Wei Ding

(as of July 2008)

Department of Computer Science
University of Massachusetts-Boston
Email: Ding@cs.umb.edu, URL: http://www.cs.umb.edu/~ding

Research Interests

My main research interests lie in the field of knowledge discovery and data mining, with applications to astronomy, geosciences, and environmental sciences. My current research projects are on discovering scientifically interesting regions of arbitrary shape and granularity from spatial datasets, on identifying novel spatial associations, and on developing scalable region discovery algorithms to cope with large real-world datasets.

Education

Ph.D., Computer Science, University of Houston, TX, 2008
M.Sc., Software Engineering, George Mason University, VA, 2000
B.E., Computer Science and Applications, Xi'an Jiao Tong University, China, 1993

Honors and Awards

- 2007: NSF Scholarship to attend the Grace Hopper Celebration of Women in Computing
- 2006: Excellent Teaching Piper Award Finalist, University of Houston – Clear Lake
- 2006: Phi Kappa Delta Honor Society
- 2006: Honorable Mention, NSF Graduate Research Fellowship
- 2005: NSF Fellowship to attend the 5th International Summer School on Biocomplexity from System to Gene, Dartmouth College
- 2004 & 2005: CRA-W Graduate Cohort Travel Award
- 2004 – 2006: Computer Science Department Scholarship, University of Houston
- 2001: Academic Excellency Award in Software Engineering, George Mason University
- 2000: Asian Heritage Month Distinction Award, George Mason University
- 1998 – 2000: Graduate Fellowship, George Mason University
- 1992: Excellent Student Leader Award, Xi'an Jiao Tong University, China
- 1989 – 1993: Excellent Student Scholarship, Xi'an Jiao Tong University, China
- 1989: National Entrance Examination Waiver, Xi'an Jiao Tong University, China

Research Experience

Visiting Scientist May 2007 – Aug. 2007
Lunar and Planetary Institute Houston, TX

- Machine detection of sub-kilometer craters in high resolution planetary images
- Data mining in terrestrial datasets

Research Assistant Aug. 2005 – May 2008
Data Mining and Machine Learning Group Houston, TX
Department of Computer Science, University of Houston

- Constructing a region discovery framework to systematically discover regional patterns and apply it to real-world applications in astronomy, geosciences, and environmental sciences: the first

project is on finding feature-based hot spots in multivariate, real-valued datasets. Our method integrates a family of clustering algorithms and is empirically evaluated on a real-world database of ground ice on Mars; the second project is on regional association rule mining and scoping, and the method is applied to the problem of arsenic contamination to identify arsenic risk patterns in the Texas water supply.

- Preparing and writing grant proposals (with Dr. Christoph Eick as PI) to the NSF Information & Intelligent Systems, the NSF Cyber-Enabled Discovery and Innovation, the Texas Advanced Research Program, and the Environmental Institute of Houston.

Summer Internship May 2007 – Aug. 2007
Lunar and Planetary Institute Houston, TX

- For Mars and terrestrial remote-sensing datasets, research work included discovering feature-based hot spots for two or more attributes, finding scientifically interesting regions of arbitrary shape and granularity, designing and implementing data preprocessing techniques to smooth, interpolate, and denoise real-world raster datasets.

Research Assistant Aug. 1999 – Aug. 2000
Information & Software Engineering Department Fairfax, VA
George Mason University

- Research work included using model checkers to test safety properties for critical systems. Specifically, used a formal method of model checking to either generate new test sets or analyze existing test sets with respect to safety properties expressed in a temporal logic, formalized notion of dangerous actions with a mutation model, and developed coverage criteria to assess test sets.

Teaching Experience

Lecturer Jan. 2002 – Aug. 2008
Computer Science and Computer Information Systems Houston, TX
University of Houston-Clear Lake

- Teaching undergraduate and graduate courses in Computer Science and Computer Information Systems; creating new classes in E-Commerce development and advanced web application development; designing and implementing new teaching and assessment methods using ePortfolio, distance learning, and peer evaluation on team project assessment; and preparing ABET (Accreditation Board for Engineering and Technology) accreditation including design of class exit surveys and course outcome assessment.

