

Rm 430BB, McClelland Hall
1130 E. Helen Street
Tucson, AZ 85721

Phone: (520) 626-2961
Fax: (520) 621-2433
Email: zhuzhang@u.arizona.edu

Zhu Zhang

EDUCATION

- 2005 Ph.D., Information and Computer Science and Engineering
University of Michigan, Ann Arbor, MI
*Thesis: Natural Language Relations: Classification and
Application*
- 1999 M.S., Information Systems
Fudan University, P. R. China
- 1996 B.E., Information Systems
Tongji University, P. R. China

PRIMARY RESEARCH INTEREST

Business intelligence, information retrieval, natural language processing, text/data mining, machine learning.

PRIMARY TEACHING INTEREST

Databases, software engineering, systems analysis and design, information retrieval, natural language processing, data mining, text mining, artificial intelligence, programming languages.

PROFESSIONAL EXPERIENCE

- 2005- Assistant Professor
Department of Management Information Systems, Eller College
of Management, the University of Arizona
- 2003 Research intern
Google Inc.
- 2002 Research intern
Natural Language Processing group, Microsoft Research
SIFTer project
- 2002 Attended *Academic Careers Workshop*
Computing Research Association
- 2002/
2003 Graduate student instructor for SI654 (DB App. Design)
University of Michigan
- 2001 Research intern
Natural Language Processing group, Microsoft Research
AMALGAM project

2000	Research intern Conversational Machines group, IBM T. J. Watson Research Speech act classification project
2000- 2004	Graduate research assistant University of Michigan
1997- 1998	Research associate HWTC Corp., P. R. China Enterprise information systems planning
1997	Research associate 863-CIMS project, P. R. China Simulation system design and implementation

HONORS AND AWARDS

2003/ 2004	Rackham Travel Grant University of Michigan
2002	Academic Careers Workshop scholarship Computing Research Association
2002	Rackham fellowship for international travel University of Michigan
1997	HWTC Scholarship Fudan University
1996	Best undergrad thesis award Tongji University
1994	Provost's special scholarship for undergrad publication Tongji University
1991- 1996	University scholarship Tongji University
1990	Third Prize National contest on mathematics, P. R. China

GRANTS

PI. Flandrau Science Center grant: Automatic Evaluation of Web Page Comprehensibility for Free-Choice Learners. \$75,000. 2006-2007.

PI. Flandrau Science Center grant: Semantic-Structure-Aware Knowledge Retrieval. \$33,000. 2008-2009.

GRANT PROPOSALS

NSF CAREER proposal, to resubmit in 2009.

Two NSF proposals, pending.

REFEREED JOURNAL PUBLICATIONS

Lin, M. and Zhang, Z. Question-Driven Segmentation of Lecture Speech Text: Towards Intelligent E-Learning Systems. *Journal of the American Society for Information Science and Technology (JASIST)*, Volume 59, Issue 2 , Pages 186-200.

Li, J., Zhang, Z., Li, X., and Chen H. Kernel-based Learning for Biomedical Relation Extraction. *Journal of the American Society for Information Science and Technology (JASIST)*, Volume 59, Issue 5 , Pages 756-769.

Zhang, Z. Mining Relational Data from Text: from Strictly-Supervised to Weakly-Supervised Learning. *Information Systems*, Volume 33, Issue 3, Pages 300-314.

Zhang, Z. Weighing Stars: Aggregating Online Product Reviews for Intelligence E-commerce Applications. *IEEE Intelligent Systems*, Volume 23, Issue 5, Pages 42-49.

Li, J., Wang, H. Zhang, Z., and Zhao, L. A Policy-based Process Mining Framework: Mining Business Policy Texts for Discovering Process Models. To appear in *Journal of Information Systems and E-Business Management*.

Li, X., Chen, H., Zhang, Z., Li, J., and Nunamaker, J. Managing Knowledge in Light of its Evolutionary Process: an Empirical Study on Citation-Network-based Patent Classification. Conditionally accepted by *Journal of Management Information Systems (JMIS)*.

