

## Weiguo (Patrick) Fan, Ph.D.

樊卫国博士

### Associate Professor Mahlon Harrell Research Fellow

Department of Accounting and Information Systems  
Pamplin College of Business  
Virginia Polytechnic Institute and State University  
3007 Pamplin Hall, Blacksburg, VA 24061, USA  
Tel: 540-231-6588 Fax: 540-231-2511 E-mail: wfan@vt.edu  
URL: <http://www.acis.pamplin.vt.edu/fac/fan.htm>

---

#### EDUCATION

- Doctor of Philosophy** (Business Administration – Information Systems) 2002  
Ross School of Business  
University of Michigan, Ann Arbor, USA
- Master of Science** (Computer Science) 1998  
School of Computing  
National University of Singapore, Singapore
- Bachelor of Engineering (Hons.)** (Information Science and Engineering) 1995  
Xi'an Jiaotong University, P. R. China

#### WORK EXPERIENCE

- Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA
  - *Associate Professor* of Information Systems, Computer Science (Since 2006)
  - *Assistant Professor* of Information Systems, Computer Science (2002-2006)

#### Courses taught:

- ACIS 1504 : *Introduction to Business Information Systems*
- ACIS/BIT 4554 : *Telecommunications and Networks in Business*
- ACIS/BIT 4564 : *Object-Oriented Systems Development in Java*
- ACIS4024/5014 : *Information System Auditing and Control*
- ACIS5704 (PMBA) : *Managing Information Systems for Executive Decision Making*

- University of Michigan Business School, Ann Arbor, Michigan, USA
  - *Lecturer*, Computer and Information Systems (2001)
- National University of Singapore, Singapore

- **Instructor**, Information Systems and Computer Science (1997)

## REFEREED JOURNAL PUBLICATIONS

1. J. Yan, N. Liu, **W. Fan**, Q. Yang, Z. Chen, W. Wei, "Trace oriented feature analysis for large scale text data dimension reduction," forthcoming at *IEEE Transactions on Knowledge and Data Engineering*, 2009.
2. **W. Fan**, P. Pathak, M. Zhou, "Genetic-based approaches in ranking function discovery and optimization in information retrieval - A framework," forthcoming, *Decision Support Systems*, 34 pages, 2009.
3. L. C. Schaupp, F. Belanger, **W. Fan**, "Examining the success of websites beyond e-Commerce: An extension of the IS Success Model", *Journal of Computer Information Systems*, 49(4), 42-52, 2009.
4. X. Lin, J. Yan, **W. Fan**, N. Liu, S. Yan, Z. Chen, "An online blog reading system by topic clustering and personalized ranking," *ACM Transactions on Internet Technology*, 9(3), 1-26, 2009.
5. H. Lin, **W. Fan**, Z. Zhang, "A qualitative study of Web-based knowledge communities: Examining success factors," *International Journal of e-Collaboration*, 5(3), 39-57, 2009.
6. R. S. Torres, A. Falcao, M. A. Goncalves, J. P. Papa, B. Zhang, **W. Fan**, and E. A. Fox, "A genetic programming framework for content-based image retrieval", 42(2), 283-292, *Pattern Recognition*, 2009.
7. R. Shen, N.S. Vemuri, **W. Fan**, E.A. Fox, "Integration of complex archaeology digital libraries: An ETANA-DL experience," *Information Systems*, 33(7-8), 699-723, 2008.
8. D. Roussinov, **W. Fan**, J. Robles, "Beyond keywords: Automated question answering on the Web," *Communications of the ACM*, 51(9), 60-65, 2008.
9. E.A. Fox, C. Andrews, **W. Fan**, J. Jiao, A. Kassahun, S.-C. Lu, Y. Ma, C. North, N. Ramakrishnan, A. Scarpa, B.H. Friedman, S. D. Sheetz, D. Shoemaker, V. Srinivasan, S. Yang, L. Boutwell, "A digital library for recovery, research, and learning from April 16, 2007 at Virginia Tech," *Journal of Traumatology*, 14(1), 2008.
10. Z. Zhang, **W. Fan**, "Web server load balancing: A queueing analysis," *European Journal of Operational Research*, 186(2), 681-693, 2008.
11. Y. Yuan, **W. Fan**, C. Li, "Spline function smooth support vector machine for classification," *Journal of Industrial and Management Optimization*, 3(3), 529-542, 2007.

