Principles, models, and methods for the characterization and analysis of lurkers in online social networks

UNIVERSITÀ DELLA CALABRIA

Dipartimento di INGEGNERIA INFORMATICA, MODELLISTICA, ELETTRONICA E SISTEMISTICA

Roberto Interdonato, Andrea Tagarelli
DIMES Dept., Università della Calabria, Italy

The 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

ASONAM 2015


HOPE
08:00

Lurk(er): what meanings

lurk [lɜːk]
vb (intr)
1. to move stealthily or be concealed, esp for evil purposes
2. to be present in an unobtrusive way; go unnoticed

n
1. (Electronics & Computer Science / Telecommunications) to read messages posted on an electronic network without contributing messages oneself
2. *Austral and NZ slang* a scheme or stratagem for success

[probably frequentative of LOUR; compare Middle Dutch *loeren* to lie in wait]

lurker *n*

lurk (lɜrk)

v.i.
1. to lie or wait in concealment, as a person in ambush.
2. to go furtively; slink.
3. to exist unperceived or unsuspected.
4. *Chiefly Computers.* to observe an ongoing discussion without participating in it.

[1250–1300; frequentative of lower[2]; compare Norwegian *lurka* to sneak away]

lurk′er, *n.*

syn: lurk, skulk, sneak, prowl suggest avoiding observation, often because of a sinister purpose. To lurk is to lie in wait for someone or to move stealthily: *The thief lurked in the shadows.* skulk has a similar sense, but usu. suggests cowardice or fear: *The dog skulked about the house.* sneak emphasizes the attempt to avoid being seen or discovered; it suggests a sinister intent or the desire to avoid punishment: *The children sneaked out the back way.* prowl usu. implies seeking prey or loot; it suggests quiet and watchful roaming: *The cat prowled around in search of mice.*

Thesaurus **Legend:** Synonyms Related Words Antonyms

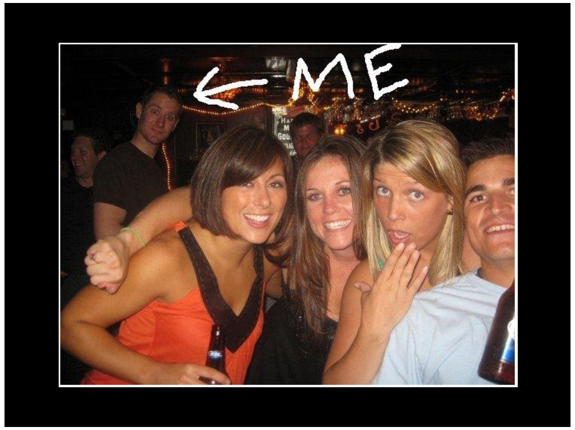
Verb 1. lurk - lie in wait, lie in ambush, behave in a sneaky and secretive manner
skulk
conceal, hide - prevent from being seen or discovered; "Muslim women hide their faces"; "hide the money"

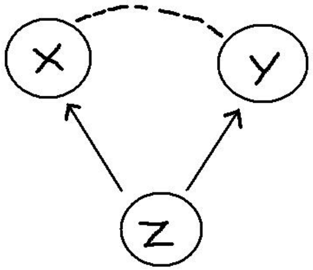
2. lurk - be about; "The high school students like to loiter in the Central Square"; "Who is this man that is hanging around the department?"
footle, hang around, lallygag, mess about, mill about, mill around, tarry, loiter, lollygag, lounge, linger, loaf
be - have the quality of being; (copula, used with an adjective or a predicate noun); "John is rich"; "This is not a good answer"
prowl, lurch - loiter about, with no apparent aim

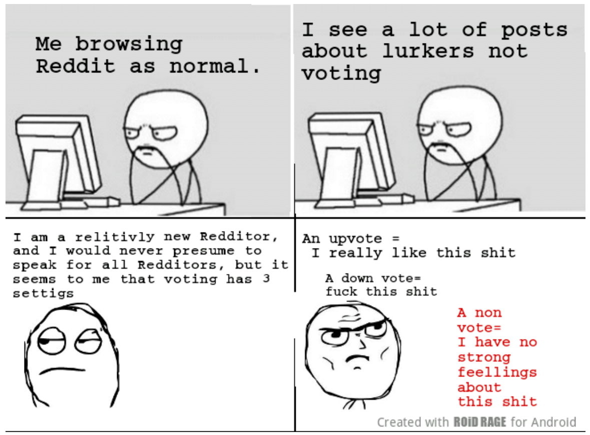
3. lurk - wait in hiding to attack
ambuscade, ambush, bushwhack, lie in wait, scupper, waylay
wait - stay in one place and anticipate or expect something; "I had to wait on line for an hour to get the tickets"

lurk
verb hide, sneak, crouch, prowl, snoop, lie in wait, slink, skulk, conceal yourself, move with stealth, go furtively *He thought he saw someone lurking above the chamber during the address.*

“Lurker”: let’s google it …







Lurk(er): what meanings

lurk [lɜːk]
vb (intr)
1. to move stealthily or be concealed, esp for evil purposes
2. to be present in an unobtrusive way; go unnoticed

n
1. (Electronics & Computer Science / Telecommunications) to read messages posted on an electronic network without contributing messages oneself
2. *Austral and NZ slang* a scheme or stratagem for success

[probably frequentative of LOUR; compare Middle Dutch *loeren* to lie in wait]

lurker *n*

lurk (lɜrk)

v.i.
1. to lie or wait in concealment, as a person in ambush.
2. to go furtively; slink.
3. to exist unperceived or unsuspected.
4. *Chiefly Computers.* to observe an ongoing discussion without participating in it.

[1250–1300; frequentative of lower[2]; compare Norwegian *lurka* to sneak away]

lurk'er, *n.*
syn: lurk, skulk, sneak, prowl suggest avoiding observation, often because of a sinister purpose. To lurk is to lie in wait for someone or to move stealthily: *The thief lurked in the shadows.* skulk has a similar sense, but usu. suggests cowardice or fear: *The dog skulked about the house.* sneak emphasizes the attempt to avoid being seen or discovered; it suggests a sinister intent or the desire to avoid punishment: *The children sneaked out the back way.* prowl usu. implies seeking prey or loot; it suggests quiet and watchful roaming: *The cat prowled around in search of mice.*

Thesaurus **Legend:** Synonyms Related Words Antonyms

Verb 1. lurk - lie in wait, lie in ambush, behave in a sneaky and secretive manner
skulk
conceal, hide - prevent from being seen or discovered; "Muslim women hide their faces"; "hide the money"

2. lurk - be about; "The high school students like to loiter in the Central Square"; "Who is this man that is hanging around the department?"
footle, hang around, lallygag, mess about, mill about, mill around, tarry, loiter, lollygag, lounge, linger, loaf
be - have the quality of being; (copula, used with an adjective or a predicate noun); "John is rich"; "This is not a good answer"
prowl, lurch - loiter about, with no apparent aim

3. lurk - wait in hiding to attack
ambuscade, ambush, bushwhack, lie in wait, scupper, waylay
wait - stay in one place and anticipate or expect something; "I had to wait on line for an hour to get the tickets"

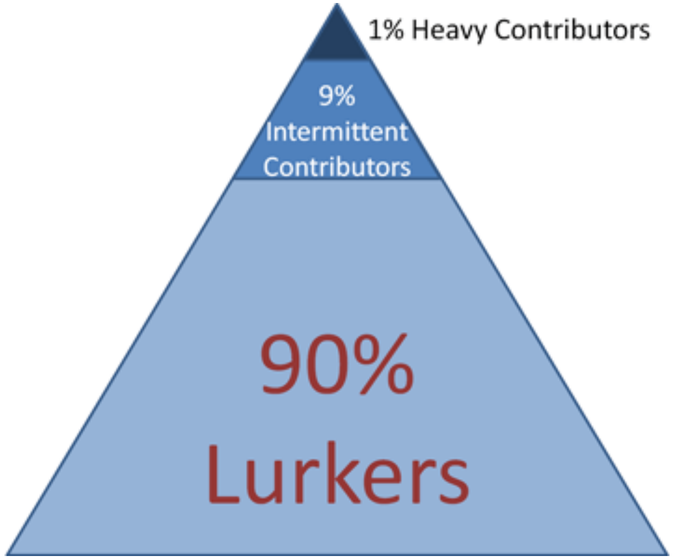
lurk
verb hide, sneak, crouch, prowl, snoop, lie in wait, slink, skulk, conceal yourself, move with stealth, go furtively *He thought he saw someone lurking above the chamber during the address.*

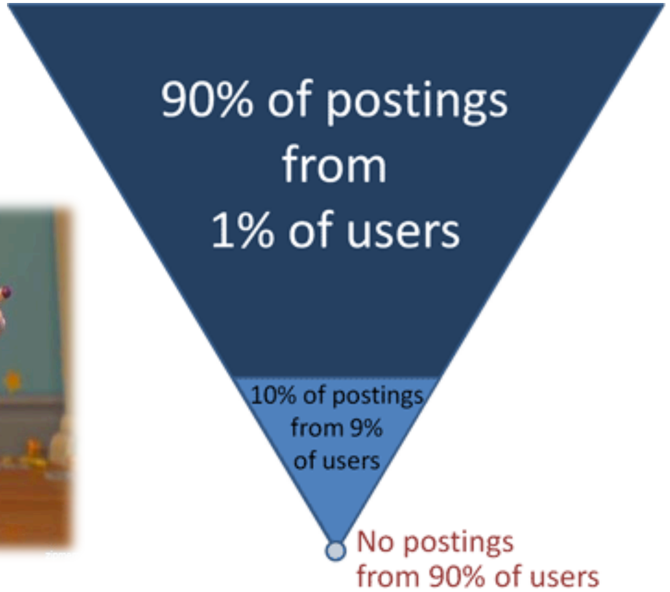
Outline

1. Lurking in online communities
   - The issue of controversial definitions
   - Lurking and online behavioral models
   - The opportunity of de-lurking
2. Modeling lurking behaviors
   - Topology-driven lurking definition
3. Lurker ranking methods
4. Experimental evaluation
   - Static scenarios
   - Dynamic scenarios
5. Applications to other domains
   - Vicariously learning
   - Lurking in social trust contexts
6. Delurking via Targeted Influence Maximization
   - The DEvOTION algorithm
7. Conclusion and future work

LURKING IN ONLINE COMMUNITIES

The 1:9:90 rule of participation inequality (1/3)





Arthur, C. (2006). What is the 1% rule? In: *The guardian*. UK: Guardian News and Media.

The 1:9:90 rule of participation inequality (2/3)

- [Nonnecke & Preece, 2000] Email-based discussion lists:
  - 77 online health support groups and 21 online technical support groups
  - 46% of the health support group members and 82% of the technical support group members are lurkers
- [Swartz, 2006] On Wikipedia: over 50% of all the edits are done by only 0.7% of the users
- [van Mierlo, 2014] On four DHSNs (AlcoholHelpCenter, DepressionCenter, PanicCenter, and StopSmokingCenter):
  - 63,990 users, 578,349 posts
  - Lurkers account for 1.3% (n=4668), Contributors for 24.0% (n=88,732), and Superusers for 74.7% (n=276,034) of content

Nonnecke, B., Preece, J. (2000). Lurker Demographics: Counting the Silent. In *Proc. SIGCHI Human Factors in Computing*.
Swartz, A. (2006). Raw thought: Who writes Wikipedia. Blog article at www.aaronsw.com/weblog/whowriteswikipedia.
van Mierlo, T. (2014). The 1% rule in four digital health social networks: An observational study. *Medical Internet Research*, 16(2).

