S. Yu, Christos Faloutsos, *Proximity Tracking on Time-Evolving Bipartite Graphs*, SDM 2008: 704–715, Atlanta, USA.  
**Winner of best paper award**
14. Xiaohui Gu, Spiros Papadimitriou, Philip S. Yu, Shu-Ping Chang, *Toward Predictive Failure Management for Distributed Stream Processing Systems*, ICDCS 2008: 825–832, Beijing, China.
15. Spiros Papadimitriou, Feifei Li, George Kollios, Philip S. Yu, *Time Series Compressibility and Privacy*, VLDB 2007: 459–470, Vienna, Austria.

16. Jimeng Sun, Spiros Papadimitriou, Philip S. Yu, Christos Faloutsos, *GraphScope: Parameter-Free Mining of Large Time-Evolving Graphs*, KDD 2007: 687–696, San Jose, USA.
17. Tsuyoshi Idé, Spiros Papadimitriou, Michail Vlachos, *Computing Correlation Anomaly Scores using Stochastic Nearest Neighbors*, ICDM 2007: 523–528, Omaha, USA.
18. Feifei Li, Jimeng Sun, Spiros Papadimitriou, George Mihaila, Ioana Stanoi, *Hiding in the Crowd: Privacy Preservation on Evolving Streams through Correlation Tracking*, ICDE 2007: 686–695, Istanbul, Turkey.
19. Spiros Papadimitriou, Jimeng Sun, Philip S. Yu, *Local Correlation Tracking in Time Series*, ICDM 2006: 456–465, Hong Kong.
20. Jimeng Sun, Spiros Papadimitriou, Philip S. Yu, *Window-based Tensor Analysis on High-dimensional and Multi-aspect Streams*, ICDM 2006: 1076–1080, Hong Kong.
21. Spiros Papadimitriou, Philip S. Yu, *Optimal Multi-Scale Patterns in Time Series Streams*, SIGMOD 2006: 647–658, Chicago, USA.
22. Michail Vlachos, Spiros Papadimitriou, Zografoula Vagena, Philip S. Yu, *Indexing and Visualization of High-Dimensional Data via Dimension Reorderings*, PKDD 2006: 407–420, Berlin, Germany.
23. Jimeng Sun, Spiros Papadimitriou, Christos Faloutsos, *Distributed Pattern Discovery in Multiple Streams*, PAKDD 2006: 713–718, Singapore.
24. Deepak S. Turaga and Michail Vlachos and Spiros Papadimitriou and Philip S. Yu, *SLIDE: Streaming and Load-Adaptive Periodicity Estimation*, ICASSP 2006: III-1000–1003, Toulouse, France.
25. Spiros Papadimitriou, Jimeng Sun, Christos Faloutsos, *Streaming Pattern Discovery in Multiple Time-Series*, VLDB 2005: 697–708, Trondheim, Norway.
26. Spiros Papadimitriou, Aristides Gionis, Panayiotis Tsaparas, Risto A. Väisänen, Christos Faloutsos, Heikki Mannila, *Parameter-Free Spatial Data Mining Using MDL*, ICDM 2005: 346–353, Houston, USA.
27. Aristides Gionis, Alexander Hinneburg, Spiros Papadimitriou, Panayiotis Tsaparas, *Dimension Induced Clustering*, KDD 2005: 51–60, Chicago, USA.
28. Kyriakos Mouratidis, Dimitris Papadias, Spiros Papadimitriou, *Medoid Queries in Large Spatial Databases*, SSTD 2005: 55–72, Angra dos Reis, Brazil.
29. Yasushi Sakurai, Spiros Papadimitriou, Christos Faloutsos, *BRAID: Stream Mining through Group Lag Correlations*, SIGMOD 2005: 599–610, Baltimore, USA.
30. Yasushi Sakurai, Spiros Papadimitriou, Christos Faloutsos, *AutoLag: Automatic Discovery of Lag Correlations in Stream Data*, ICDE 2005: 159–160, Tokyo, Japan.
31. Deepayan Chakrabarti, Spiros Papadimitriou, Dharmendra Modha, Christos Faloutsos, *Fully Automatic Cross-Associations*, KDD 2004: 79–88, Seattle, USA.
32. Cui Zhu, Hiroyuki Kitagawa, Spiros Papadimitriou, Christos Faloutsos, *OBE: Outliers By Example*, PAKDD 2004: 222–234, Sydney, Australia.
33. Spiros Papadimitriou, Anthony Brockwell, Christos Faloutsos, *Adaptive, Hands-Off Stream Mining*, VLDB 2003: 560–571, Berlin, Germany.
- Selected for best papers journal issue**
34. Spiros Papadimitriou, Christos Faloutsos, *Cross-Outlier Detection*, SSTD 2003: 199–213, Santorini, Greece.
35. Spiros Papadimitriou, Hiroyuki Kitagawa, Phillip B. Gibbons, Christos Faloutsos, *LOCI: Fast Outlier Detection Using the Local Correlation Integral*, ICDE 2003: 315–327, Bangalore, India.