12. **W. Fan**, P. Pathak, L. Wallace, "Nonlinear ranking function representations in genetic programming-based ranking discovery for personalized Web search," *Decision Support Systems*, 42(3), 1338-1349, 2006.
13. **W. Fan**, M. D. Gordon, P. Pathak, "On linear mixture of expert approaches to information retrieval," *Decision Support Systems*, 42(2), 975-987, 2006.
14. **W. Fan**, M. D. Gordon, P. Pathak, "An integrated two-stage model for intelligent information routing," *Decision Support Systems*, 42(1), 362-374, 2006.
15. **W. Fan**, L. Wallace, S. Rich and Z. Zhang "Tapping the power of text mining," *Communications of the ACM*, 49(9), 76-82, 2006. Also **reprinted** in *IT management select* 13(1), 42-49, 2007. **IT management select publishes selected prime articles from international press on IT management.**
16. F. Belanger, **W. Fan**, C. Schaupp, J. Everhart, D. Poteet, A. Krishen, K. Nakamoto, "Website success metrics: addressing the duality of goals," *Communications of the ACM*, 49(12), 114-116, 2006.
17. M. D. Gordon, **W. Fan**, P. Pathak, "Adaptive Web search: evolving a program that finds information," *IEEE Intelligent Systems*, 21(5), 72-77, 2006.
18. E. A. Fox, F. D. Neves, X. Yu, R. Shen, S. Kim, **W. Fan**, "Exploring the computing literature with visualization and stepping stones & pathways," *Communications of the ACM*, 49(4), 52-58, 2006.
19. H. Wu, M. Gordon, K. Demaagd, **W. Fan**, "Mining Web navigations for intelligence," *Decision Support Systems*, 41(3), 574-591, 2006.
20. J. Yan, B. Zhang, N. Liu, S. Yan, Q. Cheng, **W. Fan**, Q. Yang, W. Xi, Z. Chen, "Effective and efficient dimensionality reduction for large scale and streaming data preprocessing," *IEEE Transactions on Knowledge and Data Engineering*, 18(3), 320-333, 2006.
21. **W. Fan**, M. D. Gordon and P. Pathak, "Genetic programming-based discovery of ranking functions for effective Web search," *Journal of Management Information Systems (JMIS)*, 21(4), 37-56, 2005.
22. **W. Fan**, M. D. Gordon, P. Pathak, "Effective profiling of consumer information retrieval needs: a unified framework and empirical comparison," *Decision Support Systems*, 40(2), 213-233, 2005.
23. D. Radev, **W. Fan**, H. Qi, H. Wu, A. Grewal, "Probabilistic question answering on the Web," *Journal of the American Society for Information Science and Technology*, 56(6), 571-583, 2005.
24. R. Barkhi, E. Rolland, J. Butler, **W. Fan**, "Decision support system induced guidance for model formulation and solution," *Decision Support Systems*, 40(2),

- 269-281, 2005.
25. **W. Fan**, M.D. Gordon, P. Pathak, "A generic ranking function discovery framework by genetic programming for information retrieval," *Information Processing and Management*, 40(4), 587-602, 2004.
  26. **W. Fan**, M. D. Gordon, P. Pathak, "Discovery of context-specific ranking functions for effective information retrieval by Genetic Programming," *IEEE Transactions on Knowledge and Data Engineering*, 16(4), 523-527, 2004.
  27. **W. Fan**, E. Fox and P. Pathak, H. Wu, "The effects of fitness functions on genetic programming-based ranking discovery for Web search," *Journal of the American Society for Information Science and Technology*, 55(7), pp. 628-636, 2004.
  28. M. D. Gordon, **W. Fan**, S. Rafaeli, H. Wu, N. Farag, "The architecture of CommKnowledge: combining link structure and user actions to support an online community," *International Journal of Electronic Business*, 1(1), 69-82, 2003.
  29. M. D. Gordon, R. Lindsay, **W. Fan**, "Literature-based discovery on the WWW," *ACM Transactions on Internet Technology (TOIT)*, 2(4), 262-275, 2002.
  30. D. R. Radev, K. Libner, **W. Fan**, "Getting answers to natural language questions on the Web," *Journal of the American Society for Information Science and Technology (JASIST)*, 53(5), 359-364, 2002.
  31. **W. Fan**, H. Lu, S. Madnick, D. Cheung, "DIRECT: a system for mining data value conversion rules from disparate sources," *Decision Support Systems*, 34(1), 19-39, 2001.
  32. **W. Fan**, H. Lu, S. Madnick, D. Cheung, "Discovering and reconciling data value conflicts for data integration," *Information Systems*, 26(8), 635-656, 2001.
  33. K. H. Phua, **W. Fan**, Y. Zeng, "Parallel algorithms for large scale nonlinear optimization," *Journal of International Transactions in Operational Research*, 5(1), 67-77, 1998.

#### **REFEREED BOOK CHAPTERS**

34. X. Yu, M. Tungare, **W. Fan**, Y. Yuan, M. Pérez-Quiñones, E. A. Fox, W. Cameron, L. Cassel, "Automatic syllabus classification using support vector machines," in *Handbook of Research on Text and Web Mining Technologies*, Idea Group Inc, 2008.
35. X. Yu, M. Tungare, **W. Fan**, M. Pérez-Quiñones, E. A. Fox, W. Cameron, L. Cassel, "Automatic genre-specific text classification," in *Encyclopedia of Data Warehousing and Mining (2<sup>nd</sup> edition)*, Idea Group Inc, 2008.

36. B. Zhang, **W. Fan**, M. Goncalves, Y. Chen and E. A. Fox, P. Calado and M. Cristo, "A genetic programming approach for combining structural and citation-based evidence for text classification in web digital libraries," 25 pages, in *Soft Computing for Information Retrieval*, Springer, 2005.
37. **W. Fan**, P. Pathak, "Discovering ranking functions for information retrieval," in *Encyclopedia of Data Warehousing and Mining*, Idea Group Inc., 13 pages, 2005.
38. H. Wu, **W. Fan**, "E-health communities: scalability challenges to information management," in *E-Health Paradigm Shift: Perspectives, Domains and Cases*, John Wiley, 2004.
39. H. Wu, D. R. Radev, **W. Fan**, "Towards answer-focused summarization using search engines," in *New Directions in Question Answering*, edited by Mark Maybury, AAAI Press, pp. 227-236, 2004.