The 1:9:90 rule of participation inequality (3/3)

- Online learning courses:
  - No relation between interactivity (i.e., posting) and learning (i.e., earned grade)
  - Extend the notion of interactivity to include the lurking activity
    - Each of the 128 students reads at least one contribution
    - 62% of the class are lurkers—only reading posts, not contributing anything
  - No correlation between the no. of readers and the no. of writers
  - Every participant, active or lurking, reads more postings than they write $\frac{R(t)}{p} - \frac{W(t)}{p} \geq 0$
  - Active participation in an online discussion list, based on passive lurking, is expressed by reading, reflecting on the contribution of all the other members

Ebner, M., Holzinger, A. (2005). Lurking: An underestimated human-computer phenomenon. *IEEE Multimedia*, 12(4), 70–75.

Perception of lurking (1/2)

- Lurkers as "free-riders" [Kollock & Smith,1996; Morris & Ogan, 1996; Wellman & Gulia,1999; Rheingold, 2000]
- Sustainability of an online community
  - Fresh content and timely interactions
  - Lurkers contribute little value [van Mierlo, 2014]
- Lurkers may impair the virality of the community [Nielsen, 2011]

Kollock, P., Smith, M. (1996). Managing the virtual commons. *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*, 109–128.
Morris, M., Ogan, C. (1996). The internet as mass medium. *Journal of Communication*, 46(1), 39–50.
Wellman, B., Gulia, M. (1999). Net surfers don't ride alone: Virtual communities as communities. *Networks in the Global Village*, 331–366.
Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier*. MIT Press.
Nielsen, J. (2011). Participation inequality: Encouraging more users to contribute, http://www.useit.com/alertbox/participation_inequality.html.

Perception of lurking (2/2)

- Most lurkers are NOT free-riders (e.g., [Nonnecke, Preece, & Andrews, 2004; Nonnecke, Andrews, & Preece, 2006])
- Lurking can be regarded as passive participation that permits inclusion [Ferree, 2002]
- Lurking is normal and an active, participative and valuable form of online behavior [Edelmann, 2013]
- Lurkers perceive themselves as community members [Nonnecke et al., 2006]
- Lurking as a form of cognitive apprenticeship: “legitimate peripheral participation” [Lave & Wenger, 1999]

Nonnecke, B., Preece, J., Andrews, D. (2004). What lurkers and posters think of each other. In *Proc. the 37th annual Hawaii Int. Conf. on System Sciences*.
Nonnecke, B., Andrews, D., Preece, J. (2006). Non-public and public online community participation: Needs, attitudes and behavior. *Electronic Commerce Research*, 6(1), 7–20.
Ferree, M. M., Gamson, W. A., Gerhards, J., Rucht, D. (2002). *Shaping abortion discourse: Democracy and the public sphere in Germany and the United States.* New York, Cambridge University Press.
Edelmann, N. (2013). Reviewing the definitions of “Lurkers” and some implications for online research. *Cyberpsychology, Behavior, and Social Networking*, 16(9), 645–649.
Lave, J., Wenger, E. (1999). *Legitimate peripheral participation. Learners, learning and assessment.* London: The Open University, pp. 83–89.

How to identify lurkers (1/4)

- Two main features: *seldom posting*, *mostly reading contents*

- Attempts to set quantitative standards:
  - "never post in an online community" [Nonnecke et al., 2006]
  - "post messages only once in a long while" [Golder & Donath, 2004]
  - "no contribution during a 3-month period" [Nonnecke & Preece, 2000]
  - "#posts<4 from the beginning, or never posted in the last 4 months" [Ganley et al., 2012]

- Accounting for the "login" dimension [Chen, 2004]
  - Lurkers log into the community every week throughout a 6-week timespan

Golder, S. A., Donath, J. (2004). Social roles in electronic communities. *Internet Research*, 5, 19–22.
Ganley, D., Moser, C., Groenewegen, P. (2012). Categorizing behavior in online communities: A look into the world of cake bakers. In *Proc. HICSS*, pp. 3457–3466.
Chen, F. C. (2004). Passive forum behaviors (lurking): A community perspective. In *Proc. 6th Int. Conf. on Learning Sciences*, pp. 128–135.

How to identify lurkers (2/4)

- Find a certain percentage of most non-active users as lurkers
  - e.g., [Rau et al., 2008] On Microsoft's Wallop SNS, 40% of the most non-active as lurkers

- Two continuous dimensions (participation pattern) [Leshed, 2005]:
  - *Publicity*: ratio of public (i.e., posting) to non-public (i.e., reading) activities
  - *Intensity*: the frequency of total activities performed by a member
    - Lurkers tend to have higher intensity and lower publicity

Rau, P.-L. P., Gao, Q., Ding, Y. (2008). Relationship between the level of intimacy and lurking in online social network services. Computers in Human Behavior, 24(6), 2757–2770.

Leshed, G. (2005). Posters, lurkers, and in between: A multidimensional model of online community participation patterns. In *Proc. HIC*.

How to identify lurkers (3/4)

- Lurkers may be classified into: [Takahashi et al. 2003; Walker et al. 2013]
  - *Passive* lurkers: only read for their use
  - *Active* lurkers: for propagation, practical use, or personal contact

- Lurkers vs. "non-users" [Springer et al. 2015]
  - Lurking as passive participation, as opposed to commenting (active participation)
  - Non-users: read news but have no interest in the user comments/discussions

Takahashi, M., Fujimoto, M., Yamasaki, N. (2003). The active lurker: Influence of an in-house online community on its outside environment. In *Proc. ACM SIGGROUP Conf. on Supporting Group Work*, pp. 1–10.
Walker, B., Redmond, J., Lengyel, A. (2013). Are they all the same? Lurkers and posters on the net. *eCULTURE*, 3(1).
Springer, N., Engelmann, I., Pfaffinger, C. (2015). User comments: motives and inhibitors to write and read. *Information, Communication & Society*, 18(7): 798-815

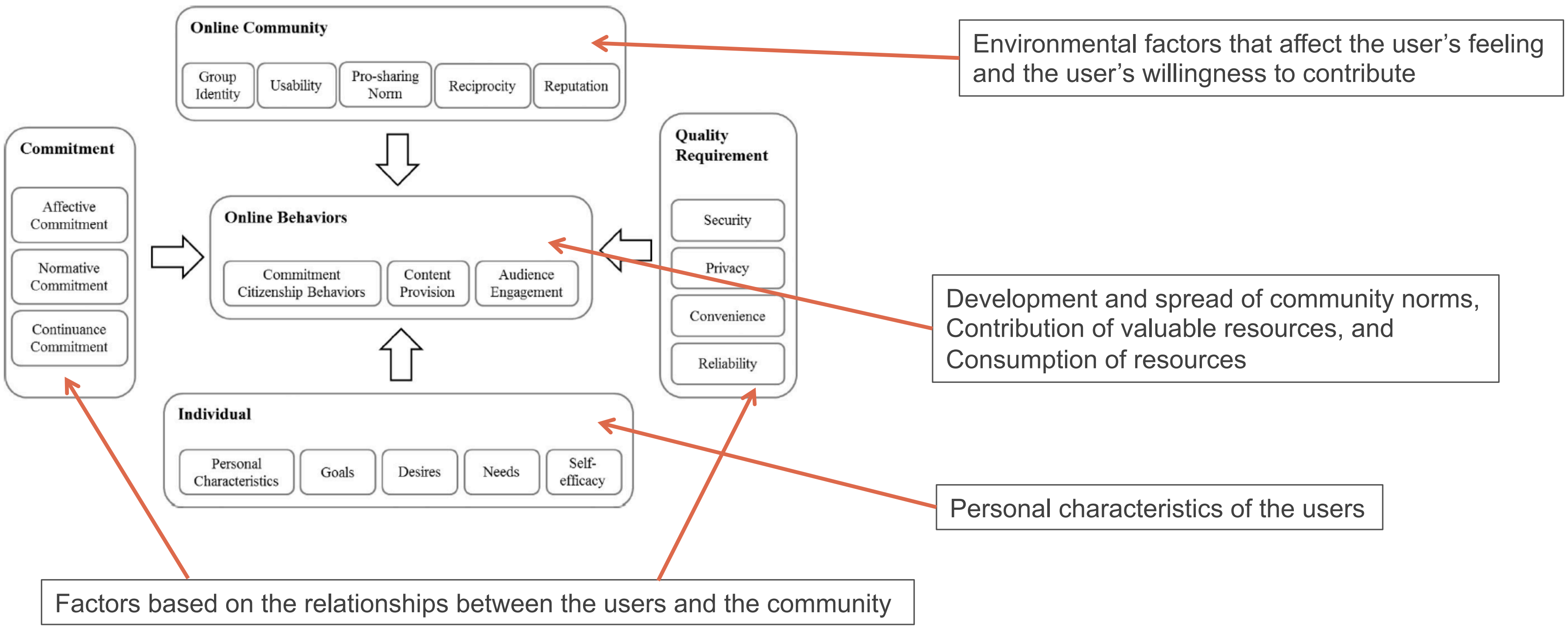
How to identify lurkers (4/4)

- Can we generalize using the previously discussed criteria?
  - No, it depends on the size, topics and culture of the online community!
  - Many factors influence online behaviors (e.g., [Bishop, 2007; Fan et al., 2009]):
    - Environmental influences
    - Personal characteristics
    - Organizational commitment

- Many lurkers: good or bad?
  - Active lurkers are beneficial for the propaganda and development of the community
  - but they have low posting rate and lack of valuable content
  - Emergence for strategies to promote *de-lurking*

Bishop, J. (2007). Increasing participation in online communities: A framework for human–computer interaction. *Computers in Human Behavior*, 23(4), 1881–1893.
Fan, Y.-W., Wu, C.-C., Chiang, L.-C. (2009). Knowledge sharing in virtual community: The comparison between contributors and lurkers. In *Proc. Int. Conf. on Electronic Business*, pp. 662–668.

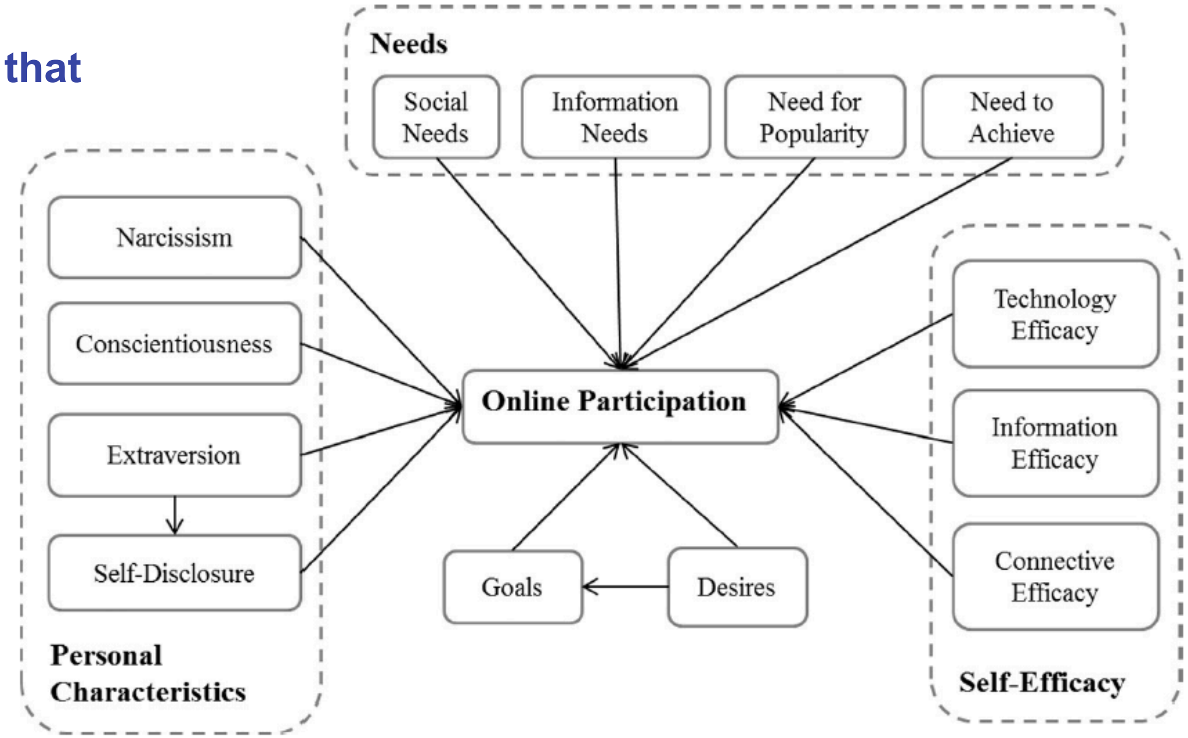
Lurking and online behavioral models (1/2)



Sun, N., Rau, P. P.-L., Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.

