YU-RU LIN

Arts, Media and Engineering Program
Arizona State University
699 South Mill Avenue Room 395
Tempe AZ, 85281, U.S.A.

Office: 480-727-0297
Fax: 480-965-0961
yu-ru.lin@asu.edu
http://www.public.asu.edu/~ylin56/

RESEARCH INTERESTS

Computational problems relating to dynamic social network analysis – community dynamics, social information summarization and representation.

EDUCATION

2005 to present — **Ph.D student, Computer Science**, Arizona State University, Tempe, Arizona
Concentration: Arts, Media and Engineering
Advisor: Prof. Hari Sundaram
GPA: currently 4.0/4.0
Expected date of graduation: May 2010

1999 to 2001 — **M.S. in Computer and Information Science**, National Chiao Tung University, Taiwan
Thesis: *The Synthesis of Trees in Chinese Landscape Painting*
Thesis advisor: Prof. Zen-Chung Shih
GPA: 3.8/4.0 (average score: 92.28/100)

1995 to 1999 — **B.S. in Computer and Information Science**, National Chiao Tung University, Taiwan
GPA: 3.6/4.0 (accumulated GPA in the last two years: 3.9/4.0, average score: 81.13/100)

AWARDS AND HONORS

IBM PhD Fellowship Award, 2009

WWW 2008 Student Travel Award

Emerging Leaders in Multimedia, 2008, nominated to present in the Emerging Leaders in Multimedia Fourth Annual Watson Workshop, IBM T. J. Watson Research Center, Hawthorne, NY

Student Excellence Award, 2007-2008, supported through the Kauffman Campuses Initiative. This award is reserved for students who exemplify the entrepreneurial and innovative nature of AME.

Best Paper Award, 2002. *Chinese Ink Rendering for Trees Using Outline Drawing and Texture Strokes*, in Proceedings of 2002 International Workshop on Advanced Image Technology.

EXPERIENCE

2005 to present — **Research Assistant**, Arts, Media and Engineering Program, Arizona State University, Tempe, Arizona
Advisor: Prof. Hari Sundaram

- Research on *community discovery and evolution* in a time-evolving heterogeneous social network. This work resulted in papers published in WWE 2006, ICME 2007, WI 2007, WWW 2008, CIKM 2008, ICASSP 2009, ICME 2009, ICWSM 2009, KDD 2009, a journal article to appear in ACM Trans. on Knowledge Discovery from Data.
- Research on *splog* (*spam blog*) *detection* based on blog temporal dynamics. This work resulted in papers published in ICME 2007, AIRWeb 2007, TREC 2006 conference proceedings, and a journal article published in ACM Transactions on the Web (February 2008).

2008 Summer **Intern**, IBM T.J Watson Research Center (Hawthorne, NY)

Supervisor: Dr. Ravi B. Konuru

- Research on community analysis in an enterprise. The challenge lies in combining people's interactions within various social media with different types of dynamic and static social relationships. We propose a novel hypergraph factorizaiton algorithm to solve the problem. First, a relational hypergraph is introduced to combine diverse social contexts. Second, we extract communities by factorizing the heterogeneous networks based on the relational hypergraph. Third, we deal with the time evolving social context by incremental updating the factorization with historic community model. Finally, extensive experiments on a collection of IBM Lotus Dogear dataset suggest that our technique is scalable and able to extract meaningful work communities. To evaluate the usefulness of our approach, we use community model to predict users' future interests. The results show that our prediction significantly outperforms baseline methods. This work resulted in a paper to be published in KDD2009.

2007 Summer **Intern**, NEC Labs America, Cupertino, CA

Supervisor: Dr. Belle Tseng

- Research on community analysis in dynamic networks. We propose a novel approach for analysis of communities and their evolution. Traditional methods for detecting dynamic communities often consider the membership consistency and community evolution separately. In our approach, the discovered communities both fit well with current observed networked data and evolve smoothly from historic data, and the evolutional relationship can be inferred from the model. Experiments on synthetic and real datasets demonstrate that the proposed method reveals interesting community dynamics which are not directly obtainable from traditional methods. This work resulted in a paper published in WWW 2008.

2006 Summer **Intern**, NEC Labs America, Cupertino, CA

Supervisor: Dr. Belle Tseng

- Research on splog (spam blog) detection based on blog temporal and link properties. The approach considers the unique temporal characteristics of splogs and introduces a novel time-sensitive evaluation framework as well as a detection method. The detection method combines traditional content features with regularity-based temporal and link features and has been tested on the annotated ground truth on real world dataset with 90% accuracy. This was the only research on splog detection that was selected in the TREC 2006 Conference Proceedings.