#### **REFEREED CONFERENCE PROCEEDINGS**

40. G. Wang, J. Jiao, **W. Fan**, "Searching for authoritative documents in knowledge-based communities," in the *Proceedings of the International Conference on Information Systems 2009*, Phoenix, AZ, 2009.
41. L. Ji, J. Yan, N. Liu, W. Zhang, **W. Fan**, Z. Chen, "ExSearch: A novel vertical search engine for online barter business", in the *Proceedings of the ACM 18<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, Hong Kong, 2009.
42. N. Liu, J. Yan, **W. Fan**, Q. Yang, Z. Chen, "Identifying vertical search intention of query through social tagging propagation," in the Proceedings of 18th International Conference on World Wide Web, Madrid, Spain, 1209-1210, 2009.
43. J. Jiao, J. Yan, **W. Fan**, H. Zhao, "An expert user ranking algorithm in online communities," in the *Proceedings of the 2009 International Conference on New Trends in Information and Service Science (NISS 2009)*, Beijing, China, 2009.
44. C. D. Ferreira, R. da S. Torres, M. A. Goncalves, W. Fan, "*Image Retrieval with Relevance Feedback based on Genetic Programming*", in the Proceedings of the XXIII Simpósio Brasileiro de Banco de Dados (Brazilian Symposium on Database), Campinas, Brazil, 2008. **Winner of the best paper award.**
45. J. Yan, N. Liu, Q. Yang, **W. Fan**, Z. Chen, "TOFA: Trace oriented feature analysis in text categorization", in the *Proceedings of the 2008 IEEE International Conference on Data Mining*, Pisa, Italy, 2008.
46. D. Chen, J. Yan, G. Wang, Y. Xiong, **W. Fan**, Z. Chen, "TransRank: A novel algorithm for transfer of rank learning," in the *Proceedings of the ICDM Workshop on Domain-Driven Data Mining*, pp. 106-115, Pisa, Italy, 2008.

47. N. Liu, J. Yan, S. Yan, **W. Fan**, Z. Chen, "Web query prediction by unifying model," in the *Proceedings of the ICDM Workshop on Spatial and Spatial-Temporal Data Mining*, pp. 436-441, Pisa, Italy, 2008.
48. S. Ai, R. Du, and **W. Fan**, "Discovering and sharing experiential knowledge through active mining and friendly delivery: A design of EKDSS," in the *Proceedings of the 38<sup>th</sup> International Conference on Computers and Industrial Engineering*, pp. 570-575, Beijing, P.R. China, 2008.
49. X. Yu, M. Tungare, **W. Fan**, Y. Yuan, M. Pérez-Quiñones, E. A. Fox, W. Cameron, L. Cassel, "Using automatic metadata extraction to build a structured syllabus repository," in the *Proceedings of the 10th International Conference on Asian Digital Libraries*, 2007.
50. X. Yu, M. Tungare, **W. Fan**, M. Pérez-Quiñones, E. A. Fox., W. Cameron, G. Teng, L. Cassel, "Automatic syllabus classification," in the *Proceedings of the Seventh ACM/IEEE-CS Joint Conference on Digital Libraries ( JCDL)*, 2007.
51. M. Tungare, X. Yu, W. Cameron, G. Teng, M. Pérez-Quiñones, E. A. Fox., **W. Fan**, L. Cassel, "Towards a syllabus repository for computer science courses," in the *Proceedings of the Special Interest Group on Computer Science Education (SIGCSE)*, 2007.
52. H. Lin, **W. Fan**, L. Wallace, Z. Zhang, "An empirical study of knowledge community success," in the *Proceedings of the 40<sup>th</sup> Hawaii International Conference on System Sciences (HICSS)*, January 3-6, Big Islands, HI, 2007.
53. M. Tungare, X. Yu, M. Pérez-Quiñones, E. A. Fox, **W. Fan**, and L. Cassel, "Towards a standardized representation of syllabi to facilitate sharing and personalization of digital library content," In the *Proceedings of the 4th International Workshop on Applications of Semantic Web Technologies for E-Learning*, Dublin, Ireland, 2006.
54. Robert E. Crossler, France Bédanger, and **W. Fan**, "Determinants of Information Security End User Behavior," Workshop on Information Security and Assurance, Dec. 10, Milwaukee, WI, 2006.
55. Rao Shen, Naga Srinivas Vemuri, **W. Fan**, Edward A. Fox, "What Is a Successful Digital Library?" in the *Proceedings of the 10<sup>th</sup> European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Alicante, Spain, 2006.
56. Naga Srinivas Vemuri, Ricardo da Silva Torres, Rao Shen, Marcos André Gonçalves, **W. Fan**, Edward A. Fox, "A Content-Based Image Retrieval Service for Archaeology Collections," in the *Proceedings of the 10<sup>th</sup> European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Alicante, Spain, 2006.
57. Johnny L. Sam-Rajkumar, Rao Shen, Naga Srinivas Vemuri, **W. Fan**, Edward A. Fox: "EtanaCMV: A Visual Browsing Interface for ETANA-DL Based on Coordinated Multiple Views," in the *Proceedings of the 10<sup>th</sup> European Conference on*