OTHER JOURNAL PUBLICATIONS (IN CHINESE)

Zhang, Z., and Zhang, F. Normalization and Integration of Construction Project Management Information Systems, *Management Information Systems*, 10/1998.

Zhang, Z. Information System Analysis and Design for Highway Construction Project Investment Control, *Journal of Highway and Transportation Research and Development*, No.2, Vol. 15.

Zhang, Z., and Zhang, F. Investment Control for Highway Construction Projects, *Journal of Highway and Transportation Research and Development*, No.1, Vol. 15.

JOURNAL PAPERS UNDER REVIEW OR READY FOR SUBMISSION

Li, X., Chen, H., Li, J., and Zhang, Z. Gene Function Prediction with Gene Interaction Networks: A Context Graph Kernel Approach. Under revision at *IEEE Transactions on Information Technology in Biomedicine*.

Abbasi, A., Zhang, Z., Zimbra, D., Chen, H., Nunamaker, J. Applying Statistical Learning Methods to Fake Website Detection. Under review at *MISQ*.

Suakkaphong, N., Zhang, Z., and Chen, H. Disease Named Entity Recognition using Semi-Supervised Learning and Conditional Random Fields. Under review at *Journal of Biomedical Informatics*.

Lin, M., Cao, J., and Zhang, Z. Segmenting Presentation Videos by Topic. To be submitted to *DSS*.

Zhang, Z. and Yan, P. Mining Websites for Instructional Value. To be submitted to *ACM JDIQ*.

WORK IN PROGRESS

Sarnikar, S., Zhang, Z., and Zhao, L. Predicting Query Performance for Improving Retrieval Effectiveness in Domain Specific Repositories. Targeting *MISQ*, to be submitted 04/09.

Aabbasi, A., Zhang, Z., and Chen, H. Mining Sentiments from Text using Feature Relation Networks. Targeting *IEEE TKDE*, to be submitted 04/09.

Zhang, Z. Competitive Intelligence Driven by Text Mining. Targeting *ISR*, to be submitted 06/09.

Li, J., Wang, H., Zhang, Z., and Zhao J. Policy-based Processing Mining. Targeting *ISR*, to be submitted 12/09.

Lu, H., Zhang, Z., and Chen H. Text-Sentiment-Driven Stock Market Prediction. Targeting *Management Science*, to be submitted 12/09.

Lu, H., Zhang, Z., and Chen H. Mining Corporate Risks from Text Repositories. Targeting *ACM TOIS*, to be submitted 12/09.

Zhang, Z., Jensen, M., Wright, K. Language Expectancy Theory and Commercial Product Reviews. Targeting *MISQ*, to be submitted 12/09.

REFEREED CONFERENCE PROCEEDINGS

Lu, H., Huang, W., Zhang, Z., and Chen T. Identifying Firm-Specific Risk Statements in News Articles. To appear in *PAISI 2009*.

Li, J., Wang, H., Zhang, Z., and Zhao J. Relation-Centric Task Identification for Policy-based Processing Mining. In *ICIS 2008*.

Hu, D., Zhang, Z., and Zhao J. Expert Recommendation in Social Networks: Development and Comparison of Three Approaches. In *WITS 2008*.

Li, J., Wang, H., Zhang, Z., and Zhao J. Process Component Identification from Business Policy Documents. In *CSWIM 2008* (best paper awardee).

Yan, P., Zhang, Z., and Garcia, R. Automatic Website Comprehensibility Evaluation. In *IEEE/WIC/ACM Web Intelligence 2007*: 191-197

Li, X., Zhang, Z., Chen, H., and Li, J. Graph Kernel-Based Learning for Gene Function Prediction from Gene Interaction Network. In *IEEE BIBM 2007*: 368-373

Li, X., Chen, H., Zhang, Z. and Li J. Automatic patent classification using citation network information: an experimental study in nanotechnology. In *ACM/IEEE JC DL 2007*: 419-427

Li, J., Wang, H., Zhang, Z., and Zhao, J.L. Mining Business Policy Texts for Discovering Process Models: A Framework and Some Initial Results. In *Proceedings of the Sixth Workshop on E-business*, Montreal, Quebec, Canada, December 9, 2007. (Nominated for Best Paper Award - 4 nominations out of 90 accepted papers).

