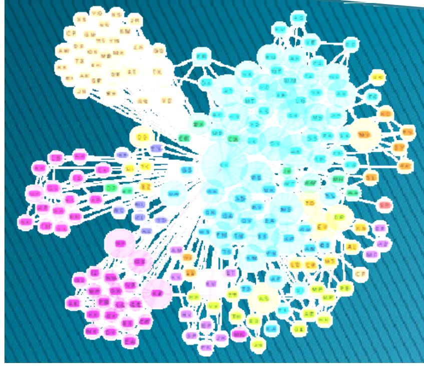
A study on social network metrics and their application in trust networks

Iraklis Varlamis
Harokopio University
Department of Informatics
and Telematics
Athens, Greece
varlamis@hua.gr

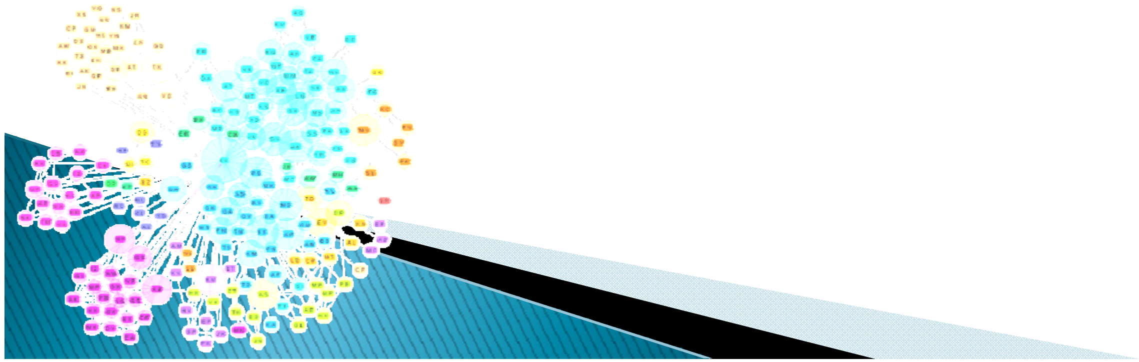
MagdaliniEirinaki
San Jose State University
Computer Engineering
Department
San Jose, CA, USA
magdalini.eirinaki@sjsu.edu

MalamatiLouta
Harokopio University
Department of Informatics
and Telematics
Athens, Greece
louta@hua.gr



Contents

- Main concepts
- Our work at a glance
- Social Network Analysis metrics
- Local and Collaborative rating
- Experiments and results
- Conclusions and future work



Social Networks

- High popularity: many participants in blogs, collaborative tagging and customer review sites etc.
- Increased research interest: sociologists, computer scientists, marketers etc.
- Unique characteristics:authorship,shared authorship, multitude of user-provided tags, inherent connectivity between users and posted items, frequent updates

Dynamics of the social network

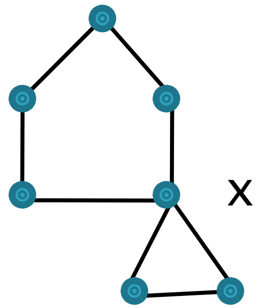
- The members of a network
  - link to members they trust
  - publish new information
  - read information published by other members
  - reference and comment on information provided by other members
- Influential members
  - Many people link to them
  - They publish first
  - They receive a lot of comments
  - They got referenced by many users
  - Valuable for viral marketing

At a glance

- Globally important members can be *influentials*
- Finding locally influential members, i.e. the members that influence the most a specific member or a group of members, is ideal for targeting this member or group
- Our model
  - creates a graph for the social network
  - employs social network analysis metrics for finding globally important members
  - combines global with local influence scores
  - provides personalized rankings of members for each community member

