Donghui Zhang

Assistant Professor
College of Computer & Information Science, Northeastern University
360 Huntington Avenue #202WVH, Boston, MA 02115
donghui@ccs.neu.edu http://www.ccs.neu.edu/home/donghui
Phone: (617)373-2177 Fax: (617)373-5121

OBJECTIVE:

Finance, quantitative, database, and web related research and development position.

CURRENT POSITION:

Assistant Professor, College of Computer & Information Science, Northeastern University (September 2002 - present).

RESEARCH INTEREST:

Indexing and query optimization in spatial, temporal, and spatio-temporal databases. Security & Privacy. Data Mining.

EDUCATION:

- **Ph.D.:** Computer Science, University of California–Riverside (Riverside, CA), August 2002. Thesis: "Aggregation Computation over Complex Objects." Advisor: Prof. Vassilis J. Tsotras.
- **M.S.:** Computer Application, Jilin University, P.R.China, July 1995.
- **B.S.:** Computer Software, Jilin University, P.R.China , July 1992.

SKILLS

I am skillful in C++, Java, and web-database application development. In 1994 I designed and implemented a test problem database software in C++. The whole project was 65,000 lines of C++ code and I contributed 35,000. The software won an Excellent Scientific Product Award issued by the Chinese Educational Committee. From 1996 to 2002 I used C++ to conduct research. Since 2002 I have been using Java to conduct research. In 1999 I developed multiple database-intensive web sites for the University of California - Riverside. At least one is still in use: UC Riverside Academic Employees (*http://www.facultydirectory.ucr.edu/*). I have been teaching database application development and kernel technologies in Northeastern University since 2002. I have strong quantitative skills. Recently I have been studying finance. Below is a selected list of my skills.

- Programming Languages: Java, C++, Perl.
- Web development: JSF/JSP, ASP.NET, AJAX, CGI.

- Database application development: JDBC, Java DB, MySQL, Oracle, SQL Server.
- Database kernel technologies: disk-based indexing, query optimization, spatial/temporal/stream data.
- Modeling: UML, ER Model and normal forms.
- IDE: MyEclipse.
- Math: Calculus, probability, statistics.
- Finance: options, futures, and other derivatives.
- Communication: oral and written.

TEACHING:

- Databases (graduate level)
  - COM3315: Principles of Database Systems (Fall'02, Spring'03)
  - CSG130: Introduction to Database Systems (Fall'05, Fall'06, Fall'07, Fall'08)
  - CSG131: Implementation of Database Systems (Spring'06, Spring'07)
- Databases (undergraduate level)
  - COM1315: Introduction to Database Systems (Winter'03)
  - CSU430: Database Design (Spring'04)
- Spatial Databases (graduate level)
  - CSG339: Special Topics in Database Management – Spatial Databases (Spring'08)
- Data Mining (graduate level)
  - CSG339: Special Topics in Database Management – Data Mining (Fall'03)
  - CSG230: Data Mining Techniques (Fall'04, Spring'07, Spring'08, Fall'08)
- Data Mining (undergraduate level)
  - CSU900: Computer Science Topics – Data Mining Techniques (Summer'05)
- Systems and Networks
  - CSU480: Systems and Networks (Spring'06)

PROFESSIONAL ACTIVITIES:

- **NSF Panel member** 2003, 2005.
- **PC Co-Chair:** International Symposium on Large Spatio-Temporal Databases (**SSTD**), 2007.
- **Program Committee member** for
  - International Conference for Very Large Databases (**VLDB**), 2005, 2008.
  - International Conference on Data Engineering (**ICDE**), 2004, 2007, 2008, 2009.