36. Mengzhi Wang, Tara M. Madhyastha, Ngai Hang Chan, Spiros Papadimitriou, Christos Faloutsos, *Data Mining Meets Performance Evaluation: Fast Algorithms for Modelling Bursty Traffic*, ICDE 2002: 507–516, San Jose, USA.
37. Agma Traina, Caetano Traina, Spiros Papadimitriou, Christos Faloutsos, *Tri-Plots: Scalable Tools for Multidimensional Data Mining*, KDD 2001: 184–193, San Francisco, USA.
38. Rakesh Agrawal, Roberto Bayardo, Daniel Gruhl, Spiros Papadimitriou, *Vinci: A Service-Oriented Architecture for Rapid Development of Web Applications*, WWW-10: 355–365, Hong Kong.  
**Selected for best papers journal issue**

### Book chapters

1. Spiros Papadimitriou, Jimeng Sun, Christos Faloutsos, Philip S. Yu, *Dimensionality Reduction and Filtering on Time Series Sensor Streams*, in *Managing and Mining Sensor Data*: 103-141, Springer 2013.
2. Spiros Papadimitriou, Jimeng Sun, Christos Faloutsos, Philip S. Yu, *Dimensionality Reduction and Filtering on Time Series Sensor Streams*, in: *Mining and Managing Sensor Data*, Kluwer Academic Publishers (to appear).
3. Hanghang Tong, Spiros Papadimitriou, Philip S. Yu, Christos Faloutsos, *Proximity Tracking on Dynamic Bipartite Graphs: Problem Denitions and Fast Solutions*, in: *Link Mining: Models, Algorithms and Applications*, Springer, 2010.
4. Anastasia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou, Bianca Schroeder, Karl Schnaitter, and Gavin Sherry, *PostgreSQL*, in *Database System Concepts, 6th ed.*, McGraw Hill, 2010.
5. Spiros Papadimitriou, *Anomaly Detection on Streams*, in: *Encyclopedia of Database Systems*, Springer, 2009.
6. Kyriakos Mouratidis, Dimitris Papadias, Spiros Papadimitriou, *Medoid Queries in Large Spatial Datasets*, in: *Geographic Data Mining and Knowledge Discovery, 2nd ed.*, CRC Press, 2008.
7. Jimeng Sun, Spiros Papadimitriou, Philip S. Yu, *Tensor Analysis on Multi-Aspect Streams*, in: João Gamma, Mohamed Medhat Gaber (eds.), *Learning from Data Streams*, Springer, 2007.
8. Spiros Papadimitriou, Jimeng Sun, Christos Faloutsos, *Dimensionality Reduction and Forecasting on Streams*, in: Charu Aggarwal (ed.), *Data Streams: Models and Algorithms*, Springer, 2007.
9. Sailesh Krishnamurthy, Spiros Papadimitriou, Bianca Schröder, Anastassia Ailamaki, *PostgreSQL*, in: Abraham Silberschatz, Henry F. Korth, S. Sudarshan, *Database System Concepts, 5th ed.*, McGraw Hill, 2006.