Lurking and online behavioral models (2/2)

Individual factors that influence online behaviors



Sun, N., Rau, P. P.-L., Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.

Why lurkers lurk (1/4)

- Four main motivational factors [Sun et al., 2014]:
  1. Environmental influence determined by the online community
  2. Personal preference related to an individual's personality
  3. Relationships between the individual and the community
  4. Security considerations

Sun, N., Rau, P. P.-L., Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.

Why lurkers lurk (2/4)

1. **Environmental influence**
   - Bad usability/interaction design
   - “Too many or too few messages to deal with”
   - Poor quality of the posted contents [Springer et al., 2015]
     - Negatively influences the affective/entertainment dimension of *gratification sought*
   - “Don’t know how to post”
     - [Nonnecke et al., 2004] Survey of 1188 users from 375 MSN online communities: 7.8% of lurkers
       - caused by poor usability and insufficient usage guidance
   - Low response rate and long response delay
   - Low reciprocity
     - [Fan et al., 2009] Survey with 207 valid responses (74% of lurkers)
     - Leads to think that “posting has no value to me”
   - “Others respond the way I would”
   - “Just reading/browsing is enough”, “No requirement to post”
     - [Kucuk, 2010] Survey of 1078 online course students: 31.1% of lurkers

Nonnecke, B., Preece, J., Andrews, D., Voutour, R. (2004). Online lurkers tell why. In *Proc. AMCIS*.
Kücük, M. (2010). Lurking in online asynchronous discussion. *Procedia-Social and Behavioral Sciences*, 2(2), 2260–2263.

Why lurkers lurk (3/4)

2. Personal reasons

- Introversion, lack of self-efficacy, bashfulness [Nonnecke et al., 2004]
- Lack of confidence in the ability to post [Lee al., 2006]
  - 40% of inactive students of an online program [Beaudoin, 2002]
- "Don't feel comfortable writing ideas online"
  - 25% of inactive students of an online program [Beaudoin, 2002]
- No need to post – only seeking for information
- Nothing to post or lack of expertise
- "Others had already posted similarly"
- Time constraints
- Missing opportunity to earn money (e.g., with commenting activities) [Springer et al., 2015]

Lee, Y.-W., Chen, F.-C., Jiang, H.-M. (2006). Lurking as participation: A community perspective on lurkers' identity and negotiability. In *Proc. Conf. on Learning Sciences*, pp. 404–410.

Beaudoin, M. F. (2002). Learning or lurking?: Tracking the "invisible" online student. *The Internet and Higher Education*, 5(2), 147–155.

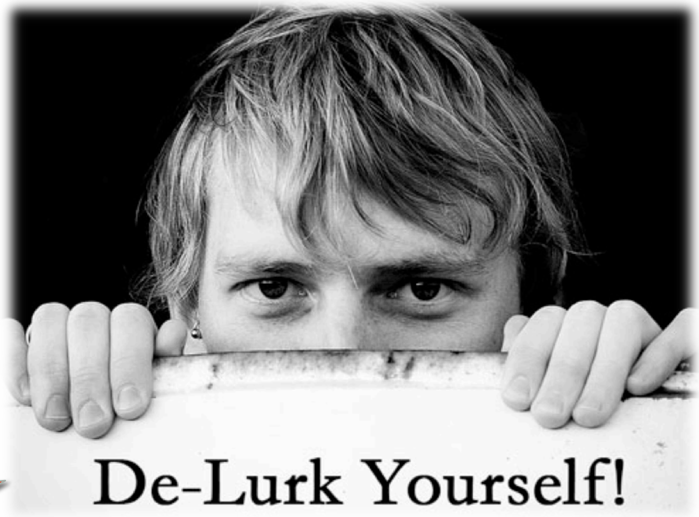
Why lurkers lurk (4/4)

3. **Relationships reasons**
   - Low verbal and affective intimacy with other members
     - *Social penetration theory* [Altman & Taylor, 1973]: intimacy develops over time to the extent that members reciprocate disclosures
   - Lack of commitment to the group
   - Fear making a commitment
   - Don’t want to spend too much time/resources to maintain a commitment
4. **Security reasons**
   - Worrying about that posting will violate privacy [Nonnecke & Preece, 2001; Springer et al., 2015
   - The community does not satisfy requirements of security and privacy, at different levels [Wang et al., 2011]

Altman, I., Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. New York: Holt, Rinehart & Winston.
Wang, Y., Norice, G., Cranor, L. F. (2011). Who is concerned about what? A study of American, Chinese and Indian users’ privacy concerns on social network sites. In *Proc. Trust and Trustworthy Computing*, pp. 146–153. Springer.

The challenge of "de-lurking"

Provide an environment that makes people's lives easier



How to promote de-lurking (1/3)

- **External stimuli** – *Social Exchange theory* [Thibaut & Kelley, 1959]
  - Providing rewards and removing negative consequences will strengthen intentions
  - Main actions:
    - Tangible or intangible rewards
    - Controlling or informative rewards

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups* (Vol. XIII). Oxford, England: John Wiley.

- **Encouragement to participate** [Nonnecke et al., 2004; Du, 2006]
  - Helps to set up a pro-sharing norm
  - Enhances users' commitment to the community
  - Improves users' confidence in expressing themselves
  - Make lurkers understand the necessity of their contribution
  - Main actions:
    - Welcome statements, introduction of reward rules, support for browsing and praise for the moderator

Sun, N., Rau, P. P.-L., Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.

How to promote de-lurking (2/3)

- **Guidance for newcomers** [Du, 2006]
  - Newcomers are likely to lurk for a while to learn the culture of the community
  - Help from elder/master users
  - Periodically provide opportunities to join conversations

- **Usability improvement** [Nonnecke et al., 2004, 2006; Du, 2006]
  - Simplify the procedures to send/respond messages
  - Rearranging the presentation of messages

Sun, N., Rau, P. P.-L., Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.

How to promote de-lurking (3/3)

- **Usability improvement** [Nazi et al., 2015]
  - Simplify the task of product/service reviewing
  - Given:
    - User feedback in textual form
    - A user and an item to review
  - Goal:
    - Recommend a set of meaningful terms (i.e., tags) to the user
  - Method:
    - Extraction of key tags from available reviews according to:
      - Relevance, Coverage, and Polarity properties
    - Formulation of top-k meaningful tags identification
      - Independent Coverage TagAdvisor
      - Dependent Coverage TagAdvisor

Nazi, A., Das, M., Das, G. (2015). The TagAdvisor: Luring the Lurkers to Review Web Items. In *Proc. ACM SIGMOD 2015*, 531–543.

Lurking as a computational problem (1/2)

- Hot topic in social science and computer-human interaction
  - Lurking conceptualized in relation to cultural capital [Soroka & Rafaeli, 2006], boundary spanning and knowledge brokering activities [Craneeld et al., 2011], group learning [Chen & Chang, 2011], epistemic curiosity [Schneider et al., 2013]
  - Focus on the identification of insights that might drive empirical evaluation of lurkers' traits
- Also becoming mature in computer science
  - Classification methods for actors in an OSN [Fazeen et al., 2011]
    - including lurkers, although treated marginally
  - Active and passive lifetime [Lang & Wu, 2013]
    - the latter however requires to know the user's last login date

Soroka, V., Rafaeli, S. (2006). Invisible participants: how cultural capital relates to lurking behavior. In *Proc. ACM WWW*, pp. 163-172.
Chen, F. C., Chang, H. M. (2011). Do lurking learners contribute less?: a knowledge co-construction perspective. In *Proc. C&T*, pp. 169-178.
Craneeld, J., Yoong, P., Hu, S. L. (2011). Beyond Lurking: The Invisible Follower-Feeder In An Online Community Ecosystem. In *Proc. PACIS*.
Schneider, A., von Krogh, G., Jager, P. (2013). "What's coming next?" Epistemic curiosity and lurking behavior in online communities. *Computers in Human Behavior* 29, 293-303
Fazeen, M., Dantu, R., Guturu, P. (2011) Identication of leaders, lurkers, associates and spammers in a social network: context-dependent and context-independent approaches. *Social Netw. Analys. Mining* 1(3), 241-254
Lang, J., Wu, S. F. (2013). Social network user lifetime. *Social Netw. Analys. Mining* 3(3), 285-297.

Lurking as a computational problem (2/2)

- Emergence for computational models, methodologies, and algorithms for
  - Understanding lurking behaviors to improve
    - User modeling, personalization and adaptation
  - Utilizing the mined knowledge in next-generation
    - Marketing-oriented applications
    - E-learning platforms
    - Collaborative systems
    - Trust systems

Next …

Modeling lurking behaviors

Topology-driven definition of lurking

In-, Out-, and InOut-neighbors driven ranking methods

Evaluation on Twitter, FriendFeed, Flickr, Google+, and Instagram

- Reciprocity, preferential attachment
- Delurking-oriented randomization model
- Percolation/resilience analysis

Lurking over time

- Lurkers vs. inactive users , and newcomers
- Responsiveness
- Preferential attachment
- Temporal trends and clustering
- Topic evolution

Vicariously Learning on RCNs

- VLRank methods

Lurking and Social Trust

- Trust-biased LurkerRank methods

Delurking via Targeted Influence Maximization

- The DEvOTION method

A. Tagarelli, R. Interdonato (2014)
Lurking in Social Networks: Topology-based Analysis and Ranking Methods.
Soc. Netw. Analys. Mining (SNAM)

A. Tagarelli, R. Interdonato (2013)
"Who's out there?" Identifying and Ranking Lurkers in Social Networks.
In *Proc. ASONAM'13*

MODELING LURKING BEHAVIORS

- In-degree, Out-degree and Lurking
- Topology-driven Lurking definition

Modeling lurking behaviors (1/4)

- Social network as a **graph**
  - Users as nodes
  - User relations as edges

- Objective:
  - Define a **lurking score** function
  - Use this function to produce a **ranking of users** at different degrees of lurking

- Assumptions:
  - edges are directed
    - i.e., user relations are asymmetric: *followships*, or *interactions*
    - In-neighbors, out-neighbors
  - nodes correspond to users only
  - (optionally) edge weights might be provided

Modeling lurking behaviors (2/4)

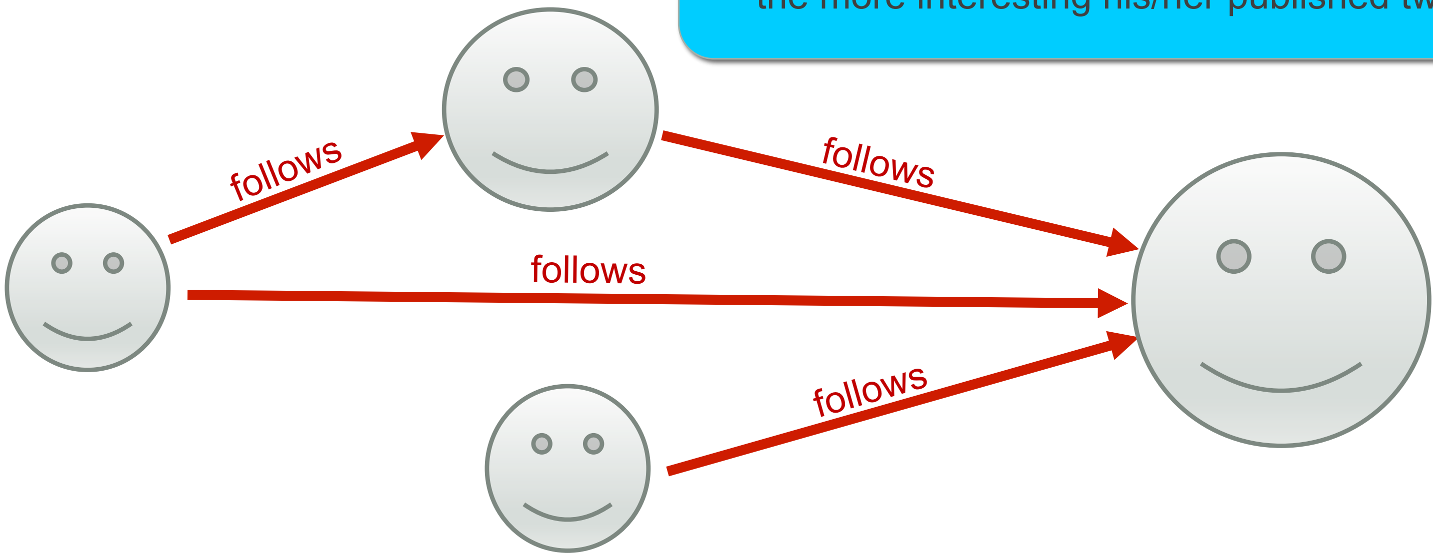
- Centrality in (social) networks
  - Many definitions, function of
    - Local topology structure
      - Degree, closeness, betweenness
    - Global topology structure
      - Propagation and attenuation of information
      - PageRank, hubs and authorities, etc.
  - Can be topic-biased
    - e.g., TwitterRank
- Other terms: prestige, importance, authoritativeness, influential status, etc.