2001 to 2004 **Supervising Engineer**, Advance and Innovation Center, Ulead Systems, Inc., Taiwan

- Video Editing and Authoring: Researched and designed automatic video summarization and editing tool. Developed an artistic image filter and a 3D transition effect for video authoring.
- 3D and Animation: Designed 3D Particle Systems to simulate natural phenomena such as fire, smoke, clouds, snow, etc. Designed object-based NPR effects to render 3D models in special styles.
- Digital Imaging: Designed 2D particle systems, an edge preserving blur filter for reducing noise of image, a visual effect to simulate night scene and to render a photorealistic moon in the sky, and image-based NPR effects for rendering images in simulated painting styles.

1999 to 2001 **Teaching Assistant**, Department of Computer and Information Science, National Chiao Tung University, Taiwan

- Courses: Computer Graphics; Advanced Computer Graphics

1997 to 2001 **Research Assistant**, National Center for High-performance Computing, Taiwan
Advisor: Dr. Fang-Pang Lin

- National Numerical Wind Tunnel Project: Developed visualization for particle trajectory and tracking for plastic extrusion simulation. The visualization included VR and animation.
- Material Database for Electronic Packages: Designed and developed a material database for analysis of numerical data of electronic package, with a user-friendly interface representing the information by graphics and tables.

2000 **Intern**, Digimax Corporation, Taiwan

- Developed 3D visual effects and 3D graphics format transcoder.

1999 **Intern**, National Chiao Tung University Library, Taiwan

- Developed a query interface for Faculty Publication Database.

1998 Student Project: MPEG-4 Application

- Developed an MPEG-4 application that allows users to interact with the video content streamed from a remote MPEG-4 server.

PUBLICATIONS

Electronic copies of my publications can be downloaded at:
http://www.public.asu.edu/~ylin56/research.html

I. Social Network Analysis – community discovery in online social networks:

1. Yu-Ru Lin, Jimeng Sun, Paul Castro, Ravi Konuru, Hari Sundaram and Aisling Kelliher, *MetaFac: Commmunity Discovery via Relational Hypergraph Factorization*, to appear in Proceedings of the 15th ACM SIGKDD Conference On Knowledge Discovery and Data Mining (KDD 2009)
2. Yu-Ru Lin, Hari Sundaram and Aisling Kelliher, *JAM: Joint Action Matrix Factorization for Summarizing a Temporal Heterogeneous Social Network*, to appear in Proceedings of International AAAI Conference on Weblogs and Social Media (ICWSM 2009)
3. Yu-Ru Lin, Hari Sundaram, Munmun De Choudhury and Aisling Kelliher, *Temporal Patterns in Social Media Streams: Theme Discovery and Evolution Using Joint Analysis of Content and Context*, to appear in Proceedings of 2009 IEEE International Conference on Multimedia and Expo (ICME 2009)
4. Munmun De Choudhury, Hari Sundaram, Yu-Ru Lin, Ajita John and Doree Duncan Seligmann, *Connecting Content to Community in Social Media via Image Content, User Tags and User Communication*, to appear in Proceedings of 2009 IEEE International Conference on Multimedia and Expo (ICME 2009)
5. Yu-Ru Lin, Jimeng Sun, Paul Castro, Ravi Konuru, Hari Sundaram and Aisling Kelliher,

Extracting Community Structure through Relational Hypergraphs, to appear in Proceedings of the 18th International World Wide Web Conference (WWW 2009) (poster).

6. Yu-Ru Lin, Hari Sundaram and Aisling Kelliher, *Summarization of Large Scale Social Network Activity*, in Proceedings of 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009).
7. Yu-Ru Lin, Yun Chi, Shenghuo Zhu, Hari Sundaram and Belle Tseng, *Analyzing Communities and Their Evolutions in Dynamics Networks*, to appear in ACM Trans. on Knowledge Discovery from Data (TKDD), special issue on Social Computing, Behavioral Modeling, and Prediction.
8. Yu-Ru Lin, Hari Sundaram and Aisling Kelliher, *Summarization of Social Activity over Time: People; Actions and Concepts in Dynamic Networks*, in Proceedings of ACM 17th Conference on Information and Knowledge Management (CIKM 2008) (poster). Acceptance rate: 16%.
9. Yu-Ru Lin, Yun Chi, Shenghuo Zhu, Hari Sundaram and Belle Tseng, *FaceNet: A Framework for Analyzing Communities and Their Evolutions in Dynamics Networks*, in Proceedings of the 17th International World Wide Web Conference (WWW 2008). Acceptance rate: 11%.
10. Yu-Ru Lin, Hari Sundaram, Yun Chi, Jun Tatemura and Belle Tseng, *Blog Community Discovery and Evolution Based on Mutual Awareness Expansion*, in Proceedings of 2007 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2007). Acceptance rate: 17%.
11. Yu-Ru Lin and Hari Sundaram, *Blog Antenna: Summarization of Personal Blog Temporal Dynamics Based on Self-similarity Factorization*, in Proceedings of 2007 IEEE International Conference on Multimedia and Expo (ICME 2007)
12. Yu-Ru Lin, Hari Sundaram, Yun Chi, Jun Tatemura and Belle Tseng, *Discovery of Blog Communities based on Mutual Awareness*, in Proceedings of the 3rd Annual Workshop on the Weblogging Ecosystem (WWE 2006)

