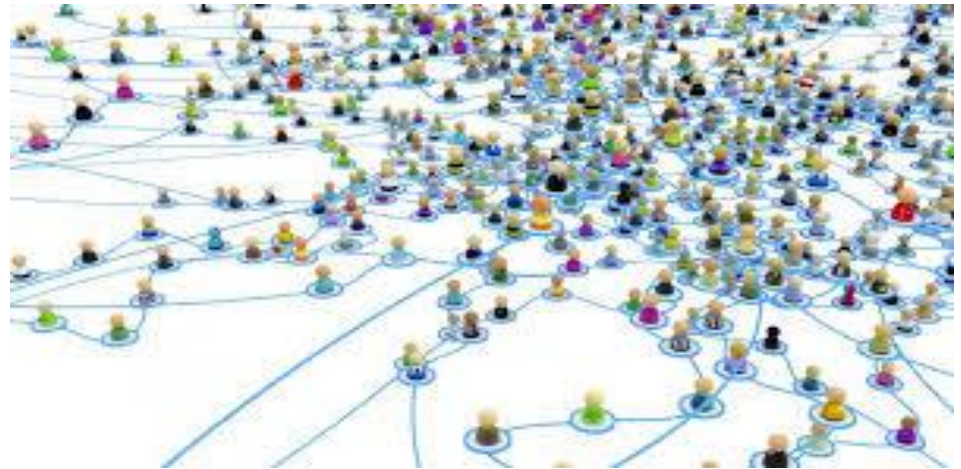


# Social Network Analysis

## **ELECTRONIC DATA SOURCES FOR NETWORK ANALYSIS**



Course Instructor: Dr.V.S.Felix Enigo

## DATA COLLECTION COMPLEXITY

- Social networks studied by observation requires close watch
- Standardized surveys minimize influence of observer but need active engagement of population to be studied
- Doubts whether responses are spontaneous and genuine
- To learn Dynamic Networks need multiple surveys
- Manual methods are labor intensive, 50% time spend on data collection
- Force to reuse the same dataset for research due to its complexity

## ALTERNATIVES

Reuse existing electronic records of social interactions not created for network analysis

*Examples:*

- Publication or project databases showing collaborations among authors or institutes
  - DB on corporate technology agreements to study networks of innovation
  - News paper archives – study of social-cognitive n/w's, terror organization structure
- Supports dynamic studies through historical analysis at significant price tag
- E.g. access to publication & patent DBs, media archives, legal and financial records

## ALTERNATIVES

- Internet is vast, diverse, dynamic and free for all resource
- Contributes to rapidly emerging field of *e-social science*
- Data collection automated as all rely on electronic networks, online information sources
- Allows to exploit dynamics of electronic data to perform longitudinal analysis

### ***Limits to e-social science:***

- social settings studied only by offline methods
- Technological limits to accuracy

## ELECTRONIC DISCUSSION NETWORKS

**First study on electronic data is from Information Dynamics Labs of Hewlett-Packard**

- Tyler et al studied communication among employees using corporate email archive
- Recreated network assigning tie between employee
- Assigned tie if exchanged certain threshold of emails
- Able to identify leaders, formal and informal communities
- Authors verified through employees of lab feedback by series of interviews

Adamic and Adar revisits question of local search problem:

- *how do people find short paths in social networks based on only local information about their immediate contacts?*

- Found that additional knowledge (E.g. physical location and position) helps to find path quickly

- More efficient than simple strategy of passing to most connected neighbor



- Study using email electronic data limited by privacy issue
- E.g. In case of HP, content of messages is ignored and data set cannot be shared with the community

## Alternatives

- Public forums, mailing lists has no privacy issues E.g. Analysis of USENET groups, WWW mailing lists
- Gloor recreated discussion networks from limited information (headers of messages)
- Gloor main contribution is dynamic visualization
- Able to quickly identify activation of entire group during key discussion not just few members
- Gloor also studied the structures that emerge over time

## BLOGS AND ONLINE COMMUNITIES

- Primary purpose of analyzing Blogs (Web logs) is trend analysis for the purposes of marketing
- Blogs allow to easily comment and react to the comments resulted communication web among bloggers
- Discussion networks in blogs formed through syndicate blogs, blog rolls and Blog walk series of meeting
- Discussion networks leads to formation of dynamic communities
- Communication and discussion network in blogs enabled for network analysis



# Social Network Analysis



 Sunday, August 20, 2006

## # Thinking and berries in Umea

I has not been blogging much last week, but this is only because I has been writing :) And, the best thing of it is where and how I has been writing.

I'm in Umea, Sweden, for [DIRN workshop](#), [presentation](#) and work/fun with [Stephanie](#). I'm happy I was able to come a few days earlier.

So far it has been almost perfect work-life balance environment. I worked on my own stuff (more productively than in my own office), discussed tons of things with Stephanie (mainly on weblog research, life and baking), enjoyed culture and nature, and all of that with picking and eating lots of berries.

Some time back [Aldo wrote](#) about thinking locations - places where you can get away from the pressures of the urgent to think your big deep thoughts - I was thinking of it while I enjoyed work and fun here in Umea.

The social component is very important, and perhaps one of the unique aspect of such a Deep Thought-network: thinkers need on the one hand to be able to concentrate, focus, and withdraw from the world. On the other hand, they very much need to be able to talk with kindred spirits, preferably people working on their own creative projects.

More on <http://thinkingcommunities.wikispaces.com>

Continued: [1 comments](#) | [TrackBacks](#) | [Links from other weblogs](#)  
More on: [life](#) [PhD](#)



Link to  
another blog

Link to  
another blog  
post

Quote

Links from  
other blogs

Comments

## **Blogs attracted research due to structured electronic data in RSS (Rich Site Summary)**

- RSS aids dynamic analysis due to the metadata ( timestamp ) of the content
- Kumar et al. and Gruhl et al. studied information diffusion in blogs based on this information
- Efimova and Anjewierden are the first to study blogs from a communication perspective
- Adar and Adamic offered visualization of communication in blogs

# Social Network Analysis



- Blogs exploited in 2004 US electoral campaign to build networks among individual activists and supporters
- Blogs used by marketers interested in understanding product choices of young demographics
- Separate blog research track is conducted in the yearly Sunbelt social networks conference

- Social networking services cater to socialization more directly than blogs
- Appealing features of social network services are social networking , messaging and photo sharing E.g. LiveJournal
- Paolillo et al studied about the interests and social networks of LiveJournal community
- Backstrom et al. studied influence of structural properties on community formation and growth, changes in membership of communities related to certain topic

# Social Network Analysis



- Most social network safeguard their and users data
- Alternative: centralized services like FOAF network
- FOAF profiles are stored on the web site of the users and linked together using hyperlinks
- Studies say FOAF network exhibits similar characteristics to other online social networks
- Drawback of FOAF is lack of tools for maintaining profiles and exploiting the network

## Summary

- Manual method of data collection for social network analysis is laborious
- Alternative is online electronic data sources
- Email electronic data limited by privacy issue
- Public forums, mailing lists has no privacy issues can be used
- Blogs and online communities are alternate sources of network analysis as they exhibit network of communication
- Social networking services are more appealing than blogs but has privacy issues
- FOAF centralized services mimic social network but they lack tools to analyze