Sarnikar, S., Zhang, Z., and Zhao, L. Predicting Query Performance in Domain-Specific Corpora. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS 2007)*, Big Island, Hawaii, 2007.

Ma, J., Zhang, Z., and Garcia, R. Automatically Determining Web Site Comprehensibility. In *WITS 2006: Sixteenth Annual Workshop on Information Technologies and Systems*, Milwaukee, Wisconsin, USA, 2006.

Zhang, Z. and Varadarajan, B. Utility Scoring of Product Reviews. In *ACM CIKM 2006: Fifteenth International Conference on Information and Knowledge Management*, Washington DC, 2006. (15% acceptance rate)

Zhang, Z. Mining Inter-Entity Semantic Relations Using Improved Transductive Learning. In *IJCNLP 2005: The Second International Joint Conference on Natural Language Processing*, Jeju Island, Korea, 2005.

Ye, Y. and Zhang, Z. Tense Tagging for Verbs in Cross-Lingual Context: A Case Study. In *IJCNLP 2005: The Second International Joint Conference on Natural Language Processing*, Jeju Island, Korea, 2005.

Zhang, Z. Weakly-Supervised Relation Classification for Information Extraction. In *ACM CIKM 2004: Thirteenth International Conference on Information and Knowledge Management*, Washington DC, 2004. (19% acceptance rate)

Zhang, Z. and Radev, D. Combining Labeled and Unlabeled Data for Learning Cross-document Structural Relationships. In *IJCNLP 2004: The First International Joint Conference on Natural Language Processing*, Sanya City, Hainan Island, P.R.China, 2004. Also to appear in *Lecture Notes in Artificial Intelligence*, Springer-Verlag.

Radev, R., Allison, T., Blair-Goldensohn, S., Blitzer, J., Çelebi, A., Dimitrov, S., Drabek, E., Hakim, A., Lam, W., Liu, D., Otterbacher, J., Qi, H., Saggion, H., Teufel, S., Topper, M., Winkel, A., and Zhang, Z. MEAD - a Platform for Multidocument Multilingual Text Summarization. In *LREC 2004: International Conference on Language Resource and Evaluation*, Lisbon, Portugal, May 2004.

Zhang, Z., Otterbacher, J., and Radev, D. Learning Cross-document Structural Relationships using Boosting. In *ACM CIKM 2003: Twelfth International Conference on Information and Knowledge Management*, New Orleans, LA, 2003. (15% acceptance rate)

Zhang, Z., Blair-Goldensohn, S., and Radev, D. Towards CST-enhanced summarization. In *AAAI 2002: 18th National Conference on Artificial Intelligence*, Edmonton, Canada, August 2002.

Gamon, M., Ringger, E., Zhang, Z., Moore, R., and Corston-Oliver, S. Extraposition: A Case Study in German Sentence Realization. In *COLING*

2002: *19th International Conference on Computational Linguistics*, Taipei, Taiwan, 2002.

Radev, D., Blair-Goldensohn, S., Zhang, Z., and Raghavan, R. Interactive, Domain-independent Identification and Summarization of Topically Related News Articles. In *ECDL 2001: 5th European Conference on Research and Advanced Technology for Digital Libraries*, Darmstadt, Germany, 2001.

Radev, D., Qi, H., Zheng, Z., Blair-Goldensohn, S., Zhang, Z., Fan, W., and Prager, J. Mining the Web for Answers to Natural Language Questions. In *ACM CIKM 2001: Tenth International Conference on Information and Knowledge Management*, Atlanta, GA, 2001.