- What about "*lurking centrality*"?

Modeling lurking behaviors (3/4)

- User interactions in a SN are typically modeled as **influence-oriented relationships**, to identify and rank **influential users**

twitter
the more followers a user has,
the more interesting his/her published tweets



Modeling lurking behaviors (4/4)

The greater the amount of information a node receives, the more likely it corresponds to a lurker

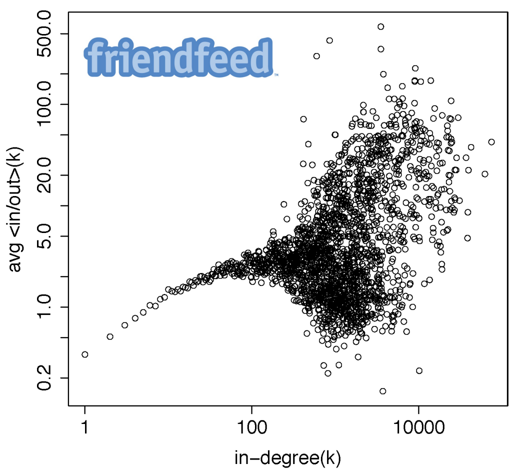
twitter

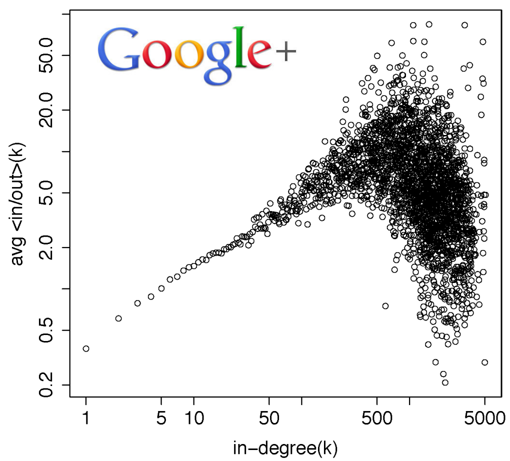
if **user *A* follows user *B***, then *A* is benefiting from *B*'s information (i.e., *A* is receiving *B*'s tweets) ➔ **link from *B* to *A***

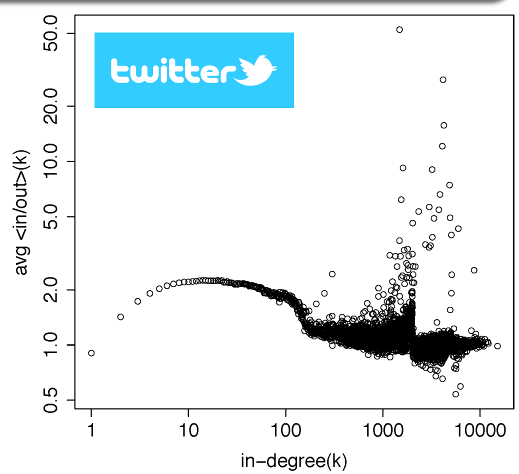
Topology-driven definition of lurking (1/3)

Modeling the mutual contribution from incoming and outgoing links through the **in/out-degree**

Is in/out-degree correlated with in-degree?







Topology-driven definition of lurking (2/3)

- Need to capitalize on a node's incoming and outgoing connections

The strength of the lurking status of a node is proportional to:

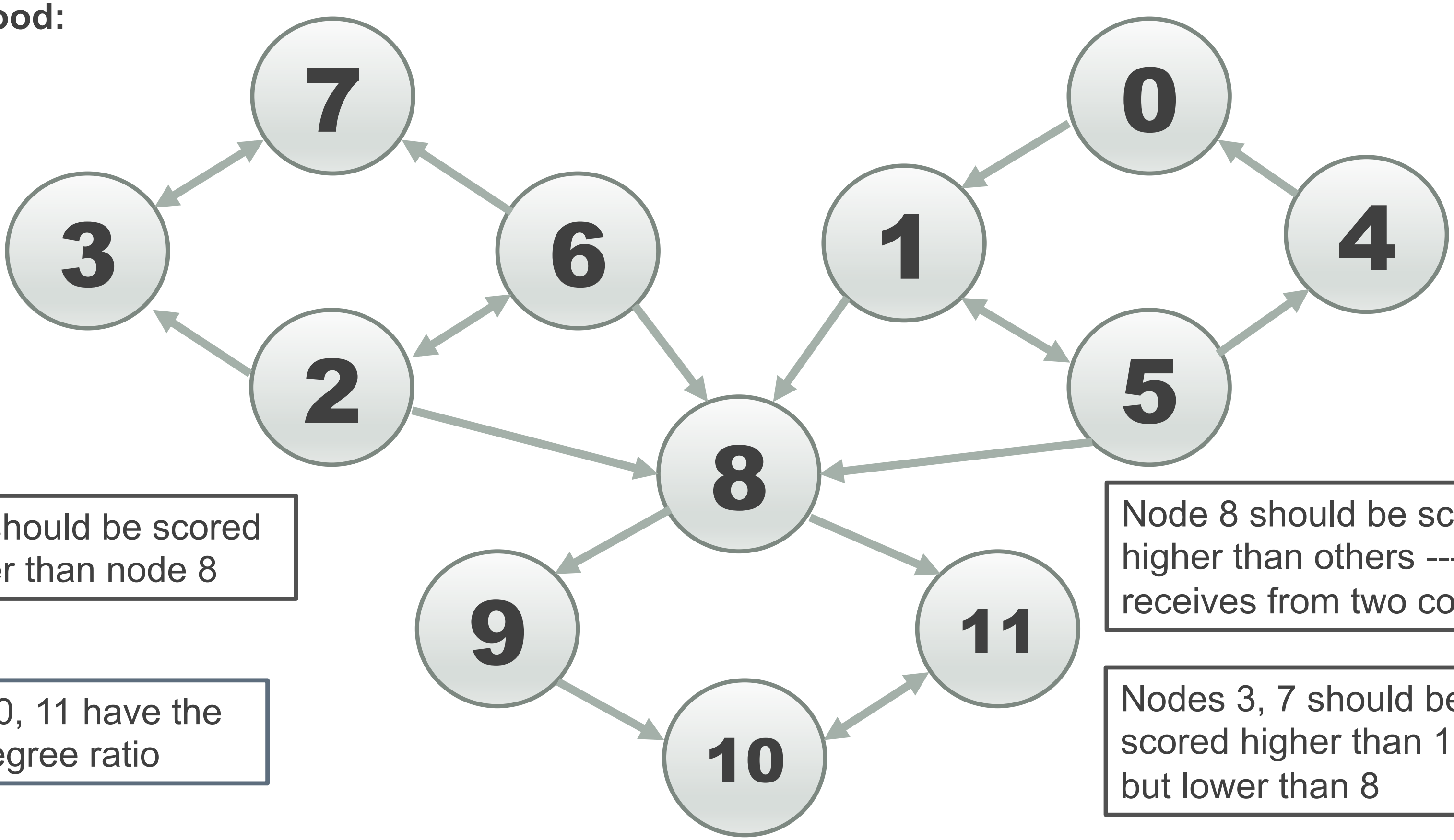
Principle I - Overconsumption:
- its own in/out-degree

Principle II - Authoritativeness of the information received:
- the influential (non-lurking) status of its in-neighbors

Principle III: Non-authoritativeness of the information produced:
- the lurking status of its out-neighbors

Topology-driven definition of lurking (3/3)

Lurking likelihood:



Nodes 10, 11 should be scored as lurkers lower than node 8

Nodes 3, 7, 8, 10, 11 have the highest in/out-degree ratio

Node 8 should be scored higher than others --- it receives from two components

Nodes 3, 7 should be scored higher than 10, 11 but lower than 8

A. Tagarelli, R. Interdonato (2014)
Lurking in Social Networks: Topology-based Analysis and Ranking Methods.
Soc. Netw. Analys. Mining (SNAM)

A. Tagarelli, R. Interdonato (2013)
"Who's out there?" Identifying and Ranking Lurkers in Social Networks.
In *Proc. ASONAM'13*

Lurker Ranking Methods

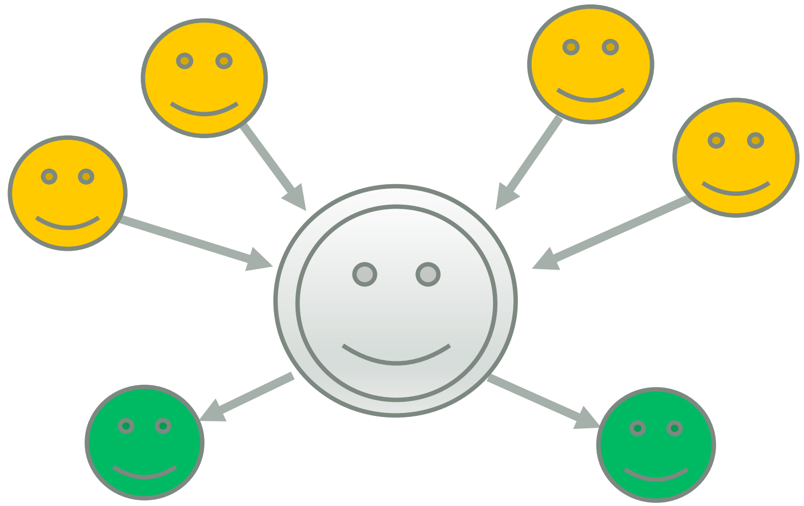
- In-neighbors- and out-neighbors-driven lurking definitions
- PageRank and AlphaCentrality based formulations

In-neighbors-driven lurking

$$r_i = \boxed{\frac{1}{out(i)}} \boxed{\sum_{j \in B_i} \frac{out(j)}{in(j)} r_j}$$

The score of a node increases with the number of its in-neighbors and with their likelihood of being non-lurkers (relatively high out/in-degree)

Factor inversely proportional to the node's out-degree accounts for its own in/out-degree property