Social Network Analysis metrics

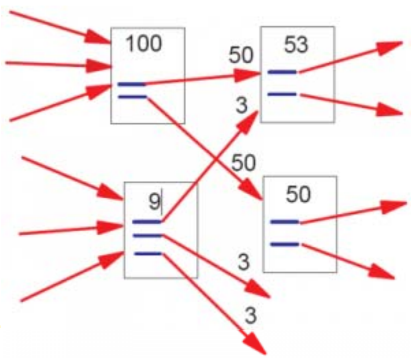
- Measures of importance or prominence
- Centrality: important actors typically occupy strategic locations in a network (undirected)
  - Degree centrality $$Gd(i) = \frac{d(i)}{n-1}$$
  - Closeness centrality $$Gc(i) = \frac{n-1}{\sum_{j=1}^{n} d(i,j)}$$
  - Betweennesscentrality $$Gb(i) = \sum_{j<k} \frac{sp_{jk}(i)}{sp_{jk}}$$

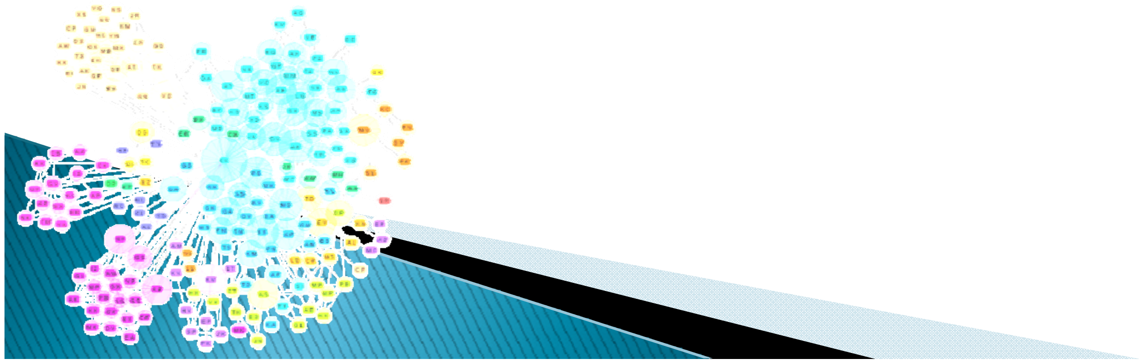


Gd(x)=4/6
Gc(x)=6/8
Gb(x)=15/21

Social Network Analysis metrics

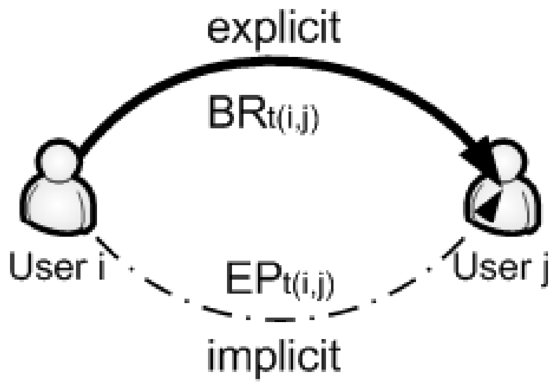
- Prestige: important actors point or are pointed by important users
  - Hub $$Ga(i) = \sum_{(j,i)\in E} Gh(j)$$
  - Authority $$Gh(i) = \sum_{(i,j)\in E} Ga(j)$$
  - PageRank $$Gp(i) = (1-d) + d \sum_{(j,i)\in E} \frac{Gp(j)}{O_j}$$





Local (direct) user score

- A user i *directly* trusts or is interested on another user j
- Direct trust or interest is based on explicit and implicit statements



$$LS_t(i,j)=w_{BR}\cdot BR_t(i,j)+w_{EP}\cdot EP_t(i,j)$$

Local accumulative user score

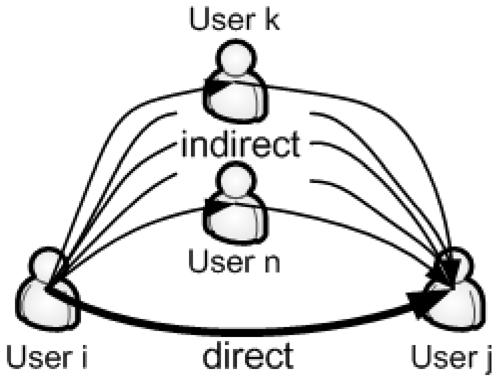
- Direct statements to user j are added by user i constantly, thus refreshing her/his interest to user j.