- Research and Advanced Technology for Digital Libraries (ECDL)*, Alicante, Spain, 2006.
58. R. Shen, N. S. Vemuri, **W. Fan**, E. A. Fox, “Exploring Digital Libraries: Integrating Browsing, Searching and Visualization,” in the *Proceedings of 2006 ACM-IEEE Joint Conference on Digital Libraries (JCDL)*, Chapel Hill, NC, 2006.
  59. N. S. Vemuri, R. Shen, S. Tupe, **W. Fan**, E. A. Fox, “ETANA-ADD: An interactive tool for integrating archaeological DL collections,” in the *Proceedings of 2006 ACM-IEEE Joint Conference on Digital Libraries (JCDL)*, Chapel Hill, NC, 2006.
  60. H. Lin, **W. Fan**, Z. Zhang, “Unveiling critical success factors for a Web-based knowledge community,” in the *Proceedings of the 16<sup>th</sup> Biennial Conference for International Telecommunication Society*, June 11-16, Beijing, P.R. China, 2006.
  61. A. Lacerda, M. Cristo, M. Gonçalves, **W. Fan**, N. Ziviani, “Learning to Advertise,” in the *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Seattle, WA, 2006.
  62. L. C. Schaupp, **W. Fan**, F. Belanger, “Determining success for different website goals,” in the *Proceedings of the 39<sup>th</sup> Hawaii International Conference on System Sciences (HICSS)*, Kauai, Hawaii, 2006.
  63. D. Roussinov, **W. Fan**, “Learning ranking vs. modeling relevance,” in the *Proceedings of the 39<sup>th</sup> Hawaii International Conference on System Sciences (HICSS)*, Kauai, Hawaii, 2006. **Nominated for the best conference paper**
  64. N. Liu, J. Yan, F. Bai, B. Zhang, W. Xi, **W. Fan**, Z. Chen, L. Ji, C. Hu, W-Yi Ma, “A similarity reinforcement algorithm for heterogeneous Web pages,” in the *Proceedings of the APWeb 2005*, 121-132, 2005.
  65. B. Zhang, M. Goncalves, **W. Fan**, Y. Chen and E. A. Fox, P. Calado and M. Cristo, “Intelligent GP fusion from multiple sources for text classification,” in the *Proceedings of the ACM 14<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, Bremen, Germany, 2005.
  66. D. Roussinov, **W. Fan**, and F. A. D. Neves, “Semantic verification for fact seeking engines,” in the *Proceedings of the ACM 14<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, 2005.
  67. R. S. Torres, A. X. Falcão, B. Zhang, **W. Fan**, E. A. Fox, M. A. Gonçalves, P. Calado, “A new framework to combine descriptors for content-based image retrieval,” in the *Proceedings of the ACM 14<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, 2005
  68. D. Roussinov, **W. Fan**, “Learning global and local weighting for information retrieval,” in the *Proceedings of 2005 HLT/EMNLP Conference*, Vancouver, Canada, 2005.

69. D. Roussinov, **W. Fan**, "Mining context specific similarity relationships using the World Wide Web," in the *Proceedings of 2005 HLT/EMNLP Conference*, Vancouver, Canada, 2005.
70. H. Li, B. Zhang, Z. Chen, Y. Liu, J. Lei, W. Xi, **W. Fan**, W. Ma, "Improving search results using affinity graph," in the *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 15-19 Aug. 2005.
71. W. Xi, E. A. Fox, **W. Fan**, B. Zhang, Z. Chen, J. Yan, D. Zhuang, "SimFusion: Measuring similarity using unified relationship matrix," in the *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 2005.
72. J. Yan, B. Zhang, N. Liu, S. Yan, Q. Cheng, **W. Fan**, Z. Chen and W. Ma, "OCFS: Orthogonal centroid feature selection," in the *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 2005.
73. A. Raghavan, N. S. Vemuri, R. Shen, M. A. Goncalves, **W. Fan**, E. A. Fox, "Incremental, semi-automatic, mapping-based integration of heterogeneous collections into archaeological digital libraries: Megiddo case study," in the *Proceedings of the 9<sup>th</sup> European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Vienna, Austria, 2005.
74. R. Shen, M. A. Gonçaves, W. Fan, E. A. Fox, "Requirements gathering and modeling of domain-specific digital libraries with the 5S framework: An archaeological case study with ETANA," in the *Proceedings of the 9<sup>th</sup> European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Vienna, Austria, 2005.
75. N. Liu, F. Bai, J. Yan, B. Zhang, W. Xi, **W. Fan**, Z. Chen, W. Ma, L. Ji and C. Hu "A similarity reinforcement algorithm for heterogeneous Web pages," in the *Proceeding of Asia Pacific Web Conference*, 2005.
76. **W. Fan**, M. Luo, L. Wang, W. Xi, E.A. Fox, "Tuning before feedback: combining ranking function discovery and blind feedback for robust retrieval," in the *Proceedings of the 27<sup>th</sup> Annual International ACM SIGIR Conference*, U. K., 2004.
77. **W. Fan**, M. D. Gordon, P. Pathak, W. Xi and E.A. Fox, "Ranking function optimization for effective web search by Genetic Programming: an empirical study," in the *Proceedings of 37th Hawaii International Conference on System Sciences (HICSS)*, 2004. **Nominated for the best conference paper.**
78. J. Yan, B. Zhang, S. Yan, Z. Chen, **W. Fan**, Q. Yang, W. Ma and Q. Cheng, "IMMC: Incremental maximum margin criterion," in the *Proceedings of the Tenth ACM SIGKDD Conference*, 2004.