Courses Taught:

Title	Enrollment *(total=1,443)*	Semester(s)
Web Application Development	915 students	Spring 2008, Fall 2007, Spring 2007, Fall 2006, Spring 2006, Fall 2005, Spring 2005, Fall 2004, Summer 2004, Spring 2004, Fall 2003, Summer 2003, Spring 2003, Fall 2002, Summer 2002, Spring 2002
Design of Database Systems	172 students	Spring 2007, Fall 2006, Spring 2006, Fall 2005, Spring 2005, Summer 2004, Spring 2002
Advanced Web Application Development	35 students	Spring 2008, Fall 2007
E-Commerce Development	39 students	Spring 2006, Fall 2004
Data Structures	69 students	Spring 2004, Fall 2003

Advanced Data Structures and Algorithms	57 students	Spring 2003, Fall 2002
Independent Study in Computer Science	9 students	Summer 2007, Summer 2003
Fundamentals of Database Systems (*invited lectures at Nankai University, Tianjin, China*)	147 students	Fall 2007

Summary of Student Evaluations:
6-year average, Spring 2002 – Fall 2007

The overall quality of the course	The instructor's overall teaching capability	Simulate and challenge students to think and to question
4.25 (out of 5.0)	4.41 (out of 5.0)	4.59 (out of 5.0)

Teaching Assistant Aug. 1998 – Aug. 1999
Information & Software Engineering Department Fairfax, VA
George Mason University

- Assisted homework grading and provided laboratory support for two graduate courses (Software Construction and Software Testing and Maintenance).

IBM DB2 Tutor Nov. 1996 – May 1998
PanSky International Holding Co. Ltd. Beijing, China

- Tutor in IBM DB2 database development and administration.

Submitted Publications

1. T. F. Stepinski, **W. Ding,** C. F. Eick, "Discovering Controlling Factors of Geospatial Variables", 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS 2008), submitted, 2008.

2. C. F. Eick, R. Parmar, **W. Ding,** T. F. Stepinski, J. P. Nicot, "Finding Regional Co-location Patterns for Sets of Continuous Variables", 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS 2008), submitted, 2008.

3. **W. Ding**, T. Stepinski, R. Parmar, D. Jiang, C. F. Eick, "Discovery of Feature-Based Hot Spots in Real-Valued Spatial Databases: an Application to Ground Ice on Mars", the International Journal of Computers and Geosciences, Elsevier, under revision.

4. **W. Ding**, P. Chen, C. Ding, "A Connectionist-based Lexical Knowledge Model", the International Journal of Cognitive Informatics and Natural Intelligence (IJCiNi), under revision.

Book Chapters

5. **W. Ding**, P. Chen, "An Interactive Visualization Model for Large High-Dimensional Datasets", Data Engineering: Mining, Information, and Intelligence, to appear, Editors: Yupo Chan, John Talburt, Terry Talley, Springer, 2008.

6. P. Chen, **W. Ding**, "Knowledge Management for Agent-based Tutoring Systems", Designing Distributed Learning Environments: With Intelligent Software Agents, pp. 146-161, Ed. F. Lin, Idea Group, Inc., 2004.

Refereed Journal Publications

7. K. Yue, A. Yang, **W. Ding,** and P. Chen, "Open Courseware and Computer Science Education", ACM Journal of Computing Sciences in Colleges, Volume 20, Issue 1, Utah, USA, October, 2004.

Refereed Conference Publications

8. **W. Ding**, R. Jiamthapthaksin, R. Parmar, D. Jiang, T. Stepinski, C. Eick, "Towards Region Discovery in Spatial Datasets", in Proc. of the Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD), Osaka, Japan, May, 2008.
   Acceptance ratio: 37/312=11%

9. **W. Ding**, C. Eick, X. Yuan, J. Wang, J.P. Nicot, "On Regional Association Rule Scoping", in Proc. of the International workshop on Spatial and Spatio-temporal Data Mining in Cooperation with IEEE ICDM 2007, Omaha, NE, USA, October, 2007.
   Acceptance ratio: 10/35=28%

10. **W. Ding**, C. Eick, J. Wang, X. Yuan, "A Framework for Regional Association Rule Mining in Spatial Datasets", in Proc. of the 6th IEEE International Conference on Data Mining (IEEE-ICDM'06), Hong Kong, China, December, 2006.
    Acceptance ratio: 152/776=19%

11. P. Chen, **W. Ding**, C. Ding, "SenseNet: A Knowledge Representation Model for Computational Semantics", in Proc. of the 5th IEEE International Conference on Cognitive Informatics (ICCI), Beijing, China, July, 2006.
    Acceptance ratio: 40/276 = 14%

12. I.A. Kakadiaris, I. Konstantinidis, E. Papadakis, **W. Ding**, D.J. Kouri, and D.K. Hoffman, "Parametric Surface Denoising", in Proc. of SPIE Wavelets XI, E. Papadakis, A. Laine, M. Unser (Eds), San Diego, CA, USA, July, 2005.
    Acceptance ratio: N/A