  - International Conference on Extending Database Technology (**EDBT**), 2004.
  - International Conference on Data Mining (**ICDM**), 2008.
  - ACM International Conference on Information and Knowledge Management (**CIKM**), 2006.
  - International Conference on Web Information Systems Engineering (**WISE**), 2006.
  - ACM International Symposium on Advances in Geographic Information Systems (**GIS**), 2007, 2008.
  - International Conference on Mobile Data Management (**MDM**), 2007, 2008.
  - International Conferences on Asia-Pacific Web Conference (**APWeb**), 2009.
  - ACM International Workshop on Data Engineering for Wireless and Mobile Access (**MobiDE**), 2006, 2007.
  - International Workshop on Spatio-Temporal Database Management (**STDBM**), 2006.
  - Workshop on Mining Spatial and Spatio-temporal Data (**SSTDM**), 2006, 2007.
  - **VLDB** Ph.D. Workshop, 2005, 2006, 2007.
  - SIGMOD Ph.D. Workshop on Innovative Database Research (**IDAR**), 2007, 2008.
  - International Workshop on Privacy-Aware Location-based Mobile Services (**PALMS**), 2007.
  - Workshop on Data Mining of Uncertain Data (**DMU**), 2007.

- **Journal referee** for
  - ACM Transactions on Database Systems (**TODS**), 2004, 2007, 2008.
  - VLDB Journal (**VLDBJ**), 2003, 2006, 2007, 2008.
  - IEEE Transactions on Knowledge and Data Engineering (**TKDE**), 2003, 2004, 2005, 2006, 2007, 2008.
  - ACM Transactions on Information Systems (**TOIS**), 2003, 2004, 2005.
  - Information Sciences (**INS**), 2003, 2004, 2005.
  - Information Systems (**IS**), 2004, 2005, 2006, 2007, 2008.
  - GeoInformatica, 2007.
  - Data & Knowledge Engineering (**DKE**), 2006, 2007.
  - Journal of Computer Science and Technology (**JCST**), 2004, 2005, 2006, 2007.
  - Journal of Electronics and Telecommunications Research Institute (**ETRI**), 2005, 2008.
  - Journal of Systems and Software (**JSS**), 2004.
  - Information Processing Letters (**IPL**), 2003, 2005.
  - Data Mining and Knowledge Discovery, 2006.
  - Computer Standards & Interfaces, 2007.
  - Computer Journal, 2008.

- **Grant proposal referee** for Research Grants Council of Hong Kong (**RGC**), 2004, 2005, 2008.
- Local Arrangement Chair for the 20^{th} International Conference on Data Engineering (**ICDE**), 2004.

- Registration Chair for the 7th Int. Symposium on Spatial and Temporal Databases (**SSTD**), 2001.
- Session Chair for ICDE'04, VLDB'05, ICDE'08.

UNIVERSITY ACTIVITIES

- M.S. Committee, 2003
- Ph.D. Committee, 2004-2007

PUBLICATIONS:

To download, visit http://www.ccs.neu.edu/home/donghui/download.htm

Book Chapters:

- **D. Zhang** and Y. Du, "**Resource Allocation Problems in Spatial Databases**", *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (editors), Springer, http://refworks.springer.com/database-systems, to appear.
- **D. Zhang**, K. P. Baclawski, and V. J. Tsotras, "**B+-tree**", *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (editors), Springer, http://refworks.springer.com/database-systems, to appear.
- **D. Zhang**, Y. Manolopoulos, Y. Theodoridis, and V. J. Tsotras, "**Extendible Hashing**", *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (editors), Springer, http://refworks.springer.com/database-systems, to appear.
- **D. Zhang**, Y. Manolopoulos, Y. Theodoridis, and V. J. Tsotras, "**Linear Hashing**", *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (editors), Springer, http://refworks.springer.com/database-systems, to appear.
- **D. Zhang**, "**Fastest-Path Computation**", *Encyclopedia of GIS*, S. Shekhar and H. Xiong (editors), Springer, ISBN: 978-0-387-30858-6, http://refworks.springer.com/geograph, 2008.
- **D. Zhang**, "**Spatial Aggregation Query**", *Encyclopedia of GIS*, S. Shekhar and H. Xiong (editors), Springer, ISBN: 978-0-387-30858-6, http://refworks.springer.com/geograph, 2008.
- M. Hadjieleftheriou, A. N. Papadopoulos, and **D. Zhang**, "**Disk Storage and Basic File Structures**", Chapter 19 of *Handbook of Database Technologies*, J. Hammer and M. Schneider (editors), CRC Press, 2007.
- **D. Zhang**, "**B Trees**", Chapter 15 of *Handbook of Data Structures and Applications*, D. P. Mehta, S. Sahni (editors), Chapman & Hall/CRC, ISBN 1-5848-8435-5, 2004.
- B. Salzberg and **D. Zhang**, "**Access Methods**", Chapter 54 of *The Computer Science and Engineering Handboook*, Second Edition, A. B. Tucker, Jr. (editor), Chapman & Hall/CRC, ISBN 1-58488-360-X, 2004.