79. G. Xue, H. Zeng, Z. Chen, W. Ma, W. Xi, **W. Fan** and Y. Yu, "Optimizing web search using web click-through data," in the *Proceedings of the 13<sup>th</sup> ACM CIKM conference*, 2004.
80. B. Zhang, M. Goncalves, **W. Fan**, Y. Chen and E. A. Fox, P. Calado and M. Cristo, "Intelligent fusion of structural and citation-based evidence for text classification," in the *Proceedings of the 13<sup>th</sup> ACM CIKM conference*, 2004.
81. U. Ravindranathan, R. Shen, M. Goncalves, **W. Fan**, E. A. Fox, and J. W. Flanagan, "Prototyping digital libraries handling heterogeneous data sources - the ETANA-DL case study," in the *Proceedings of the European Conference on Digital Libraries (ECDL 2004)*, Bath, UK, 2004.
82. U. Ravindranathan, R. Shen, M. Goncalves, **W. Fan**, E. A. Fox, and J. W. Flanagan, "ETANA-DL: a digital library for integrated handling of heterogeneous archaeological data," in the *Proceedings of 2004 ACM-IEEE Joint Conference on Digital Libraries (JCDL)*. Tucson, AZ, 2004.
83. U. Ravindranathan, R. Shen, M. Goncalves, **W. Fan**, E. A. Fox, and J. W. Flanagan, "ETANA-DL: managing complex information applications – an archaeology digital library," Demo, in the *Proceedings of 2004 ACM-IEEE Joint Conference on Digital Libraries (JCDL)*. Tucson, AZ, 2004.
84. L. Wang, **W. Fan**, R. Yang, W. Xi, M. Luo, Y. Zhou, E.A. Fox, "Ranking function discovery by genetic programming for robust retrieval," in the *Proceedings of the Twelfth TREC Conference*, 2003.
85. R. Yang, **W. Fan**, L. Wang, W. Xi, M. Luo, Y. Zhou, E.A. Fox, "VT at TREC 2003: the web track report," in the *Proceedings of the Twelfth TREC Conference*, 2003.
86. H. Wu, M. Gordon, **W. Fan**, K. Demaagd, "Principal clusters analysis: a statistical analysis of web navigations," in the *Proceedings of the 23rd International Conference on Information Systems (ICIS)*, Barcelona, Spain, December, 2002.
87. H. Wu, D. Radev, H. Qi and **W. Fan**, "Towards better answer-focused summarization," in the *Proceedings of the 1<sup>st</sup> International Conference on Information Technology & Applications (ICITA)*, Bathurst, Australia, November, 2002.
88. D. Radev, **W. Fan**, H. Qi, H. Wu, "Probabilistic question answering from the Web," with, in the *Proceedings of the 11<sup>th</sup> WWW conference*, Hawaii, May, 2002.
89. D. Radev, H. Qi, **W. Fan**, H. Wu, "Evaluating web-based question answering systems," in the *Proceedings of Third International Conference on Language Resources and Evaluation*, Spain, 2002.
90. D. Radev, H. Qi, Z. Zheng, S. B. Goldensohn, Z. Zhang, **W. Fan**, J. Prager, "Mining the web for answers to natural language questions," in the *Proceedings of 2001*

- International Conference on Information and Knowledge Management (CIKM)*, Atlanta, Georgia, 2001.
91. D. Radev, **W. Fan**, Z. Zhang “WebInEssence: A personalized web-based multi-document summarization and recommendation system,” in *the Proceedings of NAACL Workshop on Automatic Summarization*, Pittsburgh, PA, 2001.
  92. **W. Fan**, M. D. Gordon and P. Pathak, “Personalization of search engine services for effective retrieval and knowledge management,” in *the Proceedings of the International Conference on Information Systems 2000 (ICIS)*, 20-34, Brisbane, Australia, 2000.
  93. D. Radev, **W. Fan**, “Automatic summarization of search engine hit lists,” in *the Proceedings of ACL'2000 Workshop on Recent Advances in Natural Language Processing and Information Retrieval*,” Hong Kong, P. R. China, 2000.
  94. P. Pathak, M. D. Gordon, **W. Fan**, “Effective information retrieval using genetic algorithms based matching function adaptation,” in *the Proceedings of the 33rd Hawaii International Conference on System Science (HICSS)*, Hawaii, USA, 2000.
  95. **W. Fan**, M. D. Gordon and P. Pathak, “Automatic generation of matching functions by genetic programming for effective information retrieval,” in *the Proceedings of the 1999 Americas Conference on Information Systems (AMCIS)*, Milwaukee, WI, USA, August 13-15, 1999.
  96. H. Lu, **W. Fan**, C. H. Goh, S. Madnick and D. Cheung “Discovering and reconciling semantic conflicts: a data mining perspective,” in *the Proceedings of the 7<sup>th</sup> IFIP 2.6 Working Conference on Database Semantics (DS-7)*, Leysin, Switzerland, 1997. Also available as [MIT Sloan Business School Working Paper #3957](#) and [MIT 21st Century Initiative Working Paper #18](#).

## CONFERENCE PRESENTATIONS

1. “An expert user ranking algorithm in online communities,” *the 2009 International Conference on New Trends in Information and Service Science (NISS 2009)*, Beijing, China, June 30<sup>th</sup> – July 2<sup>nd</sup>, 2009.
2. “From BPR to BPO: A new marketing strategy for outsourcing vendors”, First International Conference on Service Outsourcing (ICSC), Xi’an, P.R. China, Aug. 2-3. 2007.
3. “Spline function smoothed support vector machine for data classification”, 2007 Informs International Meeting, Puerto Rico, July 8-11, 2007.
4. “Unveiling critical success factors for a Web-based knowledge community,” *the 16<sup>th</sup> Biennial Conference for International Telecommunication Society*, June 11-16, Beijing, P.R. China, 2006.

5. “Ranking function optimization for effective web search by Genetic Programming: an empirical study,” the *37th Hawaii International Conference on System Sciences (HICSS)*, Big Island, Hawaii, 2004.
6. “Personalization of search engine services for effective retrieval and knowledge management,” International Conference on Information Systems 2000 (ICIS), Brisbane, Australia, Dec. 10 –14, 2000.
7. “Automatic generation of matching functions by genetic programming for effective information retrieval,” 1999 Americas Conference on Information Systems (AMCIS), Milwaukee, WI, USA, August 13-15, 1999.