### Other (selected)

1. Jieying Chen, Jia-Yu Pan, Spiros Papadimitriou, Christos Faloutsos, *TSum: Fast, Principled Table Summarization*, ADKDD 2013 (held with KDD 2013), Chicago, USA.
2. Spiros Papadimitriou, *Scalable Graph Analysis with MapReduce: Clustering, Centrality Measures*, INFORMS ICS 2013, Santa Fe, USA.
3. Spiros Papadimitriou, *Scalable Graph Analysis: Clustering, Centrality measures*, 1st Rutgers Applied Probability Day, 2012.
4. Stavros Harizopoulos, Spiros Papadimitriou, *A Case for Micro-CellStores: Energy-efficient Data Management on Recycled Smartphones*, DaMoN 2011:50–55 (held with SIGMOD 2011), Athens, Greece.

5. Ching-Yung Lin, Nan Cao, Shixia Liu, Spiros Papadimitriou, Jimeng Sun, Xifeng Yan, *SmallBlue: Social Network Analysis for Expertise Search and Collective Intelligence*, demo in ICDE 2009: 1483–1486, Shanghai, China.
6. Jimeng Sun, Spiros Papadimitriou, Christos Faloutsos, *Online Latent Variable Detection in Sensor Networks*, demo in ICDE 2005: 1126–1127, Tokyo, Japan.

### Work in progress

1. Chuanren Liu, Hui Xiong, Spiros Papadimitriou, Yong Ge, Keli Xiao, *Proactive Workflow Modelling by Stochastic Processes with Application to Healthcare*, IEEE TKDE (under submission).
2. Yanjie Fu, Guannan Liu, Spiros Papadimitriou, Yuanchun Zhou, Hui Xiong, *Fused Latent Models for Assessing Product Return Propensity in Online Commerce*, INFORMS ISR (under preparation).
3. Spiros Papadimitriou, Evangelos Papalexakis, Bin Liu, Hui Xiong, *Remix in 3D printing: A quantitative analysis* (extended version), under preparation.
4. *Multi-attribute Homophily via Graph Autoregression Models*, with Xiaodong Lin, Tina Eliassi-Rad.
5. *Innovation Propagation in 3D Printing Remix Graphs*, with Xiaodong Lin, Hui Xiong.
6. *Compressive Sensing for Querying Irregularly Sampled Time Series*.

---

## INVITED TALKS AND TUTORIALS

---

### Tutorials

1. *Mining Smartphone Data* (with Tina Eliassi-Rad), ICWSM 2016, Cologne, Germany.
2. *Mining Mobility Data* (with Tina Eliassi-Rad), WWW 2015, Florence, Italy.
3. *Mining Data from Mobile Devices: A Survey of Smart Sensing and Analytics* (with Tina Eliassi-Rad), KDD 2013, Chicago, IL.
4. *Large-Scale Data Mining: An Introduction to Map-Reduce and Beyond*, Graph Analytics Workshop, Carnegie Mellon University, September 2012, Pittsburgh, USA (**invited tutorial**).
5. *Large-Scale Data Mining: MapReduce and Beyond* (with Jimeng Sun, and Rong Yan), KDD 2010, Washington DC, USA.
6. *Large-Scale Data Mining: MapReduce and Beyond* (with Jimeng Sun, and Rong Yan), SDM 2009, Miami, USA.
7. *Temporal Data Mining* (with Michalis Vlachos), Department of Spatial Information Science and Engineering, University of Maine, June 2008, Orono, USA (**invited tutorial**).
8. *Hands-On Time-Series Analysis with Matlab* (with Michalis Vlachos), ICDM 2006, Hong Kong.

### Invited talks

1. *Scalable Graph Analysis*, invited lecture REU 2013 program, DIMACS, Rutgers.
2. *Large Scale Estimation and Forecasting in Practice*, at LDMTA 2011 (held in conjunction with KDD 2011).

3. *Large-Scale Data Mining*, invited course lecture for *Selected Topics in Data Analytics* (Fall 2010), University of Southern California, Los Angeles, USA.
4. *Evolving Graph Mining*, **keynote talk** at ADN 2009 (held in conjunction with SDM 2009), Sparks, USA.
5. *Mining Stream and Graph Data*, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, October 2008, São Carlos, Brazil.
6. *Emerging Technologies for Large-Scale Search and Mining*, **keynote talk** at WAAMD 2008 (held in conjunction with SBBD 2008), Campinas, Brazil.