Refereed Publications:

2008
- **D. Zhang**, A. Markowetz, V. J. Tsotras, D. Gunopulos, and B. Seeger, "**On Computing Temporal Aggregates with Range Predicates**", *ACM Transactions on Database Systems (***TODS***), to appear.*
- L. Hu, Y.-C. Chang, C. A. Lang, K. A. Ross, and **D. Zhang**, "QueryScope: Visualizing Queries for Repeatable Database Tuning (Demo Paper)", *Proc. of 34th International Conference on Very Large Databases (***VLDB***)*, Auckland, New Zealand, 2008. *(Acceptance rate: 27.5%)*
- D. Kunkle, **D. Zhang**, and G. Cooperman, "**Mining Frequent Generalized Itemsets and Essential Generalized Association Rules without Redundancy**", *Journal of Computer Science and Technology (***JCST***)*, 23(1): 77-102, Jan. 2008.
- **D. Zhang**, Y. Du, and L. Hu, "**On Monitoring the top-k Unsafe Places**", *Proc. of 24th International Conference on Data Engineering (***ICDE***)*, Cancun, Mexico, 2008. *(Acceptance rate for full paper with short presentation: 19.2%)*
- Y. Tao, X. Xiao, J. Li, and **D. Zhang**, "**On Anti-Corruption Privacy Preserving Publication**", *Proc. of 24th International Conference on Data Engineering (***ICDE***)*, Cancun, Mexico, 2008. *(Acceptance rate for full paper with short presentation: 19.2%)*
- T. Xia, **D. Zhang**, and Y. Tao, "**On Skylining with Flexible Dominance Relation (poster paper)**", *Proc. of 24th International Conference on Data Engineering (***ICDE***)*, Cancun, Mexico, 2008. *(Acceptance rate: 31%)*

2007
- J. M. Kang, M. F. Mokbel, S. Shekhar, T. Xia, and **D. Zhang**, "**Continuous Evaluation of Monochromatic and Bichromatic Reverse Nearest Neighbors**", *Proc. of 23rd International Conference on Data Engineering (***ICDE***)*, Istanbul, Turkey, 2007. *(Acceptance rate: 18.5%)*
- Y. Du, T. Xia, Y. Tao, **D. Zhang**, and F. Zhu, "**On Multidimensional k-Anonymity with Local Recoding Generalization (poster paper)**", *Proc. of 23rd International Conference on Data Engineering (***ICDE***)*, Istanbul, Turkey, 2007. *(Acceptance rate: 27.6%)*