### **INVITED TALKS / PRESENTATIONS**

1. “Applying FACET to Detect Fraudulent Financial Reporting”, KPMG, New York, Aug. 4<sup>th</sup>, 2009, co-presented with Greg. Jenkins.
2. “Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms”, XiDian University, Xi’an, P.R. China, July 7<sup>th</sup>, 2009.
3. “Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms”, School of Management, Xi’an Jiaotong University, P.R. China, July 6<sup>th</sup>, 2009.
4. “Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms”, Ross School of Business, University of Michigan, Ann Arbor, May 7<sup>th</sup>, 2009.
5. “Fraud Auditing Through Textual Analysis”, Barclays Global Investors, San Francisco, Oct. 17<sup>th</sup>, 2008.
6. “Finding authoritative documents in a knowledge-based community”, XiDian University, Xi’an, May 8<sup>th</sup>, 2008.
7. “Explaining users’ continued usage of social networking service website: An empirical study”, XiDian University, Xi’an, May 8<sup>th</sup>, 2008.
8. “Explaining users’ continued usage of social networking service website: An empirical study”, Xi’an Jiaotong University, Xi’an, May 6<sup>th</sup>, 2008.
9. “Introduction to ranking function discovery by genetic programming”, Tsinghua University, Beijing, April 26<sup>th</sup>, 2008.
10. “Ranking function discovery by genetic programming, a 10-year update”, Microsoft Research Asia, Beijing, April 24<sup>th</sup>, 2008.
11. “Tapping the power of text mining”, INFORMS VT Chapter, Feb. 21<sup>st</sup>, 2007.
12. “A conceptual model for Web-based knowledge community success”, School of Management, Xi’an Jiaotong University, P.R. China, June 2<sup>nd</sup>, 2006.

13. “A conceptual model for Web-based knowledge community success”, School of Management, Xi Dian University, June 8<sup>th</sup>, P.R. China, 2006.
14. “Optimizing Web search ranking functions by genetic programming,” School of Computer Science, Xi Dian University, June 8<sup>th</sup>, P.R. China, 2006.
15. “Introduction to data mining”, Southwest Virginia Chapter of the Institute of Internal Auditors, Jan. 10<sup>th</sup>, 2006.
16. “Optimizing Web search ranking functions by genetic programming,” Microsoft Research Asia, Beijing, May 16<sup>th</sup>, 2005
17. “Optimizing Web search ranking functions by genetic programming,” Intel China Research Center, Beijing, May 17<sup>th</sup>, 2005
18. “Ranking function optimization by genetic programming,” Guanghua School of Management, Peking University, Beijing, P.R. China, May 30<sup>th</sup>, 2005.

## HONORS, AWARDS AND GRANTS

- *Visiting Scholar*, Microsoft Research Asia, Beijing, June 27<sup>th</sup> - July 4<sup>th</sup>, 2009.
- *NSF NSDL Large Grant DUE-0840719, Co-PI, 2.5 million* (VT portion, \$500,000), 09/15/2008-08/31/2010.
- *Visiting Scholar*, Microsoft Research Asia, Beijing, April 19<sup>th</sup>- 30<sup>th</sup>, 2008.
- *Faculty Award for Research Excellence*, Pamplin College of Business, Virginia Tech, 2008.
- *Google Faculty Summit Fellow*, July 26-27, 2007.
- *Mahlon Harrell Research Fellow*, Department of Accounting and Information Systems, Virginia Tech, since June, 2007.
- *Virginia Tech Scholar of the Week*, April 16 – April 29, 2007.

“The Office of the Vice President for Research recognizes information systems and computer science professor **Patrick Fan**, who studies the efficiency and effectiveness of data mining strategies for search engine optimization and text summarization.

His research applies the design and development of novel information technologies --- information retrieval, data mining, text/web mining, personalization and knowledge management techniques --- to support better business information management and decision making for both individuals and organizations. In his short time at Virginia Tech, he has proven to be a prolific scholar, with more than 25 publications in top information systems journals and more than 40 refereed conference proceedings. His work on ranking function discovery for personalized web search results and text-summarization using genetic algorithms and other techniques has been widely cited. His National Science Foundation (NSF) funded work on web-based information

extraction and question answering is extending the usefulness of the internet to a broad range of populations.”

- **Warren L. Holtzman Outstanding Educator Award**, Pamplin College of Business, Virginia Tech, 2006.
- **Certificate of Teaching Excellence Award**, Pamplin College of Business, Virginia Tech, 2005-2006.
- **Department of Accounting and Information Systems Summer Research Grants (merit-based)**, 2003, 2004, 2005, 2007, 2008.
- **Virginia Tech International Travel Grants**, 2005, 2006, 2007, 2008, \$1000 each year.
- **PricewaterhouseCoopers LLP Large Research Grant, PI**, 08/01/2007-07/31/2009, “Automated Textual Analysis for Fraud Detection,” **\$200,000**, with Greg Jenkins.
- **NSF CPATH Grant CCF-0722259, Co-PI**, 08/01/2007-07/31/2009, “Collaborative Research: CPATH CB: Living In the KnowlEdge Society (LIKES),” **\$500,000**, with Edward Fox, Carlos Evia, Christopher Zobel, Steven Sheetz, Robert Beck (Villanova), Wingyan Chung (UTEP), Edward Carr (NCA&T).
- **NSF SGER Grant IIS-0736055, CO-PI**, 08/01/2007-07/31/2008, “SGER: DL-VT416: A Digital Library Testbed for Research Related to 4/16/2007 at Virginia Tech,” \$200,000, with Edward Fox, Christopher North, Naren Ramakrishnan, Donald Shoemaker.
- **NSF DUE Grant DUE-0532825, Co-PI**, 09/01/2005-08/31/2008, “Personalization of Content: Bridging the Gap between NSDL and Its Users through the Course Website,” **\$450,000**, with Lillian Cassel (Villanova), Edward Fox, Manuel Perez.
- **NSF ITR Medium Grant IIS-0325579, Co-PI**, 09/01/2003-12/31/2005, “Managing Complex Information Applications: An Archaeology Digital Library,” **\$385,000**, with Edward Fox, James Flanagan (CWRU).
- **New City Media, Co-PI**, 2004, “Defining and Evaluating Web Site Success,” **\$9324**, with F. Belanger, K. Nakamoto.

