

Xin (Bruce) Chen

Rush Building, 30 N. 33rd Street,
Philadelphia, PA 19104
<http://www.pages.drexel.edu/~xc35/>

Email: xc35@drexel.edu
Phone: +1(573)-489-6787(cell)
+1(215)895-6360(office)

OBJECTIVE: To obtain a research and development position in the field of machine learning, data mining, web mining, and information retrieval.

QUALIFICATIONS PROFILE

- Extensive knowledge and hands-on experience in the areas of text mining, web mining, machine learning, information retrieval, social network analysis and knowledge representation, with special expertise in:
 - Nonparametric Bayesian statistics (hierarchical Dirichlet process, Dirichlet process mixture)
 - Topical modeling (PLSA and LDA-based topical modeling methods)
 - Classification (Naive Bayes Classifier, Bayesian Networks, Decision Trees, SVM, etc.)
 - Clustering (spectral clustering, agglomerative clustering, tensor canonical correlation analysis, etc.)
 - Information retrieval and Information extraction
 - Knowledge organization and representation (Ontology, metadata, and social tags)
- Demonstrated ability to develop innovative data mining approaches, methods, and algorithms
- Experienced in analyzing large-scale imperfect real world datasets
- Published 9 research papers in peer-reviewed journals and conferences including SIGKDD'10, CIKM'11, CIKM'10, CIKM'09, BIBM'11, BIBM'10, PAKDD'09, and IEEE TCBB.

EDUCATION

Ph.D. in Information Science	College of Information Science and Technology, Drexel University	Jun.2012 expected
Master in Signal and Information Processing	University of Science and Technology of China	2007
Bachelor in Electronic Engineering	University of Science and Technology of China	2004

RESEARCH EXPERIENCE

PhD Dissertation Research: Generative Latent Space Model for Automatic Image Tagging		2008-present
Research Assistant	College of Information Science and Technology, Drexel University	2010-present
Project: "Integrating and Mining Bio-Data from Multi Sources in Biological Networks", funded by NSF		
Research Assistant	Department of Electrical and Computer Engineering, Drexel University	2009-2010
Project: "A Machine Learning Framework for Metagenomic Relationships", funded by NSF		
Research Assistant	College of Information Science and Technology, Drexel University	2007-2009
Project: "A Unified Architecture for Data Mining Large Biomedical Literature Databases", funded by NSF		
Research Assistant	College of Information Science and Technology, Drexel University	2010-2011
Project: "IPL2: Internet Public Library", funded by IMLS		

INTERN EXPERIENCE

Summer Internship	Merck Research Laboratory (West Point, PA)	Jun. 2010-Aug.2010
Project: automated lipid identification using tandem mass spectrometry data & integration of BioCyc pathway ontology with metabolite ontology.		
Summer Internship	Merck Research Laboratory (West Point, PA)	Jun. 2009-Aug.2009
Project: database and software development to facilitate metabolite identification in LC-MS and NMR metabolomics.		

TEACHING EXPERIENCE

Teaching Assistant	College of Information Science and Technology, Drexel University	2011 Fall
Course: INFO153-Applied Data Management		

TECHNICAL SKILLS

Programming Language: Java, C, C++, Matlab
Databases: MySQL, Oracle, MS Access
Web Application Development: XSLT, PHP, JSP/Servlet
Statistical Software: SPSS, R
Operating systems: Windows, Linux, Mac OS.

LANGUAGE

Fluent in English and Mandarin

AWARDS

- Graduate Research Assistantship, The iSchool at Drexel University, 2007-2011
- Student Travel Award for the 19th ACM International Conference on Information and Knowledge Management (CIKM10), Toronto, Canada, 2010
- Student Travel Award for the 18th ACM International Conference on Information and Knowledge Management (CIKM09), Hong Kong, China, 2009
- Outstanding Summer Undergraduate Project Award by University of Science and Technology of China, 2003.

PROFESSIONAL AFFILIATIONS

- Institute of Electrical and Electronics Engineers (IEEE)
- Association for Computing Machinery (ACM)

SELECTED PUBLICATIONS

- **Xin Chen**, TingTing He, Xiaohua Hu, Yuan An, Xindong Wu, Inferring Functional Groups from Microbial Gene Catalogue with Probabilistic Topic Models. The 2011 IEEE International Conference on Bioinformatics and Biomedicine (**BIBM'11**). [**19.4% acceptance rate**]
- **Xin Chen**, Yuan An, Xiaohua Hu, Zunyan Xiong, Tingting He, E.K. Park, Perspective Hierarchical Dirichlet Process for User-Tagged Image Modeling. The 20th ACM Conference on Information and Knowledge Management (**CIKM'11**). [**20% acceptance rate**]
- **Xin Chen**, Xiaohua Hu, Tze-Yee Lim, X. Shen, E.K. Park, and Gail Rosen. Exploiting the Functional and Taxonomic Structure of Genomic Data by Probabilistic Topic Modeling. IEEE/ACM Transactions on Computational Biology and Bioinformatics (**IEEE TCBB**), 2011, in printing.
- **Xin Chen**, Xiaohua Hu, Xiajiong Shen, and Gail Rosen. Probabilistic Topic Modeling for Genomic Data Interpretation. The 2010 IEEE International Conference on Bioinformatics and Biomedicine (**BIBM'10**),
- **Xin Chen**, Xiaohua Hu, Zhongna Zhou, Caimei Lu, Gail Rosen, Tingting He, E.K. Park, A Probabilistic Topic-Connection Model for Automatic Image Annotation. The 19th ACM Conference on Information and Knowledge Management (**CIKM'10**). [**13.4% acceptance rate**]
- Caimei Lu, Xiaohua Hu, **Xin Chen**, Jung-ran Park, et al. (2010). The Topic-Perspective Model for Social Tagging Systems. The 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD'10**). [**13.3 % acceptance rate**]
- **Xin Chen**, Caimei Lu, Yuan An, and Palakorn Achananuparp. Probabilistic Models for Topic Learning from Images and Captions in Online Biomedical Literatures. The 18th ACM Conference on Information and Knowledge Management (**CIKM'09**). [**15% acceptance rate**]
- Caimei Lu, **Xin Chen**, and E.K. Park. (2009). Exploit the Tripartite Network of Social Tagging for Web Clustering. The 18th ACM Conference on Information and Knowledge Management (**CIKM'09**).
- **Xin Chen**, Xiaohua Hu, Xiajiong Shen, Spatial Weighting for Bag-of-Visual-Words Representation and Its Application in Content-Based Image Retrieval. The 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (**PAKDD'09**). [**33.4% acceptance rate**]