

Lei Tang

Phone: 480-252-2013 (C), 480-727-7808 (O)

Email: L.Tang@asu.edu

Homepage: <http://www.public.asu.edu/~ltang9>

Latest CV: <http://www.public.asu.edu/~ltang9/CV.pdf>

Computer Science and Engineering

Arizona State University

PO Box 878809

Tempe, AZ 85287-8809, USA

Research Interests

Social Computing: Social Network Analysis, Collective Behavior Modeling and Prediction, Network Evolution, Heterogeneous Networks, Social Media

Data Mining: Information and Knowledge Management, Text Mining, Graph Mining, Clustering, Classification, Relational Learning,

Education

Ph.D, Arizona State University Tempe, AZ, USA
Computer Science & Engineering, 08/2004 - 05/2010 (expected) GPA: 4.0/4.0

B.S., Fudan University Shanghai, China
Computer Science, 09/2000 - 07/2004 GPA: 3.7/4.0

Honors and Awards

- Doctoral Student Forum Fellowship, SIAM International Conference on Data Mining, 2009
- Student Travel Award, WWW'09, IJCAI'09, ICML'09, KDD'08, ICML'07, AAAI'05
- GPSA Research Grant on Hierarchical Content Classification, Arizona State University, 2006
- GPSA Travel Grant, Arizona State University, 2005, 2006, 2009
- Graduate College Travel Grant, Arizona State University, 2005, 2006, 2009
- Best Undergraduate Thesis, Fudan University, 2004
- Alcatel's Academic Excellence Scholarship (2 out of 180), Fudan University, 2001, 2003, 2004
- People's Scholarship 1st Prize (5 out of 180), Fudan University, 2002

Research Experience

Research Associate, Arizona State University Fall, 2007 - Present
Data Mining and Machine Learning Laboratory, Advisor: Professor Huan Liu

– Predictive Modeling of Collective Behavior in Social Media

An understanding of collective behavior presented in social media is critical to design and utilize this new form of media for targeted marketing, collective filtering, and public policy analysis. My current research, based on my proposed social-dimension model in my dissertation work and collaboration with sociologists from University of Hawaii,

focuses on constructing scalable models to accurately predict and understand online behavior. Results have been published in *KDD'09*, *SDM'09*, *CIKM'09* and are submitted to *DMKD*.

- Group Detection, Evolution and Profiling in Dynamic Heterogeneous Networks
Social media presents various kinds of networks involving heterogeneous entities and interactions, resulting in *multi-mode* or *multi-dimensional* networks. This kind of networks are rarely studied in traditional social network analysis or graph mining. Moreover, networks in social media are evolving all the time. This research addresses community detection, understanding group formation and evolution in dynamic heterogeneous networks. Results have been published in *TKDD*, *KDD'08*, *ICDM'09* and are submitted to *TKDE* and *Journal of Web Semantics*.
- Multi-Label Hierarchical Content Categorization
The explosion of online content has made the management of such content non-trivial. We-related tasks such as web page categorization, news filtering, query categorization and tag recommendation, often involve the construction of multi-label categorization systems on a large-scale. In this project, I explore different strategies to construct proper taxonomies to organize contents and exploiting the label relationship to boost the categorization performance. Results have been published in *KDD'06*, *KDD'08*, *WWW'09*, *IJCAI'09*, *ICML'09* and *TKDD*.
- Bias Analysis with Feature Selection
Feature selection, the task of selecting a subset of original features, is a core problem in many data mining and machine learning tasks. In this project, I investigate the biases associated with feature selection in terms of skewed class distribution and different cross-validation strategies. The details have been published in *ICDM'05* and *AAAI'07*.

Research Intern, Yahoo! Inc.

Summer, 2008

Data Mining and Research Group, Mentor: Dr. Vijay K. Narayanan

- Developed a multi-label hierarchical categorization system for query classification, which effectively reduces memory footprint of the model from original 7GB to 400MB; Designed a novel approach to automatically determine the labels for each query, with results published in *WWW'09*.

Research Intern, America Online Inc.

Summer, 2005

Division of Service Infrastructure, Mentor: Dr. Jianping Zhang

- Developed a hierarchical text mining system which collects documents and trains document classifiers automatically. Extended results are published in *KDD'06* and *TKDD*.

Teaching Experience

Tutorial Presenter, IEEE International Conference on Social Computing 2009

- Prepared the slides with concrete examples for tutorial on *Community Detection and Behavior Study for Social Computing*.
- Presented the major technical part of the tutorial to audience of around 60 practitioners.

Guest Lecturer, Arizona State University

2009

- Lectured for undergraduate and graduate course *Introduction to Artificial Intelligence*, on Lisp programming and projects.