$$LAS_c(i,j) = \sum_{\substack{t=c-m+1 \\ t>0}}^{c} w_t \cdot LS_t(i,j)$$

- A rating system takes into account the m most frequent ratings starting from the current period c.

Collaborative local accumulative score

- Aggregates the direct accumulative scores *$LAS_c(i,j)$*, assigned by i to any user j, with the indirect accumulative scores *$LAS_c(k,j)$* assigned to j by all users k that i trusts



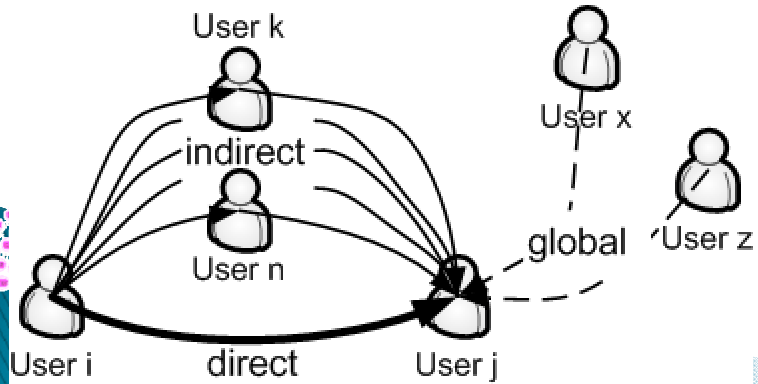
$$CLS_c(i,j) = w_i \cdot LAS_c(i,j) + \sum_{\substack{(i,j)\,\in\, E \\ (k,j)\,\in\, E \\ (i,k)\,\in\, E}} w_k \cdot LAS_c(k,j)$$

Influence

- We define global influence of a node to be the weighted sum of all the network analysis metrics

$$GI(i) = w_d \cdot Gd(i) + w_c \cdot Gc(i) + \\ w_b \cdot Gb(i) + w_h \cdot Gh(i) + \\ w_a \cdot Ga(i) + w_p \cdot Gp(i)$$

- We extend the collaborative local accumulative score to include the opinion of globally influential nodes



$$INF_c(i,j) = w_{local} \cdot LAS_c(i,j) + \\ w_{collab} \cdot \sum_{\substack{(i,j) \in E \\ (k,j) \in E \\ (i,k) \in E}} w_k \cdot LAS_c(k,j) + \\ w_{global} \cdot \sum_{(m,j) \in E} GI(m) \cdot LAS_c(m,j)$$

Experiments

- Compare the performance of local and global models of influence in providing recommendations to the users of social networks and combine them in a single model