2006
- **D. Zhang**, Y. Du, T. Xia, and Y. Tao, "**Progressive Computation of The Min-Dist Optimal-Location Query**", *Proc. of 32nd International Conference on Very Large Data Bases (***VLDB***)*, pages 643-654, Seoul, Korea, 2006. *(Acceptance rate: 13.9%)*
- T. Xia and **D. Zhang**, "**Refreshing the Sky: The Compressed Skycube with Efficient Support for Frequent Updates**", *Proc. of 25th ACM/SIGMOD Annual Conference on Management of Data (***SIGMOD***)*, pages 491-502, Chicago, IL, 2006. *(Acceptance rate: 13.0%)*
- E. Kanoulas, Y. Du, T. Xia, and **D. Zhang**, "**Finding Fastest Paths on A Road Network with Speed Patterns**", *Proc. of 18th International Conference on Data Engineering (***ICDE***)*, Atlanta, Georgia, 2006. *(Acceptance rate: 19.5%)*
- T. Xia and **D. Zhang**, "**Continuous Reverse Nearest Neighbor Monitoring**", *Proc. of 18th International Conference on Data Engineering (***ICDE***)*, Atlanta, Georgia, 2006. *(Acceptance rate: 19.5%)*
- S.-Y. Chien, V. J. Tsotras, C. Zaniolo, and **D. Zhang**, "**Supporting Complex Queries on Multiversion XML Documents**", *ACM Transactions on Internet Technology (***TOIT***)*, 6(1): 53-84, 2006.

– D. Kunkle, **D. Zhang** and G. Cooperman, "**Efficient Mining of Max Frequent Patterns in a Generalized Environment (poster paper)**", *Proc. of 15th ACM Conference on Information and Knowledge Management (***CIKM***)*, Arlington, VA, 2006. *(Acceptance rate: 25%)*

2005
– **D. Zhang** and V. J. Tsotras, "**Optimizing Spatial Min/Max Aggregations**", *VLDB Journal*, 14(2): 170-181, 2005.

– T. Xia, **D. Zhang**, E. Kanoulas, and Y. Du, "**On Computing Top-t Most Influential Spatial Sites**", *Proc. of 31th International Conference on Vary Large Data Bases (***VLDB***)*, Trondheim, Norway, 2005. *(Acceptance rate: 16.5%)*

– Y. Du, **D. Zhang** and T. Xia, "**The Optimal Location Query**", *Proc. of 9th International Symposium on Spatial and Temporal Databases (***SSTD***)*, pages 163-180, Angra dos Reis, Brazil, 2005. *(Acceptance rate: 31.2%)*

– P. Zhou, **D. Zhang**, B. Salzberg, G. Cooperman, and G. Kollios, "**Close Pair Queries in Moving Object Databases**", *Proc. of 13th ACM International Symposium on Advances in Geographic Information Systems (***GIS***)*, pages 2-11, Bremen, Germany, 2005. *(Acceptance rate: 33%)*

– T. Xia and **D. Zhang**, "**Improving the R*-tree with Outlier Handling Techniques**", *Proc. of 13th ACM International Symposium on Advances in Geographic Information Systems (***GIS***)*, pages 125-134, Bremen, Germany, 2005. *(Acceptance rate: 33%)*

2004
– H. Wu, B. Salzberg and **D. Zhang**, "**Online Event-driven Subsequence Matching over Financial Data Streams**", *Proc. of ACM/SIGMOD Annual Conference on Management of Data (***SIGMOD***)*, Paris, France, 2004. *(Acceptance rate: 16.0%)*

– **D. Zhang** and T. Xia, "**A Novel Improvement to the R*-tree Spatial Index using Gain/Loss Metrics**", *Proc. of 12th ACM International Symposium on Advances in Geographic Information Systems (***GIS***)*, pages 204-213, Washington D.C., 2004. *(Acceptance rate: 35%)*

2003
– **D. Zhang**, D. Gunopulos, V. J. Tsotras, and B. Seeger, "**Temporal and Spatio-Temporal Aggregations over Data Streams using Multiple Time Granularities**", *Journal of Information Systems*, vol. 28, no. 1-2, pages 61-84, 2003.