13. X. Wang, P. Chen, and **W. Ding**, "Web-based Interactive Visualization of Data Cubes", in Proc. the 2005 International Conference on Modeling, Simulation and Visualization Methods (MSV'05), Las Vegas, USA, June, 2005.
    Acceptance ratio: 35%

14. G. Boetticher, **W. Ding**, C. Moen, and K. Yue, "Using a Pre-Assessment Exam to Construct an Effective Concept-based Genetic Program for Predicting Course Success", In Proc. of the 36th SIGCSE Technical Symposium on Computer Science Education (ACM SIGCSE'05), pp. 500 – 504, St. Louis, Missouri, USA, Feb. 2005.
    Acceptance ratio: 104/330=32%

15. K. Yue, **W. Ding**, "Design and Evolution of an Undergraduate Course on Web Application Development", in Proc. of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education (ACM ITiCSE'04), pp. 22-26, Leeds, UK, June, 2004.
    Acceptance ratio: 46/155=29%

16. K. Yue, A. Yang, **W. Ding**, and P. Chen, "A Model for Open Content Communities to Support Effective Learning and Teaching ", in Proc. of the IADIS International Conference on Web-based Communities, pp. 533-536, Lisbon, Portugal, April 2004.
    Acceptance ratio: N/A

17. P. Chen, C. Hu, **W. Ding**, and H. Lynn, "Icon-based Visualization of Large High-Dimensional Datasets", in Proc. of the 3rd IEEE International Conference on Data Mining (ICDM'03), pp. 505-508, Melbourne, Florida, Nov. 2003.
    Acceptance ratio: 128/501=25%

18. P. Ammann, **W. Ding**, and D. Xu, "Using a Model Checker to Test Safety Properties", in Proc. of the 7th IEEE International Conference on Engineering of Complex Computer Systems, pp. 212-221, Skovde, Sweden, June 2001.
    Acceptance ratio: 45%

19. A. Abdurazik, P. Ammann, **W. Ding**, and J. Offutt, "Evaluation of Three Specification-based Testing Criteria", in Proc. of the 6th IEEE International Conference on Engineering of Complex Computer Systems, pp. 179-187, Tokyo, Japan, Sept. 2000.
    Acceptance ratio: N/A

Other Publications

20. **W. Ding**, C. Eick, "Mining Regional Knowledge in Spatial Datasets", in Proc. of Grace Hopper Celebration of Women in Computing, Orlando, FL, October 2007.

Internally Sponsored Research Grants

Total Internal Research Grants: $25,095.60

1. "Computer-aided Detection of Sub-Kilometer Craters in High Resolution Planetary Images", PI of a collaborative project awarded by the Institute for Pace Systems Operations (ISSO), Texas, $9,936, 5/2008-8/2008.
2. "Towards Region Discovery in Spatial Datasets", Faculty Development Fund, University of Houston-Clear Lake, PI, $1,876.60, 2008.
3. "Integrating Supervised and Adaptive Learning to improve Text Entry for People with Motion Impairments", Faculty Research and Support Fund, University of Houston-Clear Lake, PI, $4,500, 2007.
4. "Automatic Detection and Correction of Spelling Errors Using Knowledge Modeling", Faculty Research and Support Fund, University of Houston-Clear Lake, PI, $2,912, 2007.
5. "Developing of a Large Commonsense Knowledge Acquisition Software System", UHCL Alumni Association Program Endowment Award, University of Houston-Clear Lake, PI, $400, 2007.
6. "On Regional Association Rule Scoping", Faculty Development Fund, University of Houston-Clear Lake, PI, $1,300, 2007.
7. "A Framework for Regional Association Rule Mining in Spatial Datasets", Faculty Development Fund, University of Houston-Clear Lake, PI, $2,000, 2006.
8. "SenseNet: A Knowledge Representation Model for Computational Semantics", Faculty Development Fund, University of Houston-Clear Lake, PI, $850, 2006.
9. "Design and Evolution of an Undergraduate Course on Web Application Development", Faculty Development Fund, University of Houston-Clear Lake, PI, $1,321, 2004

Invited and Conference Presentations

1. "Towards region discovery in spatial datasets", Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD), Osaka, Japan, May 2008
2. "Discovering regional knowledge from spatial datasets", Natural Science Seminar, University of Houston-Clear Lake, January 2008.