Overview: These papers study social networking phenomena in the online social media, from two complementary perspectives: (1) macroscopic view – the community structure on the social networks, and (2) microscopic view – the dynamics of the individuals. In papers [1,5], we propose MetaFac (MetaGraph Factorization), a framework that extracts community structures from various social contexts and interactions. Our work has three key contributions: (1) metagraph, a novel relational hypergraph representation for modeling multi-relational and multi-dimensional social data; (2) an efficient factorization method for community extraction on a given metagraph; (3) an on-line method to handle time-varying relations through incremental metagraph factorization. Extensive experiments on real-world social data collected from the IBM intranet social media and Digg social network suggest that our technique is scalable and is able to extract meaningful communities based on the social media contexts. We illustrate the usefulness of our framework through prediction tasks – to predict users' future interests of tags (in IBM dataset), diggs and comments (in Digg dataset). Our prediction significantly outperforms baseline methods (including aspect model, tensor analysis), indicating the utility of metagraphs for handling time-varying social relational contexts. In papers [2,6,8] we propose a novel method for summarizing collective activity from large scale online social networks which have heterogeneous social interactions and exhibit temporal dynamics. Our approach includes (1) a novel summarization methodology that imposes a syntactic structure of the social activity to construct a semantics-rich summary; (2) an efficient unified temporal multi-graph mining algorithm for extracting activity themes over time; (3) a novel evaluation framework to quantitatively evaluate the quality of group activity summary. We have conducted extensive experiments on synthetic and real-world datasets. Experiments on real-world Flickr datasets

demonstrate that our technique significantly outperforms baseline algorithms. Paper [12] focuses on extracting blog communities based on mutually observable individual blogger actions and interactions. Our approach involves (1) developing computational models for mutual awareness, and (2) using the mutual awareness feature with a ranking based community extraction algorithm to discover communities. We validate our approach on a large blog dataset with excellent qualitative results. The extracted communities also shown to be semantically cohesive with respect to their topics of interest. In paper [10] we extract the temporal dynamics of thematic communities based on mutual awareness. Awareness arises due to observable blog actions and the expansion of mutual awareness leads to community formation. We model the process of mutual awareness expansion using a random walk process and extract communities based on the model. The evolution of community is studied by comparing community structures extracted from networks over time. In papers [7,9] we propose a novel framework for analyzing communities and their evolutions through a unified process. Our framework involves using non-negative matrix factorization to discover communities that both fit well with current observed networked data and evolve smoothly over time. We propose an iterative algorithm that is guaranteed to converge to an optimal solution with provable low time complexity. Extensive experimental studies on both synthetic datasets and real-world datasets demonstrate that our method gives more reasonable communities and provides insights that are not directly obtainable from traditional methods. Paper [11] deals with problem of analyzing and summarizing the temporal dynamics within personal blogs. Blog temporal dynamics are difficult to capture using a few class descriptors. Our approach comprises (1) a representation of blog dynamics using self-similarity matrices, (2) theme extraction using non-negative self-similarity matrix factorization, and (3) a visualization representing blog theme evolution. Summaries based on large real-world blog datasets reveals interesting temporal characteristics for four blog types – personal blog, cooperative blog, power blog and spam blogs.