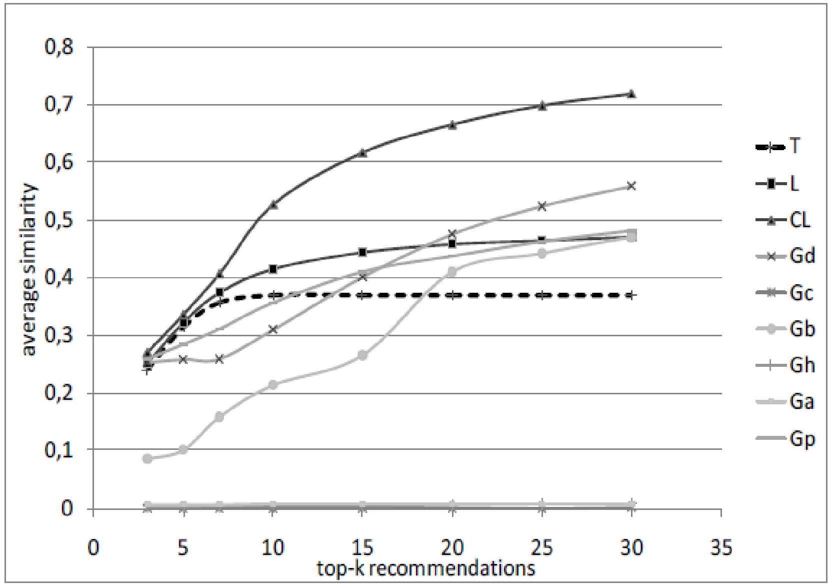
- Methodology
  - Provide for each user i a ranking for all users that i links to (directly or indirectly)
  - The ranking is based on the different rating mechanisms (combinations of local, collaborative local and global)

Dataset – Metrics

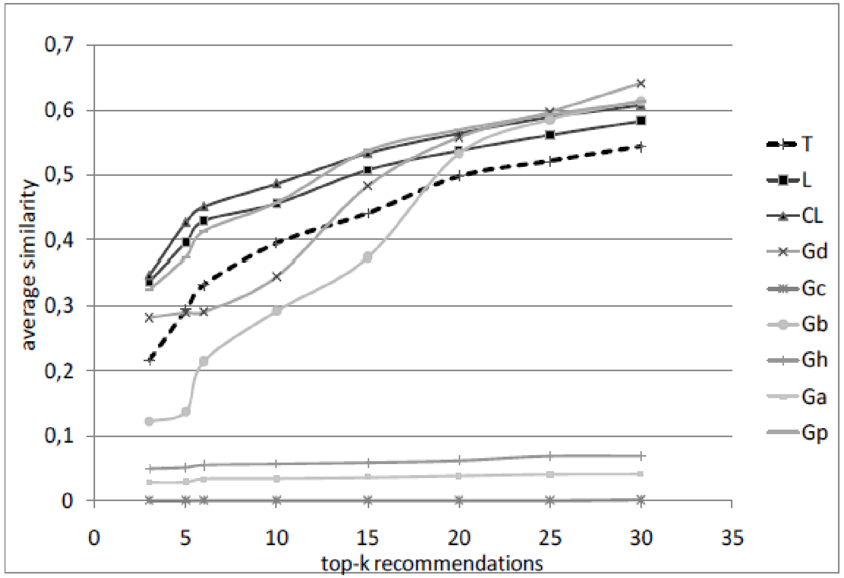
- We employed the *extended Epinions dataset*
  - 132,000 users who issued 841,372 statements
  - 717,667 positive implicit user-to-user trust ratings
  - 2 subsets of equal size (~5500 users)
    - Set A: users with few friends (5 to 10)
    - Set B: users with many friends (more than 30)
- Ratings
  - baseline: direct explicit links only (T)
  - local accumulative (L), collaborative local (CL)
  - degree centrality (Gd), closeness centrality (Gc), betweenness centrality (Gb), hub (Gh), authority (Ga), PageRank (Gp)
  - combinations: CL+individual Global , CL+combo Global

Results (local vs global)