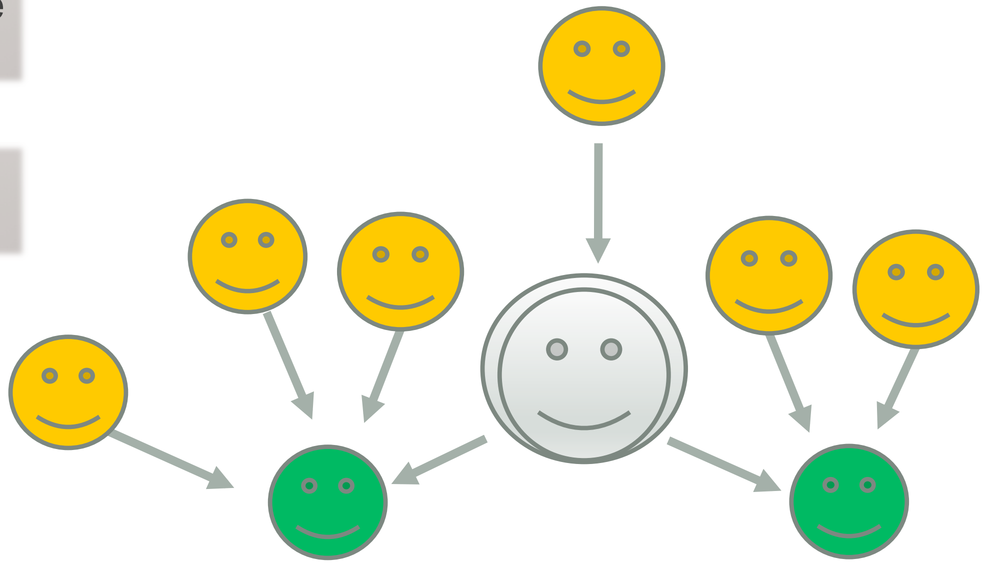


Out-neighbors-driven lurking

$$r_i = \boxed{\frac{in(i)}{\sum_{j \in R_i} in(j)}} \boxed{\sum_{j \in R_i} \frac{in(j)}{out(j)} r_j}$$

The lurking score of a node increases with the tendency of its out-neighbors of being lurkers

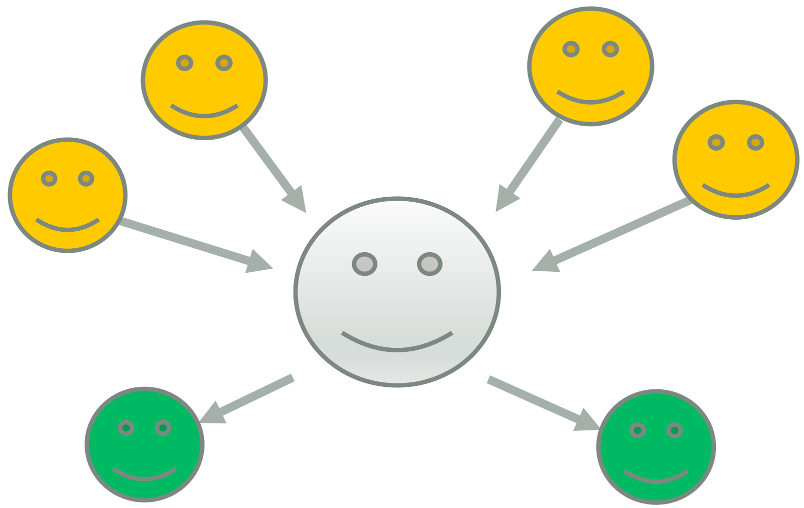
Factor scoring a node higher if it receives more than what its out-neighbors receive



In-Out-neighbors-driven lurking

$$r_i = \left(\frac{1}{out(i)} \sum_{j \in B_i} \frac{out(j)}{in(j)} r_j \right) \left(1 + \left(\frac{in(i)}{\sum_{j \in R_i} in(j)} \sum_{j \in R_i} \frac{in(j)}{out(j)} r_j \right) \right)$$

Aspect related to the strength of non-lurking behavior of in-neighbors is dominant – it's expected to have a better fit of the hypothetical likelihood function for a given node



LurkerRank methods (1/2)

- Specification in terms of classic **PageRank** [Brin & Page, 1998] and **alpha-centrality** [Bonacich & Lloyd, 2001]
  - PageRank equations

$$\mathbf{r} = \alpha \mathbf{S}\mathbf{r} + (1-\alpha)\mathbf{v} \qquad\qquad r_i = \alpha \sum_{j \in B_i} \frac{r_j}{out(j)} + \frac{1-\alpha}{N}$$

  - Alpha-centrality equations

$$\mathbf{r} = \alpha \mathbf{A}^T \mathbf{r} + \mathbf{v} \qquad\qquad r_i = \alpha \sum_{j \in B_i} r_j + v_i$$

Brin, S., Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems* 30(1-7), 107-117.
Bonacich, P., Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks* 23, 191-201.

LurkerRank methods (2/2)

- PageRank and AlphaCentrality based formulations
  - In-neighbors-driven lurking methods: **LRin**, **ac-Lrin**
  - Out-neighbors-driven lurking methods: **LRout**, **ac-LRout**
  - InOut-neighbors-driven lurking methods: **LRin-out**, **ac-LRin-out**

- e.g.,: **LRin** formulation

$$r_i = \alpha \left(\frac{1}{out(i)} \sum_{j \in B_i} \frac{out(j)}{in(j)} r_j \right) + \frac{(1-\alpha)}{N}$$

A. Tagarelli, R. Interdonato (2014)
Lurking in Social Networks: Topology-based Analysis and Ranking Methods.
Soc. Netw. Analys. Mining (SNAM)

A. Tagarelli, R. Interdonato (2013)
"Who's out there?" Identifying and Ranking Lurkers in Social Networks.
In *Proc. ASONAM'13*

Experimental Evaluation

- Data
- Assessment methodology
- Quantitative and Qualitative Results

Network datasets

data	*# nodes*	*# links*	*avg in-degree*	*avg path length*	*clustering coefficient*	*assortativity*	*# sources* *# sinks*	*LC* *wLC*
flickr	2,302,925	33,140,018	14.39	4.36*	0.107	0.015	360,416 57,424	0.573 0.248
friendfeed	493,019	19,153,367	38.85	3.82	0.029	-0.128	41,953 292,003	0.955 0.354
Google+	107,612	13,673,251	127.06	3.32	0.154	-0.074	35,341 22	0.869 0.096
Instagram	54,018	963,883	17.85	4.50	0.048	-0.067	28,719 1,310	0.910 0.273
twitter-*Kwak*	16,009,364	132,290,000	8.26	5.91*	1.26E-4	-0.095	1,067,936 10,298,788	0.914 0.435
twitter-*UDI*	24,984,590	284,884,500	11.40	5.45*	4.96E-3	-0.297	3,380,805 8,065,287	0.790 0.470

Mislove, A., Koppula, H. S., Gummadi, K. P., Druschel, P., Bhattacharjee, B. (2008). Growth of the Flickr Social Network. In *Proc. ACM WOSN*.

Celli, F., Lascio, F. M. L. D., Magnani, M., Pacelli, B., Rossi, L. (2010). Social Network Data and Practices: The Case of FriendFeed. In *Proc. SBP*.

McAuley, J. J., Leskovec, J. (2012). Learning to Discover Social Circles in Ego Networks. In *Proc. NIPS*.

Ferrara, E., Interdonato, R., Tagarelli, A. (2014). Online popularity and Topical Interest through the lens of Instagram. In *Proc. HT*.

Kwak, H., Lee, C., Park, H., Moon, S. B. (2010). What is Twitter, a social network or a news media?. In *Proc. ACM WWW*.

Li, R., Wang, S., Deng, H., Wang, R., Chang, K. C. C. (2012). Towards social user profiling: unified and discriminative in influence model for inferring home locations. In *Proc. ACM SIGKDD*.

Dealing with lack of ground-truth

- Generating a **data-driven ranking** for each evaluation dataset
- **Basic idea**:
  - Directly proportional to a node's in/out-degree
  - Inversely proportional to a SN-specific measure of influence

$$r_i^* = \frac{in(i)}{out(i)} e^{-EI(i)}$$

twitter [Bakshy, et al., 2011]

$$EI(i) = \frac{1}{out(i)} \sum_{j \in R_i} nRetweets(j)$$

friendfeed

$$EI(i) = \frac{\sum_{j \in R_i} nCom(j,i)}{out(i)} \log_{10}(10 + nPosts(i))$$

flickr DD-F

$$EI(i) = \frac{1}{out(i)} \sum_{j \in R_i} nFavorites(j)$$

flickr DD-V

$$EI(i) = \frac{1}{out(i)} \sum_{j \in R_i} nViews(j)$$

Bakshy, E., Hofman, J. M., Mason, W. A., Watts, D. J. (2011). Everyone's an influencer: quantifying influence on Twitter. In *Proc. ACM WSDM*.

Competing methods

- (baseline) In-Out distribution (IO)
- PageRank (PR)
- Alpha-centrality (AC)
- Fair-Bets [Budalakoti & Bekkerman, 2012] (FB)

$$r_i = \frac{1}{out(i)} \sum_{j \in B_i} r_j$$

  - Connections among the users are based on the number of sent and accepted invitations
  - Fair-Bets can be viewed as a model of social capital accumulation and expenditure
  - Assuming users are paying each other to accept invitations on a SN, the fair-bets score of a user is *the amount s/he can afford to pay on average*

Budalakoti, S., Bekkerman, R. (2012). Bimodal invitation-navigation fair bets model for authority identification in a social network. In *Proc. ACM WWW*.

Assessment criteria (1/2)

- **Fagin's intersection metric**: determines how well two ranking lists are in agreement with each other, accounting for top-weightedness:

$$F(L^i, L^{ii}, k) = \frac{1}{k}\sum_{q=1}^{k}\frac{L^i_{:q} \cap L^{ii}_{:q}}{q}$$

- **Kendall rank correlation coefficient**: evaluates the similarity between two rankings, expressed as sets of ordered pairs, based on the number of inversions of pairs which would be needed to transform one ranking into the other.

$$Ken(L^i, L^{ii}) = 1 - \frac{(2\Delta(P(L^i), P(L^{ii})))}{(N(N-1))}$$

Assessment criteria (2/2)

- **Bpref (binary preference)**: preference relation of whether judged relevant candidates *R* of a list L_1 are retrieved, i.e., occur in a list L_2, ahead of judged irrelevant candidates *N*:

$$Bpref(R,N) = \frac{1}{|R|}\sum_r \frac{1-(\#\,of\ n\ ranked\ higher\ than\ r)}{|R|}$$

  - *N*: set of nodes with data-driven ranking score below or equal to 1
  - *R* is selected as the set of nodes having top-k% score from the complement of *N*.