3. “Discovering Regional Patterns”, College of Software, Nankai University, Tianjin, China, December 2007.
4. “Mining Regional Knowledge in Spatial Datasets”, Grace Hopper Celebration of Women in Computing, Orlando, FL, October 2007.
5. “On Regional Association Rule Scoping”, International workshop on Spatial and Spatio-temporal Data Mining in Cooperation with IEEE ICDM 2007, Omaha, NE, USA, October 2007.
6. “A Framework for Regional Association Rule Mining in Spatial Datasets”, the 6th IEEE International Conference on Data Mining (IEEE-ICDM'06), Hong Kong, China, December 2006.
7. "SenseNet: A Knowledge Representation Model for Computational Semantics", the 5th IEEE International Conference on Cognitive Informatics (ICCI), Beijing, China, July 2006.
8. “Design and Evolution of an Undergraduate Course on Web Application Development”, the 9th Annual ACM SIGCSE Conference on Innovation and Technology in Computer Science Education, University of Leeds, UK, June 2004.

Industrial Experience

Software Engineer & Technical Consultant Nov. 2000 – Jan. 2002
VeriSign, Inc. Herndon, VA

- Developed Java and XML APIs using OO design patterns to communicate with remote servers using TCP/IP secure-socket communication
- Generated data reports from an Oracle database using Java, JDBC, Unix Shell Script and Perl
- Implemented Java APIs for wholesale partners
- Participated in domain-name registration system maintenance and performance tuning

QA Team Leader & Senior Software Engineer Aug. 2000 – Nov. 2000
MultiCity.com Vienna, VA

- Conducted testing on web applications in J2EE
- Designed and implemented testing standards and procedures
- Managed the QA team and analyzed testing results
- Developed software tools to test, track, and verify defects in web applications
- Tools used for the above were Java (JDK1.x, Java Servlet, RMI, JDBC), MySQL, Visual Café, Apache, Perforce, and Talisma

Project Manager & Senior Software Engineer Nov. 1996 – May 1998
PanSky International Holding Co. Ltd. Beijing, China

- Developed web applications for company website using Java (JDK1.0)
- Acted as a consultant for UDB and DB2/400 connection, SQL and CICS/6000 programming, ODBC (DB2/400), performance tuning of DB2/400, CICS communication via SDLC and APPC LU6.2 across AIX and VSE/ESA
- Designed Y2K solutions for a major accounting system of the Agricultural Bank of China in RPG III/400 on OS/400 3.2
- Designed and implemented integration- and acceptance-testing of a main accounting system of ShanDong Power Co. in RPG III/400
- Developed a communication system between AS/400, RS/6000, Sun workstation, and Windows via MQSeries 5.0
- Designed an IC Card system for Tianjin University
- Provided pre-sale technical support on Java (VisualAge) and DB2 UDB for the Foreign Exchange Department of the Bank of China

Testing Engineer Aug. 1996 – Oct. 1996
Microsoft (China) Ltd. Beijing, China
Designed testing requirements and performed alpha testing of PowerPoint (Chinese version) on Windows 95 and NT 4.0.

Software Engineer July 1993 – July 1996
Bank of China Yantai, China

- Chief developer of the major transaction processing systems of the Bank of China, which include financial data manipulation and analysis, transaction data sort, update and search, financial data report generation and telegram transmissions using DOS/VS COBOL 3.1, SQL(DL/I) and CICS 2.3 on VSE/ESA 1.3.3 at ES/9000-150
- Chief developer of the Call Center System of the Bank of China, using C, COBOL, Informix, CICS 2.3, VRU voice unit, and APPC LU 6.2 protocol across ES/9000-150 and RS/6000