Radev, D., Fan, W., and Zhang, Z. Webinessence: a Personalized Web-based Multi-document Summarization and Recommendation System. In *NAACL Workshop on Automatic Summarization*, Pittsburgh, PA, 2001.

Zhang, Z. Conception of Hyperobject System: a Consistent Approach to Evolutionary Information System. In *International Conference on Information Technologies & Information Services*, Shanghai, P.R.China, 1994.

REFEREED BOOK CHAPTERS

Radev, R., Qi, H., Zheng, Z., Blair-Goldensohn, S., Zhang, Z., Fan, W., and Prager, J. Query modulation for web-based question answering. In Strzalkowski, T. and Harabagiu, S. (editors), *Advances in Open-Domain Question Answering*, Kluwer Academic Publishers, 2003.

REFEREED POSTERS AND DEMONSTRATIONS

Zhang, Z. Natural Language Relations: Classification and Application. *ALISE Annual Conference*, Boston, MA, January, 2005.

Radev, R., Otterbacher, J., and Zhang, Z. CSTBank: a Corpus for the Study of Cross-document Structural Relationship (Poster). In *LREC 2004: International Conference on Language Resource and Evaluation*, Lisbon, Portugal, May 2004.

Radev, D., Blair-Goldensohn, S., Zhang, Z., and Raghavan, R. Newsinessence: a System for Domain-independent, Real-time News Clustering and Multi-document Summarization (Demo). In *HLT 2001: Human Language Technology Conference*, San Diego, CA, 2001.

NON-REFEREED CONFERENCE PROCEEDINGS

Radev, D., Blair-Goldensohn, S., and Zhang, Z. Experiments in Single and Multi-document Summarization using MEAD. In *First Document Understanding Conference*, New Orleans, LA, September 2001.

TECHNICAL REPORTS

Zhang, Z., Gamon, M., Corston-Oliver, S., and Ringger, E. Intra-sentence Punctuation Insertion in Natural Language Generation. MSR-TR-2002-58.

PATENTS

Pending patent (with colleagues at Microsoft Research): Sentence Realization Model for a Natural Language Generation System.

PROFESSIONAL AFFILIATIONS

Association for Information Systems (AIS)

Association for Computing Machinery (ACM)

Association for the Advancement of Artificial Intelligence (AAAI)

The Association for Computational Linguistics (ACL)

Association of Library and Information Science Education (ALISE)

INVITED TALKS

CST-Enhanced Text Summarization. Microsoft Research. July, 2002.

Demystifying Text Mining. Lecture given at *CSc630: Advanced Topics in Database Systems*, at the University of Arizona. November, 2005.

PROFESSIONAL SERVICES AND ACTIVITIES

Member	Departmental Undergrad Committee (2008-)
Treasurer	USWIM
EB Member	International Journal of Web Portals (IJWP)
Panelist	NSF INTEROP program (2007)
Reviewer	NSF Linguistics program (2009)
Reviewer	ACM TOIS, IPL, IEEE SMC, DSS, JMIS, Information Retrieval, JASIST, IEEE IS, DKE, ISR, JCMC
PC member	ACM SIGIR 2008
PC member	CSWIM 2007 2008
PC member	WITS 2006
PC member	InfoScale 2006
PC member	CIKM 2005
PC member	IJCNLP 2005
Reviewer	L&T 2003
Reviewer	RANLP 2003

ADVISING

Master thesis advisor: Vani Datla, Javier Espinosa, Matthew Shucker, Dan Whitlock, Adam Hancock

Doctoral thesis committee: Ahmed Abbasi, Daning Hu, Jiexun Li, Jian Ma, Robert Schumaker, Gang Wang, Runpu Sun, Siddharth Kaza, Gavin Zhang, Yan Dang

Ph.D. minor advisor: Huoping Chen (ECE), Seungho Lee (SIE)

LANGUAGES

Chinese:	native language
English:	proficient
German:	intermediate level
French and Japanese:	beginner's level