Evaluation goals

- Lurking reciprocity: how lurkers relate to each other?
- Lurkers-active users attachment: how lurker distribution grows w.r.t. active users (and vice versa)?
- Ranking evaluation:
  - Correlation analysis w.r.t. data-driven rankings
  - Comparative evaluation with LurkerRank methods
  - Efficiency performance
- Delurking-oriented randomization
- Percolation analysis

- Qualitative analysis
  - Manually inspecting web profiles of top-lurkers

Reciprocity (1/2)

Impact of the presence of lurkers on measures of reciprocity, based on top-25%, top-10%, and top-5% of a LR solution

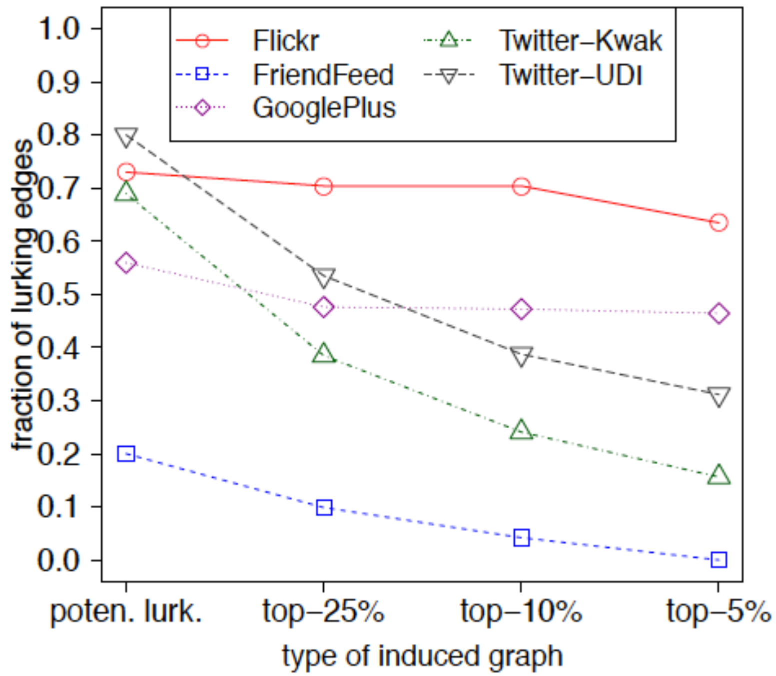
		top-25% of the LRin-out solution			top-10% of the LRin-out solution			top-5% of the LRin-out solution		
	# recip. edges (full graph)	*# edges (induced graph)*	*# reciprocal lurking edges*	*% rle*	*# edges (induced graph)*	*# reciprocal lurking edges*	*% rle*	*# edges (induced graph)*	*# reciprocal lurking edges*	*% rle*
Flickr	20,603,483	23,352,367	16,440,872	49.61	12,349,595	8,704,922	26.27	5,030,759	3,192,712	9.63
FriendFeed	3,014,306	340,935	33,654	0.18	1,096	46	<0.01	2	0	0.00
GooglePlus	2,870,336	1,413,468	667,422	4.88	49,481	23,562	0.17	5,310	2,624	0.02
Twitter-Kwak	52,137,192	7,293	2,806	<0.01	216	52	<0.01	64	10	<0.01
Twitter-UDI	191,858,256	18,839,845	10,078,339	3.54	3,094,341	1,198,615	0.42	872,332	271,751	0.10

- Small or negligible
  - fraction of reciprocal lurking edges to the total no. of edges in the original graph (rle)
  - fraction of reciprocal edges in the original graph that connect lurkers to each other
- LRin performed very similarly to LRin-out
- LRout achieved much higher values (as expected)

Reciprocity (2/2)

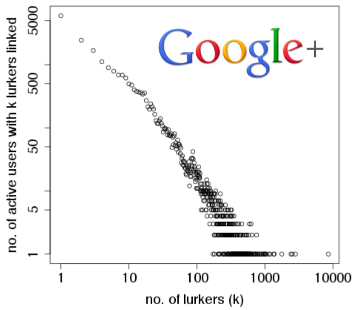
Fraction of edges that connect lurkers to each other in a lurking-induced subgraph

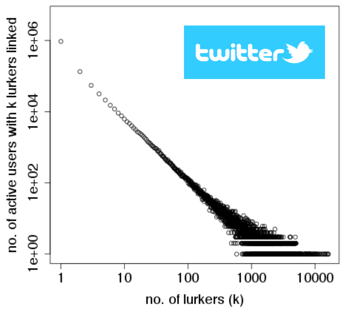
- Decreasing trend for lurking reciprocity (LRin-out)
  - stagnant on Flickr, GooglePlus, and FriendFeed
  - steeper on Twitter
- Inverse trend when using LRout



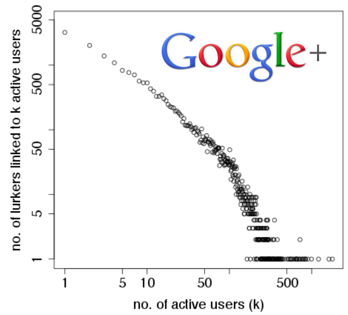
Attachment

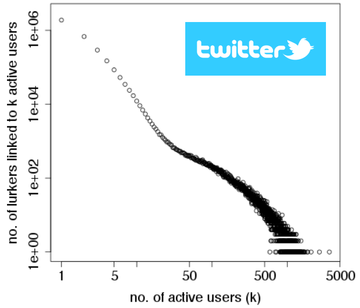
Distribution of active users as a function of the lurkers-followers





Distribution of lurkers as function of the active users-followees





Active users who already are followed by a large number of lurkers, are likely to attract even more lurkers

Lurkers who already follow a large number of active users, are more likely to do so

Ranking evaluation:

Correlation analysis w.r.t. data-driven rankings

Kendall tau correlation (95% confidence intervals)

dataset	IO	PR	AC	ER	LRin	LRout	LRin-out	ac-LRin	ac-LRout	ac-LRin-out
FriendFeed	.169 (± .003)	.128 (± .004)	.230 (± .005)	.373 (± .004)	.661 (± .003)	-.169 (± .005)	.497 (± .003)	**.664** (± .003)	-.189 (± .005)	.470 (± .003)
Flickr vs DD-V	.046 (± .008)	.043 (± .005)	.043 (± .008)	.047 (± .002)	.247 (± .007)	-*.007* (± .013)	.239 (± .014)	.234 (± .014)	*.011* (± .014)	**.251** (± .013)
Flickr vs DD-F	.052 (± .007)	.049 (± .005)	.049 (± .008)	.053 (± .002)	.231 (± .006)	*.003* (± .012)	.260 (± .013)	.255 (± .013)	*.011* (± .014)	**.273** (± .012)
Twitter-Kwak	.171 (± .006)	*.004* (± .011)	.215 (± .010)	.235 (± .012)	**.671** (± .007)	-.082 (± .004)	.559 (± .008)	.659 (± .008)	-.073 (± .004)	.560 (± .008)

- Highest correlation for LRin and LRin-out (and their ac- counterparts)
- Low correlation for LRout and ac-LRout
  - Hint: Principle III tends to weight less than Principles I-II in effectively lurker ranking
- Poor correlation shown by the other methods
  - Hint: in/out-degree cannot approximate well LurkerRank

Ranking evaluation:
Comparative evaluation with LurkerRank methods

twitter	F $k = 10^2$ // 10^3 // 10^4					$Bpref$ $l = 10$ // 25 // 50				
	DD	IO	PR	AC	FB	DD	IO	PR	AC	FB
LRin	.527 .289 .581	.404 .209 .617	0.0 0.0 .001	0.0 0.0 .001	.112 .127 .068	.997 .995 .985	.992 .989 .962	.121 .473 .521	.790 .914 .866	.441 .704 .606
LRout	.030 .008 .003	.032 .008 .002	.181 .351 .437	.010 .024 .048	.034 .015 .005	.045 .055 .109	0.0 .001 .007	.754 .757 .641	.311 .650 .678	.313 .600 .648
LRin-out	.475 .314 .666	.364 .277 .688	0.0 0.0 .001	0.0 0.0 .001	.064 .063 .032	.968 .979 .961	.981 .977 .925	.039 .387 .453	.826 .929 .878	.204 .524 .489
ac-LRin	.583 .573 .767	.459 .570 .810	0.0 0.0 .001	0.0 0.0 .001	.174 .122 .048	.993 .992 .982	.990 .988 .967	.072 .443 .501	.808 .921 .872	.339 .653 .575
ac-LRout	.038 .009 .003	.032 .008 .002	.244 .319 .362	.006 .017 .042	.036 .011 .004	.049 .059 .120	0.0 0.0 .081	.796 .775 .654	.339 .659 .687	.307 .598 .643
ac-LRin-out	.473 .278 .663	.363 .234 .685	0.0 0.0 .001	0.0 0.0 .001	.062 .062 .031	.957 .975 .957	.981 .976 .933	.039 .386 .453	.828 .930 .880	.203 .464 .454

Similar remarks for *Bpref* evaluation.

One difference: **LRin**, **LRin-out** and their **ac-** counterparts

… and very low *F* intersection w.r.t. **FB** and nearly empty w.r.t. **PR** and **AC**

LRout and **ac-LRout** show some correlation w.r.t. **PR** and nearly null with other methods

Ranking evaluation:
Statistical testing

Twitter

	Fagin evaluation			*Bpref evaluation*		
	PR	AC	FB	PR	AC	FB
LRin	2.9E-59	2.9E-59	3.2E-52	5.2E-110	1.1E-25	2.1E-65
LRout	1.1E-28	1.0E-68	1.2E-106	3.2E-50	5.5E-79	9.2E-71
LRin-out	3.0E-204	3.0E-204	2.0E-192	1.5E-89	6.7E-21	7.6E-65
ac-LRin	9.1E-193	9.1E-193	5.4E-185	1.2E-91	2.1E-25	2.7E-65
ac-LRout	4.3E-21	4.0E-61	8.3E-112	4.1E-54	1.8E-71	2.3E-73
ac-LRin-out	4.3E-197	1.1E-201	9.2E-123	5.8E-85	2.1E-21	1.0E-64

FriendFeed

	Fagin evaluation			*Bpref evaluation*		
	PR	AC	FB	PR	AC	FB
LRin	4.0E-103	1.9E-102	5.9E-86	2.6E-93	1.7E-52	3.7E-34
LRout	5.6E-66	3.1E-37	1.2E-96	1.0E-53	3.9E-59	3.4E-174
LRin-out	2.1E-178	5.0E-165	1.6E-97	7.1E-101	2.2E-50	5.8E-44
ac-LRin	5.3E-105	2.1E-106	2.2E-81	1.1E-103	3.0E-52	3.1E-35
ac-LRout	2.2E-66	6.1E-37	7.9E-96	2.9E-69	5.3E-64	9.3E-176
ac-LRin-out	7.1E-151	3.2E-161	2.0E-97	4.0E-102	5.4E-51	9.6E-49

Unpaired two-tail t-test

Samples: performance scores obtained by a ranking method w.r.t. DD for each iteration

Null hypothesis: no difference in performance w.r.t. data-driven ranking between a LurkerRank method and a competing method

Useful to confirm that the difference in performance between the LurkerRank methods and the competing ones holds on FriendFeed as well, (despite the high *Bpref* scores observed in most cases)

Ranking evaluation: Statistical testing – second stage (1/2)

- Data preparation (network-specific): 100 subgraphs extracted, each with a randomly picked seed node and roughly covering a fixed number of nodes (around 1/100 of the original network size)

- **Goal**: to stress the ranking methods performing over a pool of subnetworks with different characteristics

- Paired two-tail t-test, with samples *F* scores respectively obtained by two ranking methods w.r.t. **DD** over the same randomly generated subgraph
    - *k* was set to 10^4, hence very high for such network sizes (i.e., around 200,000 nodes)

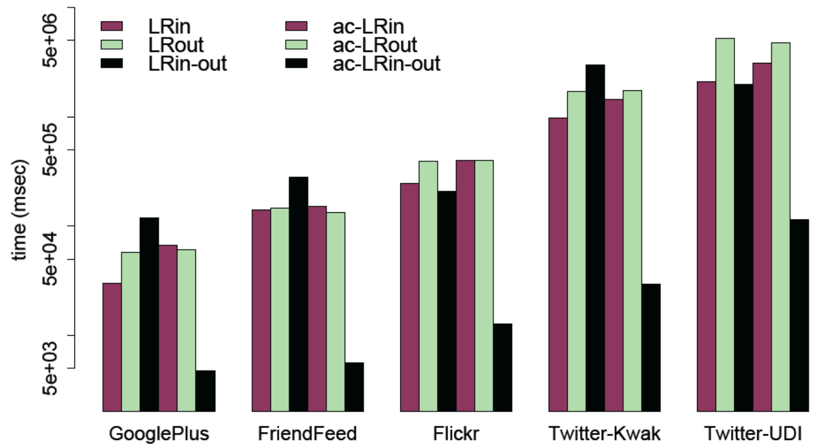
Ranking evaluation: Statistical testing – second stage (2/2)

For each pair of LurkerRank method vs. competing method, the null hypothesis of equal means was rejected at 1% significance level (p-values ranging from 1.4E-3 to 2.8E-19 on Twitter)

- Close behavior of the LurkerRank methods (except **LRout** and **ac-LRout**) and **AC** (e.g., around 0.19 *F* on average, on Twitter)
- Close behavior of **PR** and **FB**, which however achieved a lower average *F* (e.g., 0.029, on Twitter)