set A



set B

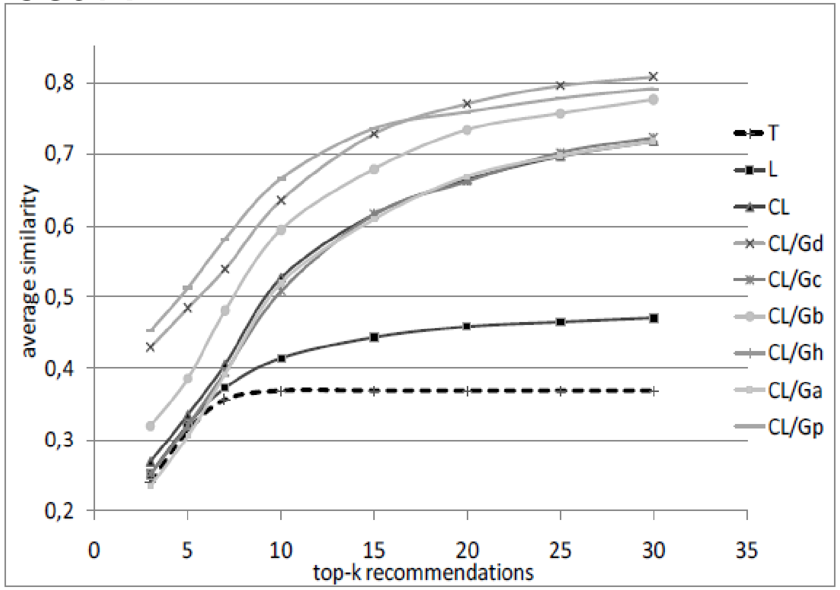


- CL significantly improves the performance of the baseline (T), especially for users with a small circle of trust (set A)
- It is useful to check for suggestions beyond the direct neighbors of a node, in the extended neighborhood of users

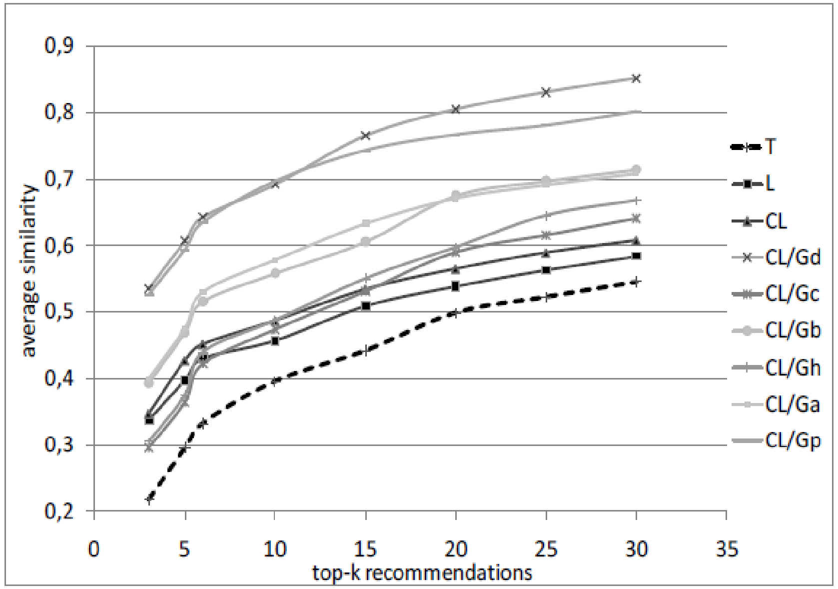
- For users with many neighbors (set B), certain global models (i.e. degree, betweenness and PageRank) perform better than local models examined

Results (CL plus global)