---

## PATENTS

---

### Granted

1. *Identifying Optimal Multi-scale Patterns in Time-series Streams*, with Philip S. Yu (granted 5/17/2011, US7945570; filed 8/31/2009, IBM docket no. YOR920060244).
2. *Systems and Methods for Simultaneous Summarization of Datacube Streams*, with Jimeng Sun and Philip S. Yu (granted 3/17/2009, US7505876; filed 1/7/2007, IBM docket no. YOR920060647)
3. *Method of Privacy-Preserving Data Stream Publishing using Dynamic Correlations*, with Yuan Chi Chang, Feifei Li, George Mihaila, Ioana Stanoi, Jimeng Sun, Philip S. Yu (granted 12/14/2010 US7853545; filed 2/26/2007, IBM docket no. YOR920060786).
4. *Method of Privacy-Preserving Data Stream Publishing using Dynamic Autocorrelations*, with Yuan Chi Chang, Feifei Li, George Mihaila, Ioana Stanoi, Jimeng Sun, Philip S. Yu (granted 11/23/2010, US7840516; filed 2/26/2007, IBM docket no. YOR920060787).
5. *Systems and Methods for Predictive Failure Management*, with Shu-Ping Chang, Xiaohui Gu and Philip S. Yu (granted 6/1/2010 US7730364; filed 4/5/2007, IBM docket no. YOR920060859).
6. *Methods Involving Computing Correlation Anomaly Scores*, with Tsuyoshi Idé (granted 1/27/2009, US7483934; filed 12/18/2007, IBM docket no. JP820070157).
7. *Methods and Apparatus for Perturbing an Evolving Data Stream for Time Series Compressibility and Privacy*, with Philip S. Yu (granted 12/27/2011, US8086655; filed 9/14/2007, IBM docket no. YOR820070447).
8. *Determining the Importance of Data Items and their Characteristics Using Centrality Measures*, with Ching-Yung Lin, Hanghang Tong, Jimeng Sun, U Kang (granted 8/26/2014, US8818918; filed 4/28/2011, IBM docket no. YOR920110240).

### Pending

9. *Content-Based and Time-Evolving Social Network Analysis*, with Ching-Yung Lin, Jimeng Sun, Kun-Lung Wu (published 3/3/2011, US20110055379 A1; filed 9/2/2009, IBM docket no. YOR820080476).
10. *Systems and Methods for Indexing and Visualization of High-dimensional Data via Dimension Reorderings*, with Zografoula Vagena, Michail Vlachos, and Philip S. Yu (published 3/20/2008, US20080071843 A1; filed 9/14/2006, IBM docket no. YOR920060578).
11. *Alert Summarization System for Online Click Fraud Filters*, with Jieying Chen, Jia-Yu Pan, Christos Faloutsos (Google reference no. GP-10572-00 / 2011).

---

## TEACHING

---

### Courses taught

- Fall 2015 Database and Web Applications (33:136:440), Rutgers University, Business School.
- Fall 2015 Foundations of Business Programming (33:136:388), Rutgers University, Business School.
- Spring 2015 Large Scale Data Analysis (33:136:487), Rutgers University, Business School.
- Spring 2015 Business Data Management (33:136:470), Rutgers University, Business School.
- Fall 2014 Database and Web Applications (33:136:440), Rutgers University, Business School.
- Fall 2014 Foundations of Business Programming (33:136:388), Rutgers University, Business School.
- Spring 2014 Large Scale Data Analysis (33:136:487), Rutgers University, Business School.
- Spring 2014 Business Data Management (33:136:470), Rutgers University, Business School.
- Fall 2013 Data Intensive Analytics (26:711:685), Rutgers University, Business School.
- Fall 2013 Foundations of Business Programming (33:136:388), Rutgers University, Business School.
- Spring 2013 Business Data Management (33:623:470), Rutgers University, Business School.
- Fall 2012 Object Oriented Programming in Finance I (22:839:614), Rutgers University, Business School.
- 
- Spring 2003 Database Applications (15-415), Lead Teaching Assistant, Carnegie Mellon University, Computer Science Department.
- Fall 2001 Database Management Systems (15-721), Teaching Assistant, Carnegie Mellon University, Computer Science Department.
- Spring 1999 Introduction to Computer Systems (15-213), Teaching Assistant and Recitation Lecturer, Carnegie Mellon University, Computer Science Department, Spring 1999.