Ranking evaluation:
Efficiency performance of LurkerRank methods

- LRin and LRout have pretty similar runtime
- LRin-out slower than the others
  - on 3 out of 5 networks
  - about twice more than LRin and LRout



- All methods reach ranking stability quickly
  - 35 to 75 iterations
  - much fewer iterations for ac-LRin-out
    (at the cost of poor diversification of the ranking scores)

Qualitative Evaluation

rank	PR			AC			FB			LRin		
	user	*score*	*#rt*	*user*	*score*	*#rt*	*user*	*score*	*#rt*	*user*	*score*	*#rt*
1	B.O.	4.85E-03	17811	B.O.	1.59E-04	17811	D.W.S.	1.15E-05	0	R.F.	7.78E-06	0
2	W.F.	3.57E-03	1676	ZAI.	1.41E-04	10902	n.a.	6.35E-06	0	R.J.	7.72E-06	0
3	ZAP.	2.47E-03	8707	ZAP.	1.36E-04	8707	APA.	6.33E-06	0	R.M.K.	7.49E-06	0
4	TH.	1.86E-03	7169	AS.	1.35E-04	1172	T.S.C.	5.41E-06	1	B.B.P.	7.35E-06	0
5	L.E.	5.77E-04	683	M.M.	1.31E-04	7	n.a.	5.07E-06	0	TR.	6.84E-06	0
6	J.B.	5.64E-04	1248	W.F.	1.23E-04	1676	CON.	4.97E-06	0	MU.	6.04E-06	0
7	M.S.	4.87E-04	476	M.K.	1.20E-04	48	K.T.	4.95E-06	0	B.R.	5.37E-06	0
8	AS.	4.25E-04	1172	P.B.	1.10E-04	328	n.a.	4.78E-06	0	AZ.	5.30E-06	0
9	OH.	3.53E-04	1009	W.A.	1.07E-04	2814	S.M.	4.36E-06	0	O.L.	5.25E-06	0
10	H.T.	3.19E-04	43	C.B.	1.06E-04	11943	n.a.	4.06E-06	0	N.T.	5.20E-06	0
11	E.T.	3.17E-04	2435	EL.	1.04E-04	902	n.a.	3.83E-06	0	FR.	5.15E-06	0
12	SCH.	3.02E-04	3277	SCO.	1.03E-04	6970	M.P.	3.82E-06	0	D.W.S.	5.15E-06	0
13	RE.	2.93E-04	1467	WI.	1.02E-04	811	n.a.	3.81E-06	0	AW.	4.96E-06	0
14	H.S.	2.89E-04	1346	O.W.	1.02E-04	1803	n.a.	3.79E-06	0	O.B.	4.68E-06	0
15	M.M.	2.89E-04	7	T.B.B.	9.84E-05	102	M.E.	3.69E-06	0	N.C.	4.56E-06	0
16	ZAI.	2.85E-04	10902	T.S.	9.82E-05	74	B.B.P.	3.68E-06	0	D.P.	4.43E-06	0
17	SCO.	2.84E-04	6970	S.S.	9.72E-05	789	n.a.	3.68E-06	0	AU.	4.30E-06	0
18	M.K.	2.63E-04	48	M.W.	9.17E-05	363	n.a.	3.61E-06	0	EM.	4.28E-06	0
19	WI.	2.59E-04	811	H.R.	8.89E-05	750	n.a.	3.58E-06	0	DI.	4.12E-06	0
20	W.A.	2.56E-04	2814	A.K.	8.69E-05	1572	n.a.	3.57E-06	0	M.A.	3.91E-06	0

Top-20 by FB:
Most users have never been retweeted

Most of them are spammers or with profiles suspended by Twitter due to violation of terms of service

certainly can be considered lurkers

Delurking-oriented randomization (1/2)

Using randomized model to enable "**self-delurking**" of a network

- Randomization-like model to simulate introducing of lurkers to active users
- Inserting new links from active users to lurkers
- Requires:
  - cut-off thresholds for the selection of the sets of active users and lurkers
  - probability to control the degree of lurking
- Note both the size of the network and the degree of vertices may change

Algorithm Delurking-oriented randomization

Input: The topology graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ of an OSN. The ranking L corresponding to a LR solution for $\mathcal{G}$. Cut-off percentage thresholds t_1, t_2 of ranking order in L. Probability p. Maximum fraction d of new edges to add to $\mathcal{G}$.

Output: A randomized graph $\mathcal{G}'$.

1: $\mathcal{E}' \leftarrow \emptyset$
2: Sort L by decreasing lurking score
3: Let L_{top} (resp. L_{bottom}) be the top-t_1 (resp. bottom-t_2) of the sorted L
4: $E_{al} \leftarrow \{e = (a, l) \in \mathcal{E} \mid a \in L_{\mathrm{bottom}}, l \in L_{\mathrm{top}}\}$
5: **repeat**
6: Pick randomly with probability p an edge $(a_1, l_1) \in E_{al} \setminus \mathcal{E}'$
7: Pick randomly with probability p an edge $(a_2, l_2) \in E_{al} \setminus \mathcal{E}'$, with $a_2 \neq a_1, l_2 \neq l_1$
8: $\mathcal{E}' \leftarrow \mathcal{E}' \cup \{(l_1, a_2), (l_2, a_1)\}$ /* *add the new edges* */
9: **until** $(|\mathcal{E}'| \geq d|E_{al}|)$
10: $\mathcal{G}' \leftarrow \langle \mathcal{V}, \mathcal{E}' \cup \mathcal{E} \rangle$

Delurking-oriented randomization (2/2)

- Setting: p = 0.5, $t_1 = t_2$ = 25%, d in [0.2, 1.0] (increment by 0.2)

- Correlation analysis of LR solutions (resp. in/out ranking) *before/after* randomization
  - Poor when sinks/sources are discarded
- The top-ranked lurkers can significantly change
  - w.r.t. the original configuration of the network, and
  - also for different degrees of delurking-oriented randomization
  - Less evident on Twitter (larger size, lower CC, higher avgPL)

- **Negligible impact on the in/out-degree distribution**
  - Moderate to high correlation between:
    - in/out ranking in the original network and each of the in/out rankings of the randomized networks
    - the randomized in/out rankings pairwise

Percolation analysis (1/3)

- Assessing topological integrity properties
  - typically via edge removal strategies based on topological overlap measures
- Removing edges by increasing order of topological overlap has shown to effectively detect edges that act as bridges between different communities [Girvan & Newman, 2002]

Using percolation analysis to explain relationships between **lurkers** and **community bridges**

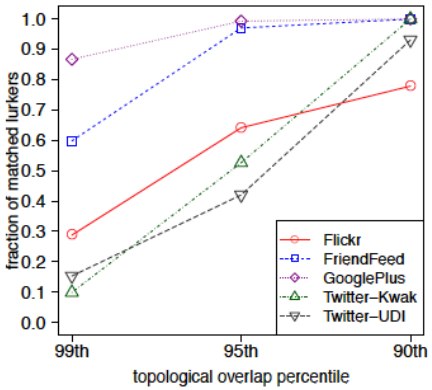
- *Directed* topological overlap: $O(i,j) = \dfrac{|R_i \cap B_j|}{(|R_i|-1) + (|B_j|-1) - |R_i \cap B_j|}$

Girvan, M., Newman, M. E. J. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. (PNAS)* 99(12), 7821-7826.

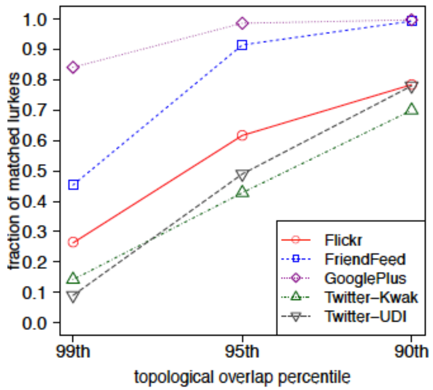
Percolation analysis (2/3)

- Comparison between
  - set of vertices resulting from edge removal based on increasing topological overlap
  - set of top-ranked lurkers

- Matching of top-25% lurkers to the sets of vertices included in the 99th, 95th and 90th percentile of the edges with lowest directed topological overlap
  - At 90th percentile, almost all top-lurkers matched on FriendFeed, GooglePlus, and (by LRin and LRin-out) on the two Twitter networks
  - On FriendFeed and GooglePlus, most top-lurkers matched at 95th

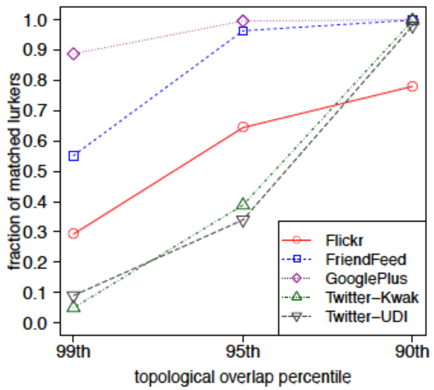


LRin



LRout

LRin-out

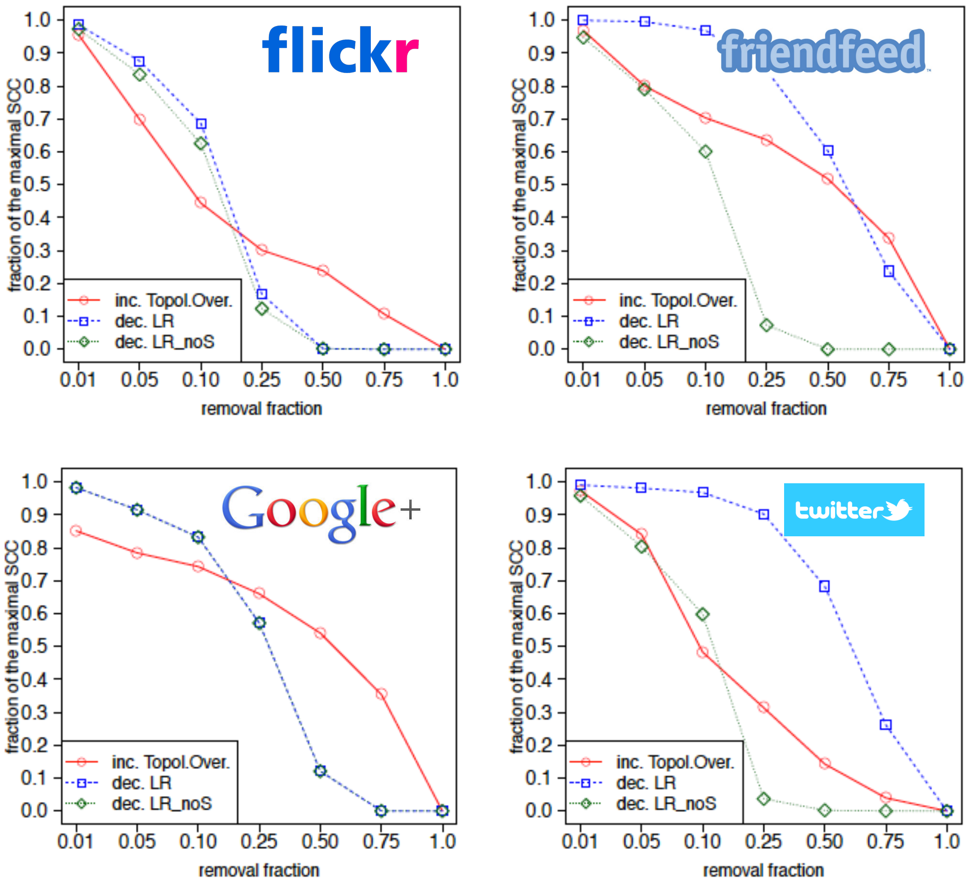


Percolation analysis (3/3)

- What fraction of the vertex set?
  - above 90% on FriendFeed and GooglePlus
  - but below 27% on Twitter

Lurkers can act as community bridges!