set A



set B

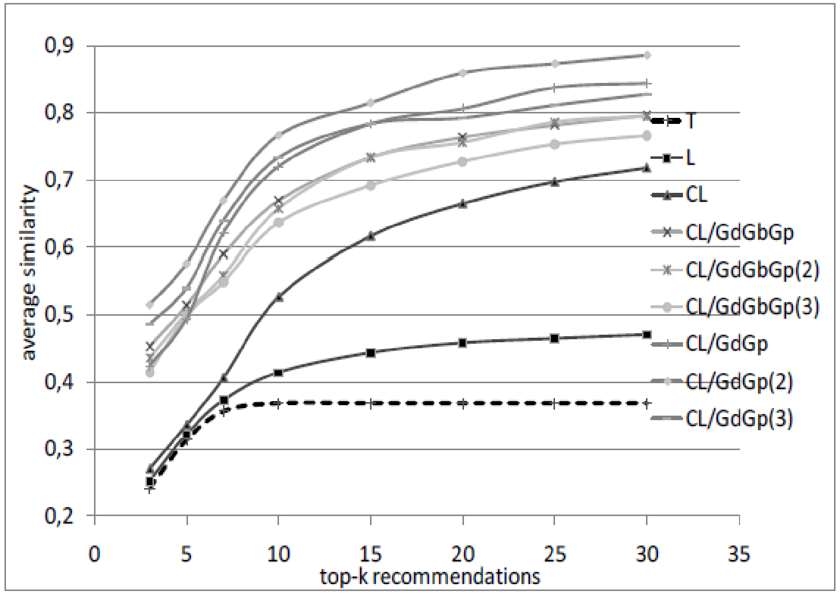


- highly ranked users (i.e. influential users) may provide additional recommendations which are useful to all authors
- The average improvement for all the values of k is 0.12, 0.13 and 0.06 for (CL/Gd), (CL/Gp) and (CL/Gb) respectively

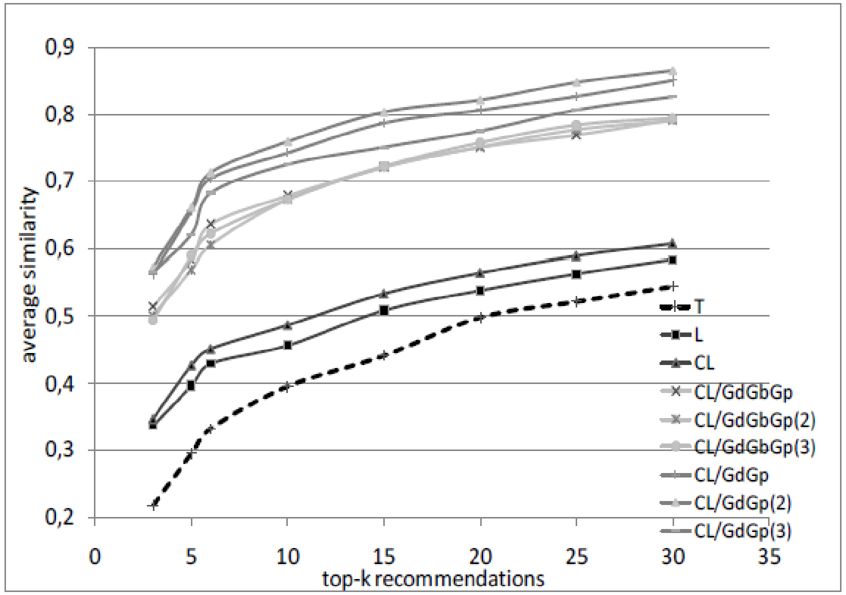
- the local methods demonstrate slightly improved results for set B in comparison to set A (average improvement is 0.037)
- the combined methods further increase this improvement (average improvement for PageRank and degree is 0.05)

Results (CL plus combo global)

set A



set B



Combinations of global metrics:

- CL/GdGbGp: wd = 0.2, wb = 0.2, wp = 0.6
- CL=GdGbGp(2): wd =1/3, wb =1/3, wp =1/3
- CL=GdGbGp(3): wd = 0.2, wb = 0.4, wp = 0.4
- GL=GdGp: wd = 0.5, wp = 0.5
- CL=GdGp(2): wd = 1/3, wp = 2/3
- CL=GdGp(3): wd = 2/3, wp = 1/3

- most of the combinations improve the results of the baseline and the collaborative local model with the combinations of PageRank and degree to outperform all other combinations

Conclusions

- We studied the contribution of various measures in identifying similar or influential actors in a social network in order to recommend them to a specific user
- Global measures are not very useful by themselves in providing recommendations to users
- When combined with the collaborative local measures have a positive impact in the final recommendation set
- Especially for users with few “friends”