### Curriculum development

Class webserver (with links to syllabi): <https://io.clusterhack.net/>

#### Business Data Management

- Designed new curriculum based on modern practices, focusing on database system use and applications.
- Set up and maintain database server and web server for class.
- Developed software for online answer testing and live leaderboards for SQL assignment.
- Designed assignments for web application development; implemented, set up, and maintained necessary supporting technical infrastructure.

#### Large-Scale Data Analysis

- Designed new theoretical curriculum based on current data science practices, focusing on understanding and applying state-of-the-art data analysis methods on real data.
- Designed completely new curriculum on practical Python (with NumPy and scikit-learn) for data analysis.
- Developed online contest scoreboard software (a-la Kaggle) for homework assignments.



- Designed assignments on real-world data analysis, using real-world data (UCI data repository, and 500M Amazon product reviews from the SNAP repository); implemented necessary supporting technical infrastructure to make task tractable for students.

#### Database and Web Applications

- Proposed course and designed completely new curriculum, based on modern practices.
- Curriculum follows three-pronged approach: (i) semi-structured data on the web (HTML/XML, CSS, XPath and XSLT, JSON); (ii) front-end technologies (Javascript, HTML5, AJAX, jQuery, AngularJS); (iii) back-end technologies (web application frameworks, object-relational mappers).
- Set up and maintain server infrastructure for students.
- Developed software for web-based semi-structured data processing (XPath and XSLT), online answer testing, and live leaderboards for semi-structured data assignment.
- Developed web applications to illustrate key concepts, via hands-on coding with students participating in-class.
- Wrote notes to introduce students to basics of the Unix shell, for server-based development.

#### Foundations of Business Programming

- Developed Java-based curriculum based on an objects-last, interdisciplinary approach to teaching programming.
- Designed and developed assignments with unit-test infrastructure (based on JUnit), to provide immediate feedback to students while working on solutions.

#### Object-oriented Programming in Finance I (with Farid Alizadeh)

- Developed C++-based objects-last curriculum for programming.
- Designed and developed assignments with unit-test infrastructure.

### Students supervised

#### Ph.D.

#### Committees

1. Nilofar Varzgani, Rutgers University (advisor: Michael Katehakis)
2. Zhongmou Li, Rutgers University (advisor: Hui Xiong)
3. Yanjie Fu, Rutgers University (advisor: Hui Xiong)
4. Bin Liu, Rutgers University (advisor: Hui Xiong)
5. Chuanren Liu, Rutgers University (advisor: Hui Xiong)

#### Advisor

1. John O'Donnell, Capstone Project (2015), Rutgers Business School.
2. John Misarti, Independent study (2015), Rutgers Business School.
3. William Dakota Commons, Honors Project (2015), Rutgers Business School.
4. Xiang (David) Cui, MIT Capstone Project (2014), Rutgers Business School.

#### Ph.D. Student Mentor

1. Zhe (Shirley) Chen, Internship (2012), Google.
2. U Kang, Summer Internship (2010), IBM Research.
3. Feifei Li (co-mentor), Summer Internship (2006), IBM Research.
4. Jimeng Sun, Summer Internship (2006), IBM Research.

---

## SERVICE

---

### University service

MSIS Faculty Hiring Committee (2015).

Business Analytics and Information Technology (BAIT) major, co-coordinator (2014–).

Johnson and Johnson Case Competition, faculty mentor (2014 and 2015).

Technology Taskforce, Rutgers Business School, member (2013–14).

MSIS department seminar, webmaster and schedule coordinator (2013–).

Business Information Technology Society (BITS), faculty advisor and escort, trip to Big Data Conference, Washington DC (11/2013).

Technology Policy Committee, Rutgers Business School, member (2012–13).