***II. Adversarial Information Retrieval on the Web – splog detection*:**

13. Yu-Ru Lin, Hari Sundaram, Yun Chi, Jun Tatemura and Belle Tseng, *Detecting Splogs via Temporal Dynamics using Self-similarity Analysis*, in ACM Transactions on the Web (TWEB) Volume 2, Issue 1 (February 2008)
14. Yu-Ru Lin, Hari Sundaram, Yun Chi, Jun Tatemura and Belle Tseng, *Splog Detection Using Self-similarity Analysis on Blog Temporal Dynamics*, in Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web (AIRWeb 2007)
15. Yu-Ru Lin, Hari Sundaram, Yun Chi, Jun Tatemura and Belle Tseng, *Splog Detection Using Content, Time and Link Structure*, in Proceedings of 2007 International Conference on Multimedia and Expo (ICME 2007)
16. Yu-Ru Lin, Wen-Yen Chen, Xiaolin Shi, Richard Sia, Xiaodan Song, Yun Chi, Koji Hino, Hari Sundaram, Jun Tatemuran and Belle Tseng. *The splog detection task and a solution based on temporal and link properties*, in The Fifteeth Text REtrieval Conference Proceedings (TREC 2006)

Overview: These papers deal with detecting spam blogs (splogs), a major problem in the blogosphere. Splogs are undesirable blogs meant to attract search engine traffic, used solely for promoting affiliate sites. Papers [13-15] address the problem of splog detection using temporal and structural regularity of content, post time and links. Experiments based on the annotated ground truth on real world dataset show excellent results on splog detection tasks with 90% accuracy. In paper [16] we propose a comprehensive framework to address the issues of splog detection, including: (1) identifying unique

characteristics of splog, (2) a novel time-sensitive detection evaluation that captures the unique characteristics of splog, (3) a splog detection solution that combines new temporal and link features with traditional content features and uses machine learning technique (SVM-based classifier), and (4) an annotation tool to generate ground truth on the TREC-Blog dataset. Experiments for offline (traditional IR evaluation) and our proposed online splog detection task based on the annotated ground truth set show excellent results on both tasks.

***III. Computer Graphics / Image Processing*:**

17. Der-Lor Way, Yu-Ru Lin, Zen-Chung Shih, *The Synthesis of Trees in Chinese Landscape Painting Using Silhouette and Texture Strokes*, in Proceedings of The 10-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2002), 499-506
18. Der-Lor Way, Yu-Ru Lin, Zen-Chung Shih, *Chinese Ink Rendering for Trees Using Outline Drawing and Texture Strokes*, in Proceedings of 2002 International Workshop on Advanced Image Technology (IWAIT 2002, **BEST PAPER AWARD**)
19. Master Thesis: *The Synthesis of Trees in Chinese Landscape Painting* (2001)

Overview: The research considers non-photorealistic rendering (NPR) technique that deals with problem of reducing the visual complexity while maintaining the aesthetic value of synthesized images. NPR stems from a key insight: artistic works are not necessarily photorealistic, whereas the observer can still comprehend the information delivered by the artists, based on the artistic context of the work. Theses papers present a framework for automatically drawing three-dimensional trees (one of the essential painting subjects in Chinese landscape painting) in Chinese ink painting style. We address the problem by leveraging computer rendering techniques with knowledge about Chinese painting. Our approach includes image-based outline rendering and texture generation according to the information captured from the three-dimensional objects. We introduce novel reference maps to analyze the information for creating the brush strokes, textures, and inking parameters of washing tone. Stylized textures are created by procedurally defining the texture patterns. The analytical painting framework can be extended to draw other subjects in Chinese ink painting style. We demonstrate the results of our method with excellent 2D/3D rendering results.

PATENTS

2004 Yu-Ru Lin and Chun-Yi Wang, *Slideshow composition method*, U.S. Patent filed June 2004, Appl. No. 20050275805

Yu-Ru Lin, *Method for edge detection and contour stroke generation*, U.S. Patent filed April 2004, Appl. No. 20050238234

Yu-Ru Lin, Shu-Fang Hsu, Chun-Yi Wang, *System and method for the automatic and semi-automatic media editing*, U.S. Patent filed Feb. 2004, Appl. No. 20050182503

2002 Yu-Ru Lin and Alpha Wu, *Method for Rendering Outlines of 3D Objects*, U.S. Patent filed Nov. 2002, Appl. No. 20040085314

SERVICES

Program Committee: ICWSM Data Challenge Workshop 2009
Reviewer: TMM, TOMCCAP, TOIS, DAPD

SKILLS Extensive experience in computer graphics and animation, and image processing. In depth understanding of object-oriented programming paradigms and windows based MFC programming.
Language/Script: C++/C, JAVA, Python, JavaScript, PHP, SQL, etc.
Package/Application: MATLAB, Eclipse, MFC, OpenGL, Processing, VTK, Tcl/TK, WIN32, C++ STL, etc.

ACTIVITIES AME Student Association, 2006-2007.

CITIZENSHIP Taiwan citizenship. Student with F-1 visa.

REFERENCE Available upon request.