Teaching Assistant, Arizona State University

Fall 2004 - Spring 2007

- Assisted undergraduate or graduate courses including
 - * Principles of Programming Languages
 - * Data Structures and Algorithms
 - * Introduction to Artificial Intelligence
 - * Data Mining
- Assisted the instructor to design projects and exam questions.
- Assisted students with recitation of course materials and evaluated them on individual or group projects, assignments, quiz and homework grading.

Instructor, Summer Camp

2001

- Taught English in a summer camp with around 20 students from a local middle school.

Other Leadership Experience

Co-advisor of two new PhD students, October 2008 - present

Coordinator of Data Mining and Machine Learning Lab at ASU, 2005 - present

Coordinator of Fudan Alumni Association at Arizona State University, 2005.

Publications and Talks

Tutorial

Huan Liu, **Lei Tang** and Nitin Agarwal. *Tutorial on Community Detection and Behavior Study for Social Computing*. Presented in The 1st IEEE International Conference on Social Computing (SocialCom'09), 2009.

Book

Lei Tang and Huan Liu. *Community Detection and Mining in Social Media*, Morgan & Claypool Publishers, to be published in 2010.

Refereed Book Chapters

Lei Tang and Huan Liu. *Understanding Group Structures and Properties in Social Media*. In Link Mining: Models, Algorithms and Applications, Editors: Christos Faloutsos, Jiawei Han and Philip S. Yu. In press.

Lei Tang and Huan Liu. *Graph Mining Applications to Social Network Analysis*. In Managing and Mining Graph Data, Editors: Charu Aggarwal and Haixun Wang. Springer, in press.

Payam Refaeilzadeh, **Lei Tang** and Huan Liu. *Cross Validation*. In Encyclopedia of Database Systems, Editors: M. Tamer Özsu and Ling Liu. Springer, 2009.

Lei Tang, Huan Liu and Jianping Zhang. *Bridging Taxonomic Semantics to Accurate Hierarchical Classification*. In Encyclopedia of Data Warehousing and Mining, 2nd Edition, Editor: John Wang. Idea Group, 2008.

Refereed Journal Articles

Shuiwang Ji, **Lei Tang**, Shipeng Yu and Jieping Ye. *A Shared-Subspace Learning Framework for Multi-Label Classification*. ACM Transactions on Knowledge Discovery from Data (TKDD), To appear.

Lei Tang, Huan Liu, Jianping Zhang, Nitin Agarwal and John Salerno. *Topic Taxonomy Adaptation for Group Profiling*. ACM Transactions on Knowledge Discovery from Data (TKDD), vol 1, issue 4, 1-28, 2008.

Journal Articles under Review

Lei Tang and Huan Liu. *Toward Collective Behavior Prediction via Social Dimension Extraction*. Submitted to IEEE Intelligent Systems.

Lei Tang, Xufei Wang and Huan Liu. *Community Detection in Multi-Dimensional Networks*. Submitted to Journal of Web Semantics (JWS).

Lei Tang and Huan Liu. *SocioDim: A Social-Dimension Approach to Classification in Social Media*. Submitted to Journal of Data Mining and Knowledge Discovery (DMKD).

Nitin Agarwal, Huan Liu, **Lei Tang** and Philip S. Yu. *Modeling Blogger Influence in a Community*. Submitted to the International Journal on Very Large Data Bases (VLDB Journal).

Lei Tang, Huan Liu and Jianping Zhang. *Identifying Evolving Groups in Dynamic Multi-Mode Networks*. Submitted to IEEE Transactions on Knowledge and Data Engineering (TKDE).

Refereed Conference Papers

Lei Tang, Xufei Wang and Huan Liu. *Uncovering Groups via Heterogeneous Interaction Analysis*. In Proceedings of IEEE International Conference on Data Mining (ICDM'09), 2009.

Lei Tang and Huan Liu. *Scalable Learning of Collective Behavior based on Sparse Social Dimensions*. In Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM'09), Pages 1107-1116, 2009.

Lei Tang and Huan Liu. *Relational Learning via Latent Social Dimensions*. In Proceedings of The 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'09), Pages 817-826, 2009.

Lei Tang. *Collective Behavior Prediction in Social Media*. In Doctoral Student Forum, SIAM International Conference on Data Mining (SDM'09), 2009.

Lei Tang, Jianhui Chen and Jieping Ye. *On Multiple Kernel Learning with Multiple Labels*. In Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'09), Pages 1255-1260, 2009.