Students Directed

1. Josue Salazar, "Spatial Data Mining on Terrestrial Datasets", Research Assistant, University of Houston–Clear Lake, 2008
2. Simmons, Timothy R, "Question Answering Based On Structural Matching" , Research Assistant, University of Houston-Clear Lake, 2008
3. Srikanth Kotagiri, "Integrating Supervised and Adaptive Learning to improve Text Entry for People with Motion Impairments", Graduate Research Assistant, University of Houston–Clear Lake, 2008
4. Sai Srinivas Pabbathi, "A Study on an ePortfolio Pilot Project", Graduate Independent Study, University of Houston–Clear Lake, 2008
5. Thinh Nguyen, "E-Commerce Web Site Development on a Real-World Case Study", Undergraduate Independent Study, University of Houston–Clear Lake, 2008
6. Ashutosh Raval, "Integrating Supervised and Adaptive Learning to improve Text Entry for People with Motion Impairments", Graduate Research Assistant, University of Houston–Clear Lake, 2008
7. Anurag Nagar, "Biomedical Information Retrieval using Multiple Ontologies", Graduate Thesis Committee member, University of Houston–Clear Lake, 2008
8. Michael Baldauf, "Domain Driven Causal Financial Engineering in the Context of Evolutionary Computing", Graduate Thesis Committee member, University of Houston–Clear Lake, 2008
9. Timothy R. Simmons, "Search Engine with Natural Language Interface", Undergraduate Independent Study, University of Houston–Clear Lake, Spring 2008
10. Georgi S. Tanev, "Design of Database Systems", Undergraduate Independent Study, University of Houston–Clear Lake, Spring 2008
11. Sasya Kodali, "Advanced Web Application Development Using ASP.NET and ADO.NET Framework", Graduate Independent Study, University of Houston–Clear Lake, Summer 2007
12. Insia Iftiqhar, "Lexical Knowledge Acquisition and Representation", Graduate Research Assistant and Graduate Independent Study, University of Houston–Clear Lake, Summer 2007
13. Sukumar Bollineni, "Advanced Web Application Development Using J2EE", Graduate Independent Study, University of Houston–Clear Lake, Summer 2007
14. Srinivas Veesam, "Web Application Development Using AJAX", Graduate Independent Study, University of Houston–Clear Lake, Summer 2007
15. Suman Tedla, "Web Bias", Graduate Thesis Committee member, University of Houston–Clear Lake, 2006
16. Tuan Anh Nguyen, Graduate Thesis Committee member, "Evaluations of Secure MANET Routing Protocols in Malicious Environments", University of Houston–Clear Lake, 2005
17. Karthik Sadasivam, Graduate Thesis Committee member, "Performance and Security in a Distributed Wireless Networking Environment", University of Houston–Clear Lake, 2004
18. Kerry K. Lawson, Graduate Thesis Committee member, "Website Re-engineering", University of Houston–Clear Lake, 2004

19. Sreenivasan Alakappan, Graduate Thesis Committee member, "A Framework of Inlining Algorithms for Mapping XML Document Type Definitions to Relational Schemas", University of Houston–Clear Lake, 2003
20. Hatal Mahendra Malkan, Undergraduate Independent Study, "Computer-Aided Library System", University of Houston–Clear Lake, Summer 2003

Professional Service

- 2008: PC member of the International Workshop on Spatial and Spatiotemporal Data Mining (SSTDM-08) In Cooperation with IEEE ICDM 2008
- 2008: PC member of the 20th International Conference on Tools with Artificial Intelligence, Dayton, Ohio, USA
- 2008: PC member of the 17th International Conference on Software Engineering and Data Engineering, Los Angeles, CA, USA
- 2007: Session Chair for the IEEE International Workshop on Spatial and Spatio-temporal Data Mining in cooperation with IEEE ICDM 2007, Omaha, NE, USA

Reviewer:
- 2007: IEEE International Workshop on Spatial and Spatio-temporal Data mining in cooperation with IEEE ICDM
- 2007: International Conference on Machine Learning and Data Mining (MLDM)
- 2007: Springer Volume on "Data Engineering"
- 2006 & 2007: IEEE International Conference on Data Mining (ICDM)
- 2005: The Automated Verification of Critical Systems of the Journal Formal Aspects of Computing (FAC)
- 2005: The Internet Encyclopedia, John Wiley & Sons, Inc.
- 2005 – Present: ACM Special Interest Group on Computer Science Education (SIGCSE)
- 2005 – Present: ACM Integrating Technology into Computer Science Education (ITiCSE)
- 2005 – Present: Journal of Information Systems Education (JISE)
- 2004: The Scientia Iranica
- 2002 – Present: Prentice Hall Publishing
- 2002 – Present: John Wiley & Sons Ltd.
- 2002 – Present: Thomson Learning (EMEA) Ltd.

University Service:
- 2006 – 2008: Computer Science Admission Committee, University of Houston–Clear Lake
- 2006 – 2008: Mentor for NSF undergraduate scholars, University of Houston–Clear Lake
- 2002 – 2008: Course Coordinator for Design of Database Systems and Web Application Development, University of Houston–Clear Lake
- 2002 – 2008: Faculty advisor of 20 undergraduate students in Computer Science and Computer Information Systems, University of Houston–Clear Lake
- 2003: Course Coordinator for Advanced Data Structures and Algorithms, University of Houston–Clear Lake

Recruitment:
- Programming Contest, San Jacinto Community College, Houston, Texas, 2008
- UHD Scholars Day, University of Houston–Clear Lake, November 2007
- Texas Work Source, Texas City, Texas, June 2007
- Deer Park High School, Deer Park, Texas, May 2007