– J. Shan, **D. Zhang** and B. Salzberg, "**On Spatial-Range Closest-Pair Query**", *Proc. of 8th International Symposium on Spatial and Temporal Databases (***SSTD***)*, pages 252-269, Santorini Island, Greece, 2003. *(Acceptance rate: 27%)*

2002
– S.-Y. Chien, Z. Vagena, **D. Zhang**, V. J. Tsotras, and C. Zaniolo, "**Efficient Structural Joins on Indexed XML Documents**", *Prof. of 28th International Conference on Very Large Data Bases (***VLDB***)*, Hong Kong, China, 2002. *(Acceptance rate: 18%)*

– **D. Zhang**, V. J. Tsotras and D. Gunopulos, "**Efficient Aggregation over Objects with Extent**", *Proc. of 21th ACM International SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (***PODS***)*, Madison, Wisconsin, 2002. *(Acceptance rate: 22%)*

– **D. Zhang**, V. J. Tsotras and B. Seeger, "**Efficient Temporal Join Processing using Indices**", *Proc. of 18th International Conference on Data Engineering (***ICDE***)*, San Jose, California, 2002. *(Acceptance rate: 19%)*

– **D. Zhang**, D. Gunopulos, V. J. Tsotras, and B. Seeger, "**Temporal Aggregation over Data Streams using Multiple Granularities**", *Proc. of 8th International Conference on Extending Database Technology (***EDBT***)*, Prague, Czech Republic, 2002. *(Acceptance rate: 17%)*

- S.-Y. Chien, V. J. Tsotras, C. Zaniolo, and **D. Zhang**, "**Efficient Complex Query Support for Multiversion XML Documents**", *Proc. of 8th International Conference on Extending Database Technology (***EDBT***)*, Prague, Czech Republic, 2002. *(Acceptance rate: 17%)*
- **D. Zhang** and V. J. Tsotras, "**Index Based Processing of Semi-Restrictive Temporal Joins**", *Proc. of 9th International Symposium on Temporal Representation and Reasoning (***TIME***)*, Manchester, UK, 2002. *(Acceptance rate: 34.8%)*
- **D. Zhang**, "**Advanced Database Aggregation Query Processing**", *Proc. of Extending Database Technology (***EDBT***) Ph.D. Workshop*, Prague, Czech Republic, 2002.

2001

- **D. Zhang**, A. Markowetz, V. J. Tsotras, D. Gunopulos, and B. Seeger, "**Efficient Computation of Temporal Aggregates with Range Predicates**", *Proc. of 20th ACM International SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (***PODS***)*, Santa Barbara, California, 2001. *(Acceptance rate: 26.3%)*
- **D. Zhang** and V. J. Tsotras, "**Improving Min/Max Aggregation over Spatial Objects**", *Proc. of 9th ACM International Symposium on Advances in Geographic Information Systems (***GIS***)*, Atlanta, Georgia, 2001. *(Acceptance rate: 30.5%)*
- S.-Y. Chien, V. J. Tsotras, C. Zaniolo, and **D. Zhang**, "**Storing and Querying Multiversion XML Documents using Durable Node Numbers**", *Proc. of 2nd International Conference on Web Information Systems Engineering (***WISE***)*, Kyoto, Japan, 2001. *(Acceptance rate: 32%)*

Invited Publications:

- **D. Zhang**, "NEUStore: A Simple Java Package for the Construction of Disk-based, Paginated, and Buffered Indices", *ODBMS.ORG – Object Database Management Systems: The Resource Portal for Education and Research*, 2005.
- Z. Chen, C. Li, J. Pei, Y. Tao, H. Wang, W. Wang, Jiong Yang, Jun Yang, and **D. Zhang**, "Recent Progress on Selected Topics in Database Research – A Report from Nine Young Chinese Researchers Working in the United States", *Journal of Computer Science and Technology, P. R. China*, 2003.