## **MEDIA MENTIONS**

- ACM TechNews on my CACM text mining article  
<http://technews.acm.org/archives.cfm?fo=2006-09-sep/sep-27-2006.html>
- Automatic textual analysis for fraud detection (see [VT news](#))
  - [Newswise Business News](#)
  - [Director of Finance News](#) (U.K.)

- [The Engineer Online News](#) (U.K.)
- [Scientific Frontline News](#)
- [Richmond Times](#) (indexed by LexisNexis, Factiva)
- [Collegiate Times](#)
- [American Society for Quality](#)
- [ID Theft Protect News](#)
- [Homeland Security Daily Wire News](#)
- [Chronicle](#)
- [CFO Magazine \(April 2008\)](#)
- [Monterey Wealth Report](#) (also featured in many weekly market commentaries, [online blogs](#), [online communities](#))
- [Research Milestone News](#)
- [SmartPros Accounting](#)
- [Individual.com](#)
- [StickyMinds.com](#)

## **RESEARCH INTERESTS**

- Business Intelligence
- Financial Statement Fraud Detection
- Applying Text Mining to Capital Market Research
- Customer Satisfaction and Loyalty in e-Services
- Competitions in the Digital Industry
- Information Integration
- Information Retrieval / Text Mining
- Information System Success
- Innovation Assessment
- Knowledge Innovation, Dynamics, and Diffusion in Knowledge Communities
- Knowledge Discovery / Data Mining
- Knowledge Management
- Quality metrics for Information Systems
- Social Computing
- Service Computing
- Customer Intelligence
- User Behaviors in Web 2.0

## TEACHING INTERESTS

- Business Data Communications
- Information System/Technology Auditing
- Object-Oriented System Development
- Management Information Systems
- Business Intelligence and Data Mining
- Knowledge Management
- Strategic Information Management
- E-Commerce/Business Technologies
- Database Management Systems

## PROFESSIONAL SERVICE

### ➤ Editorial Board Member

- ❖ *Journal of Database Management* (since June, 2004)
- ❖ *Journal of Electronic Commerce Research and Applications* (since July, 2008)

### ➤ Guest Editor, *Journal of Electronic Commerce Research*, 2010 May issue.

### ➤ National Science Foundation Panel Review Member

- ❖ REU 2004
- ❖ IIS 2005, 2007
- ❖ CDI, 2009

### ➤ External Review Member for Hong Kong Research Grant Council, 2008, 2009

### ➤ Ad Hoc Reviewer for

- ❖ *ACM Transactions on Information Systems (TOIS)*
- ❖ *ACM Transactions on Internet Technologies (TOIT)*
- ❖ *ACM Transactions on Computer-Human Interaction (TOCHI)*
- ❖ *Communications of the ACM (CACM)*
- ❖ *Decision Support Systems (DSS)*
- ❖ *Data and Knowledge Engineering (D&KE)*
- ❖ *IEEE Transactions on Evolutionary Computation (TEC)*
- ❖ *IEEE Transactions on Knowledge and Data Engineering (TKDE)*
- ❖ *IEEE Transactions on Systems, Man, Cybernetics (TSAC)*
- ❖ *Information Systems (IS)*
- ❖ *Information Sciences (IS)*
- ❖ *Information Systems Research (ISR)*

- ❖ *Information Systems Journal (ISJ)*
- ❖ *Information Processing and Management (IP&M)*
- ❖ *International Journal of Electronic Commerce (IJECE)*
- ❖ *Journal of the American Society for Information Science and Technology (JASIST)*
- ❖ *Journal of Digital Library (JDL)*
- ❖ *Journal of Database Management (JDM)*
- ❖ *Journal of Natural Language Processing (JNLP)*
- ❖ *Management Information System Quarterly (MISQ)*
- ❖ *Management Science (MS)*
- ❖ *International Conference on Information Systems (ICIS)*
- ❖ *Extending Database Technologies (EDBT)*
- **Services for International Conferences and Workshops**
  - ❖ *Program committee member of the 4th International Conference on Information Systems, Technology and Management, Bangkok, Thailand, March 10-12th, 2010.*
  - ❖ *Program committee member of the 2009 Genetic and Evolutionary Computation Conference, Montreal, Canada, 2009.*
  - ❖ *Associate Editor of the 2009 International Conference on Information Systems, Phoenix, AZ, 2009.*
  - ❖ *Track Chair of the InfoScale 2008 (The Third International ICST Conference on Scalable Information Systems), Campania, Italy, 2008.*
  - ❖ *Program Committee member of the 3rd International Symposium on Trustworthiness, Reliability and Services in Ubiquitous and Sensor Networks (TRUST-08), Shanghai, China, 2008.*
  - ❖ *Program Committee member of third IEEE Asia-Pacific Services Computing Conference (APSCC 2008), Yilan, Taiwan, 2008.*
  - ❖ *Program committee member of the 2008 Genetic and Evolutionary Computation Conference, Atlanta, Georgia, 2008.*
  - ❖ *Program committee member of the ACM 17<sup>th</sup> Conference on Information and Knowledge Management (CIKM), Napa Valley, California, 2008.*
  - ❖ *Program committee member of the ACM 16<sup>th</sup> Conference on Information and Knowledge Management (CIKM), Lisbon, Portugal, 2007.*
  - ❖ *Program committee member of the IEEE International Conference on Services Computing (SCC 2007), Salt Lake City, Utah, 2007.*
  - ❖ *Program committee member of the Americas Conference on Information Systems (AMCIS), 2007.*
  - ❖ *Advisory board member, and program committee member of the Pre-ICIS SIG-*