### Professional activities and service

Organizing committees

1. Registrations chair, 15th IEEE International Conference on Data Mining (ICDM 2015)
2. Web chair, 13th IEEE International Conference on Data Mining (ICDM 2013)
3. Demonstrations co-chair, 15th European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2011)

Editorial boards

1. Journal of Information and Data Management (JIDM), 2015–

Program committees (62)

1. ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD): 2016, 2015, 2014 (industry & research tracks), 2013, 2010, 2009 (demos committee), 2008, 2007, and 2006
2. European Conf. on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD): 2015, 2014, 2013, 2012, 2008, 2007, and 2006
3. IEEE International Conference on Data Mining (ICDM): 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, and 2006
4. SIAM Intl. Conference on Data Mining (SDM): 2015, 2014, 2013, 2012, 2011, and 2010
5. Intl. Conference on Information and Knowledge Management (CIKM): 2015, 2014, 2012, and 2010
6. IEEE International Conference on Data Engineering (ICDE): 2013, 2010, 2009 (demos committee), and 2007
7. International Conference on Very Large Databases (VLDB): 2014, and 2010 (industrial track)
8. ACM International Conference on Management of Data (SIGMOD): 2013
9. Intl. Conf. on Extending Database Technology (EDBT): 2006 (demos committee)
10. International Symposium on Spatial and Temporal Databases (SSTD): 2015, 2013, 2009, and 2007

11. International World Wide Web Conference (WWW): 2015 (web mining track), and 2013
12. ACM International Conference on Web Search and Data Mining (WSDM): 2015, 2014, and 2013
13. International Joint Conference on Artificial Intelligence (IJCAI): 2016
14. IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM): 2015, and 2012
15. IEEE International Conference on Big Data (BigData): 2015 (industry track), and 2013 (industry track)
16. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD): 2010, 2009
17. 1st Workshop on Spatial and Spatio-temporal Data Mining (SSTDM): 2006
18. International Workshop on Mining and Learning with Graphs (MLG): 2009
19. Intl. ACM Workshop on Data Engineering for Wireless and Mobile Access (MobiDE): 2013

Journal  
referee

1. IEEE Transactions on Knowledge and Data Engineering (TKDE)
2. IEEE Transactions on Knowledge Discovery from Data (TKDD)
3. Knowledge and Information Systems (KAIS)
4. Data Mining and Knowledge Discovery (DAMI)
5. International Journal on Very Large Databases (VLDBJ)
6. IEEE Transactions on Parallel and Distributed Systems (TPDS)
7. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)

Other

1. NSF panelist: 2016, 2015 (twice), 2013, 2012, 2009, and 2007.
2. Reviewer for Google University Awards: 2012 and 2011.
3. NSF workshop participant, on Sustainable Energy-Efficient Data Management (SEEDM), 2011.

---

## OTHER

---

### Open-source software

XMLlab	<a href="https://github.com/spapadim/xmllab">https://github.com/spapadim/xmllab</a> Web-based interactive XPath and XSLT processor, for teaching semi-structured data management.
HCC	<a href="http://bitquill.net/trac/wiki/HCC/Start">http://bitquill.net/trac/wiki/HCC/Start</a> Hierarchical Co-Clustering, with GUI (in Matlab).
PCC	<a href="http://bitquill.net/trac/wiki/PCC/Start">http://bitquill.net/trac/wiki/PCC/Start</a> Distributed Co-Clustering with MapReduce (in Java).

### Funding experience

Participated in proposal preparation to the NSF and to the AFOSR, in 2015 and 2013.

Participated in proposal for multi-institution Network Science Collaborative Technology Alliance (NS-CTA); proposal was subsequently funded (2010, IBM portion \$200K).

IBM liaison (with responsibility to write IBM-internal parts of proposal) for IBM Faculty Awards (renewable annually) for Christos Faloutsos (four years, 2006–9, \$20–40K each).

Co-authored Pennsylvania Infrastructure Technology Alliance (PITA) proposals (*Automatic Mining on Sensor Data*, 2003–4, \$39K, and *Distributed Mining in Co-evolving Streaming Sensor Data*, 2005–6, \$64K).

**References**      Available upon request.

Jersey City, June 2016