GRANTS:

- PI for "CAREER: Fast Query Support for Emerging Spatial Database Applications". Grant duration: 7/1/04-6/30/09, National Science Foundation (NSF), CAREER Award. Grant amount: $506,000.
  URL: http://www.ccs.neu.edu/home/donghui/research/CAREER_grant/
- PI for the submitted proposal "III-CXT: Tools and Techniques for Privacy Enhancement and (Street and Cyber) Crime Prevention", co-PIs Jennifer Robinson and Ravi Sundaram, submitted to NSF-IIS-III.
- Participated in "IGERT: INTELLIGENT DIAGNOSTICS for Aging Civil Infrastructure". Sara Wadia-Fascetti (PI) with Bernal (CEE), Bluestone (Urban and Regional Policy), Rappaport (ECE), Padilla (CEE - Puerto Rico). Funded by the National Science Foundation (Award #DGE-0654176). September 1, 2007 - August 31, 2012.

HONORS AND AWARDS:

- CAREER Award, National Science Foundation, 2004.
- Graduate Dean's Dissertation Grant Award, University of California, Riverside, 2001.
- GSA Mini-Grant, University of California, Riverside, 2001.
- Phi Beta Kappa Alumni Scholarship Award, 2000.
- Excellent Attitude Award, Center for Visual Computing, University of California, Riverside, 1999.
- Excellent Scientific Product Certification, Chinese Educational Committee, China, 1994.

RECENT TALKS:

- Algorithms in Spatial Databases (a short course)
  - Renmin University, P. R. China, May 26 - June 1, 2007.
- The Skyline Query in Databases: Which Objects are the Most Important?
  - CCIS Ph.D. seminar, March 22, 2007.
- The Min-Dist Optimal-Location Query
  - 32nd International Conference on Very Large Data Bases (VLDB), Seoul, Korea, 2006.
  - UMass Boston, October 17, 2006.
  - CCIS Ph.D. Seminar, November 30, 2006.
  - UMass Lowell, December 6, 2006.
- Introduction to Spatial Databases
  - NEU ACM Talk, Feburary 21, 2007.
  - CCIS Ph.D. Open House, Mar 23, 2007.
- The optimal-location query
  - 9th International Symposium on Spatial and Temporal Databases (SSTD), Angra dos Reis, Brazil, 2005.
  - CCIS Ph.D. seminar, November 3, 2005.
  - Boston University, April 19, 2006.
  - WPI, April 28, 2006.
- Spatial Database Query Processing using Indices
  - CCIS Ph.D. Seminar, October 21, 2004.
- On Spatial-Range Closest-Pair Query
  - UMass Boston, March 12, 2003.
- ER-Miner: A New Method to Mine Essential Rules and Constrained Essential Rules
  - CCIS Ph.D. Seminar, 2003.

STUDENTS:

- Ph.D. thesis advisor:
  - Tian Xia, 2007. Thesis: "Subspace and Relaxed Skyline Query Processing". Joined Oracle.
  - Yang Du, 2008. Thesis: "Optimal-Location Queries in Spatial Databases".
  - Ling Hu.
  - Jian Wen.
  - Tianhua Zheng.
  - Lin Zhao
- Ph.D. thesis committee:
  - Xiaowei Sun, Northeastern University, 2004.
  - Rui Wang, Northeastern University, 2006.
  - Panfeng Zhou, Northeastern University, 2006.
  - Xiaoqin Ma, Northeastern University, 2006.
  - Jing Shan, Northeastern University, 2007.
  - Feifei Li, Boston University, 2007.
  - Wei Li, University of Massachusetts, Lowell, 2007.
- Ph.D. thesis external examiner:
  - Hua Lu, National University of Singapore, 2006.
  - Ke Deng, University of Queensland, Australia, 2007.
  - Yidong Yuan, University of New South Wales, Australia, 2007.
- M.S. students:
  - Feizhi Li, 2003. M.S. project: "Using Object-Oriented Databases to Store and Query XML Data".
  - Tilak Palanisamy, 2003. M.S. project: "On Enabling Schema Evolution in a Relational Database System".
  - Madhuri Velidi, 2007, M.S. project: "Web-Database Application Development For Training Blackjack Skills".

MEMBERSHIP:

ACM member and SIGMOD member.

IMMIGRATION STATUS

Permanent Resident of USA