*DSS workshop, Milwaukee, Wisconsin, 2006.*

- ❖ *Program committee member of the 18th IEEE International Conference on Tools with Artificial Intelligence, Washington, DC, 2006.*
  - ❖ *Program committee member of the 2006 International Conference on Innovations in Information Technology, Dubai, 2006.*
  - ❖ *Program committee member of the International Conference on Asian Digital Libraries (ICADL), Kyoto, Japan, 2006.*
  - ❖ *Associate Editor and Program committee member of the 2006 International Conference on Information Systems, Milwaukee, Wisconsin, 2006.*
  - ❖ *Program committee member of the ACM 15<sup>th</sup> Conference on Information and Knowledge Management (CIKM), Washington, DC, 2006.*
  - ❖ *Program committee member of the 9<sup>th</sup> Pacific Rim International Conference on Artificial Intelligence, Guilin, China, 2006.*
  - ❖ *Program committee member of the 3rd Asia Information Retrieval Symposium (AIRS 2006), Singapore, 2006.*
  - ❖ *Program committee member of the Pacific Asia Conference on Information Systems (PACIS), 2006.*
  - ❖ *Program committee member of the IEEE International Conference on Services Computing (SCC 2006), Chicago, 2006.*
  - ❖ *Program committee member of the HLT-NAACL 2006 Conference, New York, 2006.*
  - ❖ *Program committee member of the ICCSA'2006 Workshop on Visual Computing and Multimedia, Scotland, 2006.*
  - ❖ *Program committee member of the 2005 International Conference on Information Systems, Las Vegas, 2005.*
  - ❖ *Program committee member of the 15th Workshop on Information Technology and Systems (WITS), Las Vegas, 2005.*
  - ❖ *Program committee member of the International Conference on Asian Digital Libraries (ICADL), Bangkok, Thailand, 2005.*
  - ❖ *Program committee member of the 14th Workshop on Information Technology and Systems (WITS), Washington D.C. 2004.*
- Member of the IEEE Technical Committee on Digital Libraries (since June, 2004).

#### **COLLEGE/DEPARTMENTAL SERVICES**

- MBA advisory committee, 2009-2010
- ACIS department research committee member (since July, 2002)
- ACIS department academic program committee (since July, 2002)

- Faculty advisor, information system option (since July, 2002)
- Faculty advisor board member for Information Systems Society (since July, 2002)
- Faculty advisor for Asian Chinese Association (2004-2005)
- Faculty advisor for Beta Alpha Psi (2005-2006)
- Faculty advisor for Information Systems Society (2006-now)
- CS department faculty search committee (2004)
- Faculty advisor for Upward Bound Program at VT (2007) - advising one high school student Ellie Trump from Christiansburg High School, VA.

### **STUDENTS ADVISED/BEING ADVISED**

- Ph.D. students
  - ❖ Lemuria Carter (ACIS) (Graduated in 2006, Now at Mississippi State Univ.)
  - ❖ Yuxin Chen (CS) (Graduated in 2007. Now at Microsoft)
  - ❖ (David) Jing Dai (CS)
  - ❖ Bobby George (CS)
  - ❖ Marcos Andre Goncalves (CS) (Graduated in 2005. Now at UFMG, Brazil)
  - ❖ Jian Jiao (CS)
  - ❖ Seonho Kim (CS)
  - ❖ Hui Lin (ACIS) (Graduated in 2008. Now at DePaul Univ.)
  - ❖ Bing Liu (CS) (Now at Google)
  - ❖ Pengbo Liu (CS) (Now at Ebay)
  - ❖ Xiaomo Liu (CS)
  - ❖ James Long (ACIS) (Now at Auburn Univ.)
  - ❖ Ming Luo (CS) (Now at Yahoo)
  - ❖ Yi Ma (CS)
  - ❖ Michele Meckfessel (ACIS)
  - ❖ James Penner (ACIS)

- ❖ Christian Schaupp (ACIS) (Graduated in 2005. Now at WVU)
- ❖ Rao Shen (CS) (Graduated in 2006. Now at Yahoo)
- ❖ Wensi Xi (CS) (Graduated in 2006. Now at Google)
- ❖ Xiaoyan Yu (CS) (Now at MicroStrategy)
- ❖ Baoping Zhang (CS) (Graduated in 2006. Now at Microsoft)
- Master Students
  - ❖ Yi Cao (CS) (Now at Seagate)
  - ❖ Li Wang (now at BGI)
  - ❖ Doug Gordon (CS)
  - ❖ Rohit Kelapure (CS) (Now at IBM)
  - ❖ V. Vikram Raj (CS) (Now at Microsoft)
  - ❖ Unni Ravindranathan (CS) (Now at Microsoft)
  - ❖ Johnny L.S. Rajkumar (CS) (Now at Microsoft)
  - ❖ Srinivas Vemuri (CS) (Now at Microsoft)
  - ❖ Ye Zhou (Now at Google)
  - ❖ Haibei Zhao (ACIS)
  - ❖ Xiaoyu Zhang (CS) (Now at IBM)
- Undergraduate Students
  - ❖ Stephanie Rich (CS) (Now at Vanguard)

Last revision: 08/21/2009