Computational Aspects of Social Machines

Jie Tang
Tsinghua University
Social Networks

• >900 million users
• the 3rd largest “Country” in the world
• More visitors than Google

• >500 million users

• More than 5 billion images

• 2009, 2 billion tweets per quarter
• 2010, 4 billion tweets per quarter
• 2011, 25 billion tweets per quarter

• 2012, 300 million users, 300% yearly increase

• Pinterest, with a traffic higher than Twitter and Google
Social networks already become a bridge to connect our daily **physical** life and the **virtual** web space

*On2Off* [1]

[1] Online to Offline is trillion dollar business
What are the **fundamentally new** things in social networks?

**Web 1.0**

Hyperlinks between web pages
Examples:
Google search (information retrieval)
What are the fundamentally new things in social networks?

Collaborative Web

(1) personalized learning
(2) collaborative filtering
What are the fundamentally new things in social networks?

Social Web

Interactions

1. Influence
2. Collective intelligence

(1) interactions
(2) information diffusion
Iceberg Model for Social Network

- User behaviors
- Network structure
- Social Tie
- Influence
- Collective Intelligence
- Information Diffusion
Social Ties Analysis

Collaborate with
John Hopcroft, Jon Kleinberg (Cornell), Tiancheng Lou (Google)

KDD 2010, PKDD 2011 (Best Paper Runnerup), WSDM 2012, DMKD
Example: finding **boss** in email networks
(PKDD’11, Best Paper Runnerup)

Enterprise email network

How to infer

- CEO
- Manager
- Employee

User interactions may form *implicit groups*
Networks

- **Epinions**: a network of product reviewers: 131,828 nodes (users) and 841,372 edges
  - trust relationships between users
- **Slashdot**: 82,144 users and 59,202 edges
  - “friend” relationships between users
- **Mobile**: 107 mobile users and 5,436 edges
  - to infer friendships between users
- **Coauthor**: 815,946 authors and 2,792,833 coauthor relationships
  - to infer advisor-advisee relationships between coauthors
- **Enron**: 151 Enron employees and 3572 edges
  - to infer manager-subordinate relationships between users.
Social Influence Analysis

Collaborate with
Jimeng Sun (IBM TJ Watson), Jiawei Han and Chi Wang (UIUC)

KDD 2009, KDD 2011, DMKD
Who are the opinion leaders in a community?

Find $K$ nodes (users) in a social network that could maximize the spread of influence (Domingos, 01; Richardson, 02; Kempe, 03)
Opinion Leader

Who are the opinion leaders in a community?

Challenge: How to quantify the strength of social influence between users?

Find $K$ nodes (users) in a social network that could maximize the spread of influence (Domingos, 01; Richardson, 02; Kempe, 03)
Topic-level Social Influence Analysis

Input: coauthor network

Social influence analysis

Output: topic-based social influences

Several key challenges:
- How to differentiate the social influences from different topics?
- How to incorporate different information (e.g., topic distribution and network structure) into a unified model?
- How to estimate the model on real-large networks?

Emotion/Action Prediction:
Can we infer users’ emotions and social action?

Collaborate with
Jinghai Rao (Nokia), Jimeng Sun (IBM TJ Watson), Yuan Zhang (MIT)

KDD 2010, ICDM 2010, ACM TKDD, IEEE TAC
It's an emotional world we live in!

Emotion stimulates the mind 3000 times quicker than rational thought!!!
It's an emotional world we live in!

Six degree vs. Three degree [Nature; BMJ]
“Happy” System

Can we predict users’ emotion?
Observations

Location correlation (Red-happy)

Activity correlation

The Old Summer Palace

Dorm

class room

GYM

KO

Activity correlation
Observations (cont.)

(a) Social correlation

Social correlation

(a) Temporal correlation

Temporal correlation

(a) Implicit groups by emotions
MoodCast: Dynamic Continuous Factor Graph Model

Our solution
1. We directly define continuous feature function;
2. Use Metropolis-Hasting algorithm to learn the factor graph model.
Social Action Modeling and Prediction

Action Prediction:
Will John post a tweet on “Haiti Earthquake”?

Personal attributes:
1. Always watch news
2. Enjoy sports
3. ....
“Social Machine”

• **Deploy** a “machine” on Weibo.com, the largest “Twitter” in China;
• **Act** as a person by auto follow/retweet/reply;
• **Attracted** thousands of fans.
Summaries

• Social network brings a trillion dollar opportunity

• Computational models
  – Social tie analysis
  – Social influence analysis
  – Emotion and Action prediction
Representative Publications

- Lu Liu, Jie Tang, Jiawei Han, and Shiqiang Yang. Learning Influence from Heterogeneous Social Networks. *DMKD*, 2012, Vol 25, Issue 3, pages 511-544.
- Chenhao Tan, Jie Tang, Jimeng Sun, Quan Lin, and Fengjiao Wang. Social Action Tracking via Noise Tolerant Time-varying Factor Graphs. *KDD’10*.
- Chi Wang, Jiawei Han, Yuntao Jia, Jie Tang, Duo Zhang, Yintao Yu, Jingyi Guo. Mining Advisor-Advisee Relationships from Research Publication Networks. *KDD’10*.
- Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. Social Influence Analysis in Large-scale Networks. *KDD’09*.
- Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. *KDD’08*. 
Thanks!

HP: http://keg.cs.tsinghua.edu.cn/jietang/
System: http://arnetminer.org