Jianhui Chen, **Lei Tang**, Jun Liu and Jieping Ye. *A Convex Formulation for Learning Shared Structures from Multiple Tasks*. In Proceedings of The 26th International Conference on Machine Learning (ICML'09), Pages 137–144, 2009.

Lei Tang, Suju Rajan and Vijay K. Narayanan. *Large Scale Multi-Label Classification via MetaLabeler*. In Proceedings of The 18th International World Wide Web Conference (WWW'09), Pages 211–220, 2009.

Shuiwang Ji, **Lei Tang**, Shipeng Yu and Jieping Ye. *Extracting Shared Subspace for Multi-label Classification*. In Proceedings of The 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'08), Pages 381–389, 2008.

Lei Tang, Huan Liu, Jianping Zhang and Zohreh Nazeri. *Community Evolution in Dynamic Multi-mode Networks*. In Proceedings of The 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'08), Pages 677–685, 2008.

Nitin Agarwal, Huan Liu, **Lei Tang** and Philip Yu (alphabetical order). *Identifying the Influential Bloggers in a Community*. In Proceedings of The 1st ACM Conference on Web Search and Data Mining (WSDM'08), Pages 207–218, 2008.

Payam Refaeilzadeh, **Lei Tang** and Huan Liu. *On Comparison of Feature Selection Algorithms*. In Proceedings of AAAI Workshop on Evaluation Methods for Machine Learning II (AAAI'07), Pages 34–39, 2007.

Lei Tang, Jianping Zhang and Huan Liu. *Acclimatizing Taxonomic Semantics for Hierarchical Content Categorization*. In Proceedings of The 12th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'06), Pages 384–393, 2006.

Lei Tang and Huan Liu. *Bias Analysis in Text Classification for Highly Skewed Data*. In Proceedings of The 5th IEEE International Conference on Data Mining (ICDM'05), Pages 781–784, 2005.

Ronglu Li, Xiaopeng Tao, **Lei Tang** and Yunfa Hu. *Using Maximum Entropy Model for Chinese Text Categorization*. In Proceedings of the 6th Asia Pacific Web Conference (AP-Web'04), Pages 578–587, 2004.

Papers in Progress

Lei Tang, Xufei Wang and Huan Liu. *Understanding Emerging Social Structures: A Group-Profiling Approach*. Submitted to the 19th International World Wide Web Conference (WWW'10), under review.

Lei Tang, Geoffrey Barbier, Huan Liu and Jianping Zhang. *A Social Network Analysis Approach to Detecting Suspicious Financial Activities*. Submitted to International Conference on Social Computing, Behavioral Modeling and Prediction (SBP'10), under review.

Lei Tang, Ali Abbasi, Huan Liu and Lei Wang. *Multi-Resolution Extraction of Sparse Social Dimensions*. In preparation for SIGIR'10.

Reza Zafarani, **Lei Tang**, and Huan Liu. *Maximum-Margin Block Models*. In preparation for ICML'10.

Lei Tang and Huan Liu. *Network Feature Selection*. In preparation for KDD'10.

Lei Tang, Reza Zafarani and Huan Liu. *Social Signatures across Social Media Sites*. In preparation for KDD'10.

Research Grants and Proposals

Actively involved in proposals for NSF, ARO, AFOSR and ONR Grants with Dr. Huan Liu.
ASU GPSA Research Grant on Hierarchical Content Categorization, 2006 - 2007.

Professional Services

Program Committee Member

- International Conference on Social Computing, Behavioral Modeling & Prediction, 2010.

Organizing committee Member

- Publicity Chair, International Workshop on Social Computing, Behavioral Modeling, and Prediction, 2008, 2009.

Reviewer

- IEEE Transactions on Knowledge and Data Engineering, 2009.
- ACM Transactions on Knowledge Discovery from Data, 2008, 2009.
- ACM International Conference on Web Search and Data Mining, 2010.
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2006, 2007, 2009.
- SIAM International Conference on Data Mining, 2006, 2008, 2009.
- IEEE International Conference on Data Mining, 2005, 2006, 2007, 2009.
- ACM Conference on Information and Knowledge Management, 2006–2009.
- International Joint Conference on Artificial Intelligence, 2009.
- IEEE International Conference on Social Computing, 2009.
- European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2008.
- International Conference on Machine Learning, 2005, 2006, 2008.
- AAAI Conference on Artificial Intelligence, 2005, 2007.
- Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2005.

Volunteer

- International Joint Conference on Artificial Intelligence, 2009.
- International Conference on Knowledge Discovery and Data Management, 2008.
- International Conference on Machine Learning, 2007.

Professional Memberships

- Member of Association for Computing Machinery
- Member of IEEE Computer Society