Feedback Effects between Similarity and Social Influence in Online Communities

Authors: David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, Siddharth Suri

Presented by: Nedyalko Borisov
(CPS296.3 Spring 2009)

(slides borrowed from David Crandall KDD'08 presentation)
Homophily in social networks

“Birds of a feather flock together”

Caused by two related social forces [Friedkin98, Lazarsfeld54]

Social influence: People become similar to those they interact with
Selection: People seek out similar people to interact with

Both processes contribute to homophily, but

Social influence leads to community-wide homogeneity
Selection leads to fragmentation of the community
Importance to online communities

Together these forces shape how a community develops important for understanding health, trajectory of community

Applications in online marketing
viral marketing relies upon social influence affecting behavior recommender systems predict behavior based on similarity like selection vs. social influence, these are in tension

We study two questions in large online communities
How do selection & social influence interact to create social networks? Is similarity or interaction more predictive of future behavior?
Main questions

How do similarity and social ties compare as predictors of future behavior?

*Viral marketing* relies upon social influence affecting behavior
*Recommender systems* predict behavior based on similarity
like social influence and selection, these are in tension

Can we quantify and model the way in which selection and social influence interact to create social networks?
Important for understanding health, trajectory of a community
Interplay between influence and similarity

How does first interaction affect similarity?

Wikipedia: a large collaborative social network
users interact by posting to each others’ user-talk pages
user interests revealed by article edit patterns
rich, publicly-available, fine-grained log
Results
Results

Slower, long-term increase after first interaction (social influence)

Rapid increase in similarity before first interaction (selection)
Results

Effect is qualitatively stable across populations (admins/non-admins, heavy/light users, etc.) across different time scales, similarity metrics, languages, etc.
Holme and Newman Model

Each node has a single categorical attribute (one out of G possible opinions)
In each step, a node changes its opinion to
  - match a neighbor's opinion
  - re-wires one of its links to connect to someone of the same opinion
Not able to model Wikipedia users (too simple)
A model of user behavior

We model systems where people interact & do activities
each user $u$ has a history of $k$ most recent activities, $E_k(u)$

At each time step, user $u$ either,
interacts with another user, choosing someone who has engaged in a common activity or someone at random
performs an activity, choosing as follows:
Simulation results

We used the model to simulate user behavior in Wikipedia using maximum-likelihood estimates of the parameters. Simpler models (e.g. [Holme-Newman06]) do not produce this effect.
Predictive value

[Backstrom06] found that the more friends in a community, the higher a user’s probability of joining that community.

We compare similarity and social ties in predicting behavior in Wikipedia, social ties are more predictive.

In LiveJournal, interest similarity is more predictive.
Conclusions

We studied the interplay between selection and social influence in online communities

Modeled the feedback between activities and interactions; models individual behavior; explains aggregate phenomenon

Compared social ties, similarity as predictors of behavior; social ties better in Wikipedia, similarity better in LiveJournal
Discussion

Can this framework compare different social networks?
Can it suggest alternative/optimal designs?
Is this framework sufficient for social networks like Facebook?
  - Can all users fit the model?
  - Or just some of them?