- Resilience evaluation: fraction of the maximal strongly CC as function of removed vertices (w/ and w/o sinks)
  - Most disruptive removal strategy based on decreasing LR (w/o sinks)



Main findings

- LurkerRank abilities:
  - Effectiveness in detecting and ranking lurkers confirmed by qualitative examination made on the evaluation SN websites
  - Higher correlation with data-driven ranking than competing and baseline methods
  - Competing methods fail in identifying lurkers:
    - PageRank and alpha-centrality still detect influential users,
    - Fair-Bets tends rather to identify spammers

- Lurking-oriented network analysis:
  - Lurkers are not very prone to reciprocate each other
  - Lurkers may be related to users playing the role of bridges between communities (under lurking-oriented graph model)
  - Self-delurking randomization can be useful to change the top-ranked lurkers in the network, while scarcely affecting the in/out degree distribution

A. Tagarelli, R. Interdonato (2015)
Time-aware Analysis and Ranking of Lurkers in Social Networks.
Soc. Netw. Analys. Mining (SNAM)

A. Tagarelli, R. Interdonato (2014)
Understanding lurking behaviors in social networks across time.
In *Proc. ASONAM'14*

EXPERIMENTAL EVALUATION

Understanding lurking behaviors over time

Understanding lurkers over time

Do lurkers match to zero-contributors?

Do lurkers match newcomers?

How frequently do lurkers respond to the others' actions?

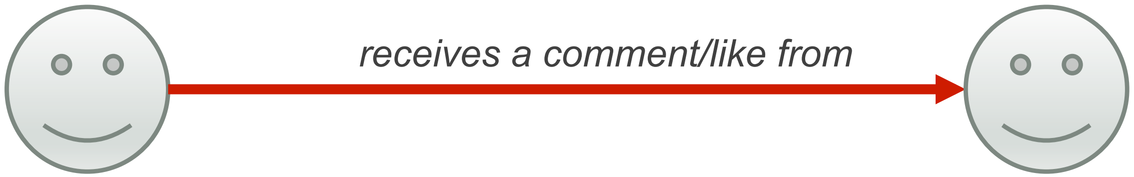
Do lurkers create preferential relations with active users?

How does lurking behavior evolve?

How do topical interests of lurkers evolve?

Time-varying snapshot graphs

- **Interaction graph:** Useful to represent evolving/dynamic lurking behavior:
  - Subgraphs of the *static* followship graph
  - Edges represent interactions among users in a certain time interval
    - Friendfeed and Instagram: **comment-based interactions**
    - Flickr: **favorite-based interactions**



- **Timestamped followship graph** [only for Flickr]

data	*# nodes*	*# links*	*avg in-deg.*	*avg path len.*	*clust. coef.*	*assorta- -tivity*
averages over time-varying snapshot graphs						
Flickr-social	1,889,102	25,265,343	13.25	4.41	0.108	0.009
Flickr	215,429	1,483,462	6.85	4.69	0.025	-0.013
FriendFeed	6,962	64,509	5.15	5.89	0.071	-0.043
Instagram	10,353	31,215	2.94	5.83	0.083	0.217

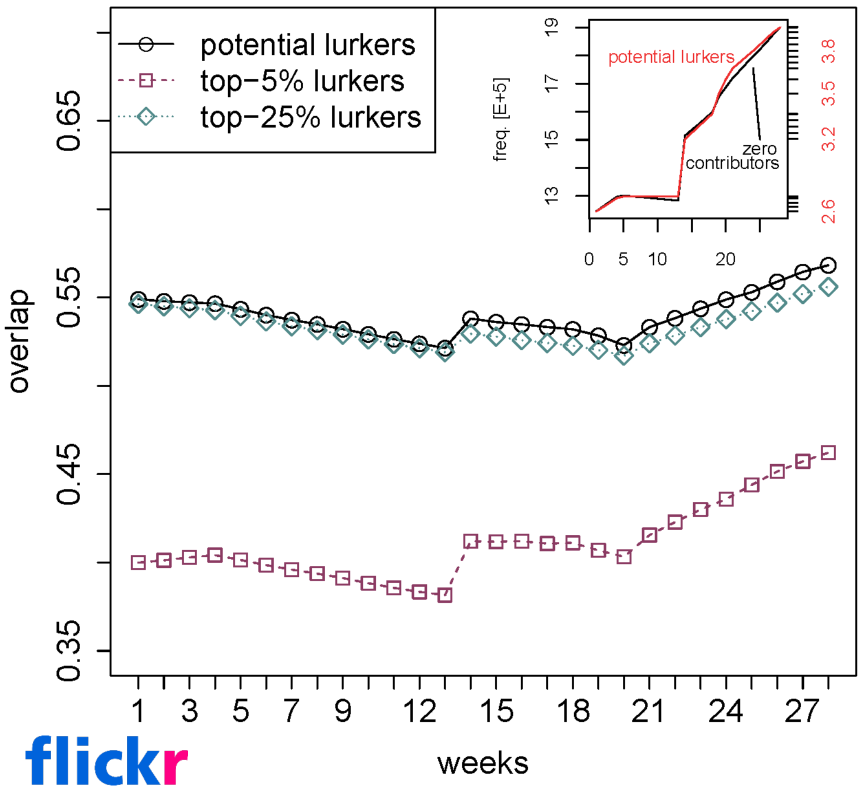
data	*start date*	*end date*
Flickr-social	2006-11-02	2007-05-17
Flickr	2006-09-08	2007-03-22
FriendFeed	2010-04-09	2010-09-30
Instagram	2012-06-28	2013-12-18

Lurkers vs. inactive users: static analysis

- How much the set of zero-contributors overlaps with the set of "*potential lurkers*" (i.e., users with in/out >1)?
  - 12% (favorite-based interaction network in Flickr)
  - 72% (comment-based interaction network in FriendFeed)
  - 95% (comment-based interaction network in Instagram)

(Potential) lurkers are more likely to behave similarly to inactive users when lurkers' activity is regarded in terms of "comments"

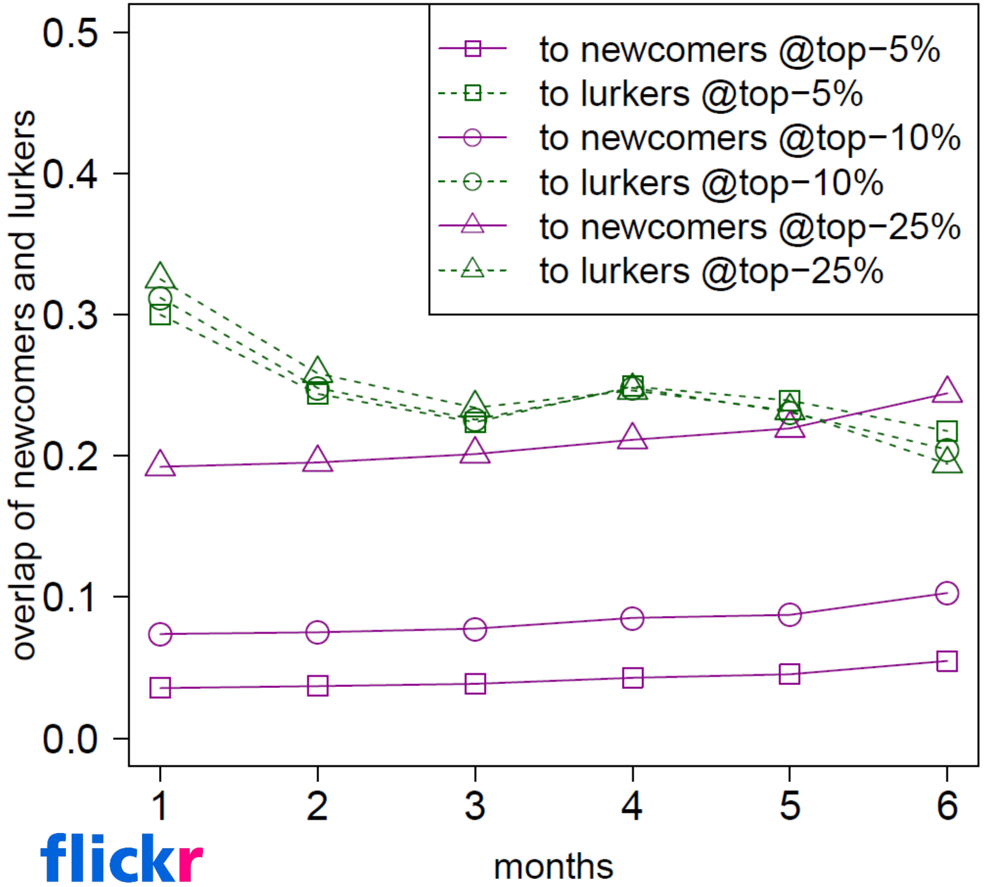
Lurkers vs. inactive users: temporal analysis



- Temporal trends of overlap ratios w.r.t.:
  - potential lurkers
  - top-5% ranked lurkers by LRin-out
  - top-25% ranked lurkers by LRin-out
- Inset: distributions of potential lurkers and zero-contributors follow close trends (at different scales)

Overlap ratios remain rather unaffected over time

Lurkers vs. newcomers (1/3)

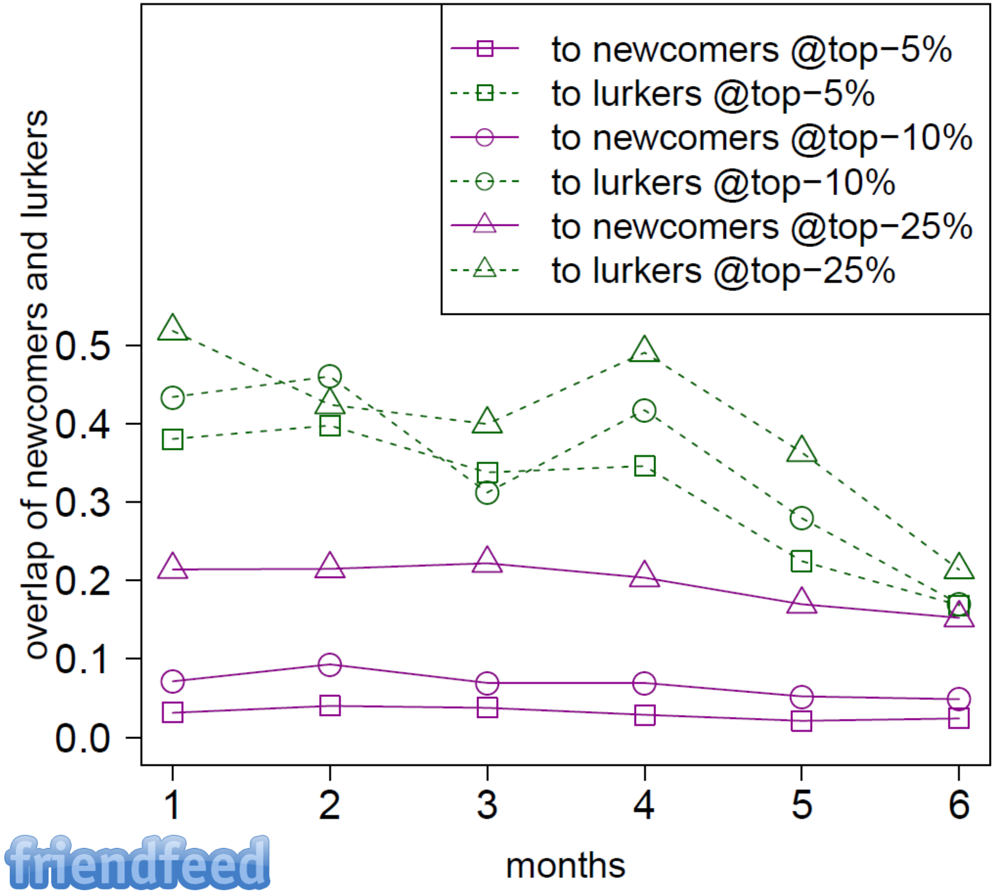


flickr

- A user is a newcomer at time *t* if she is not involved in any discussion at any time *t'<t*
- Lurkers identified at each time *t*

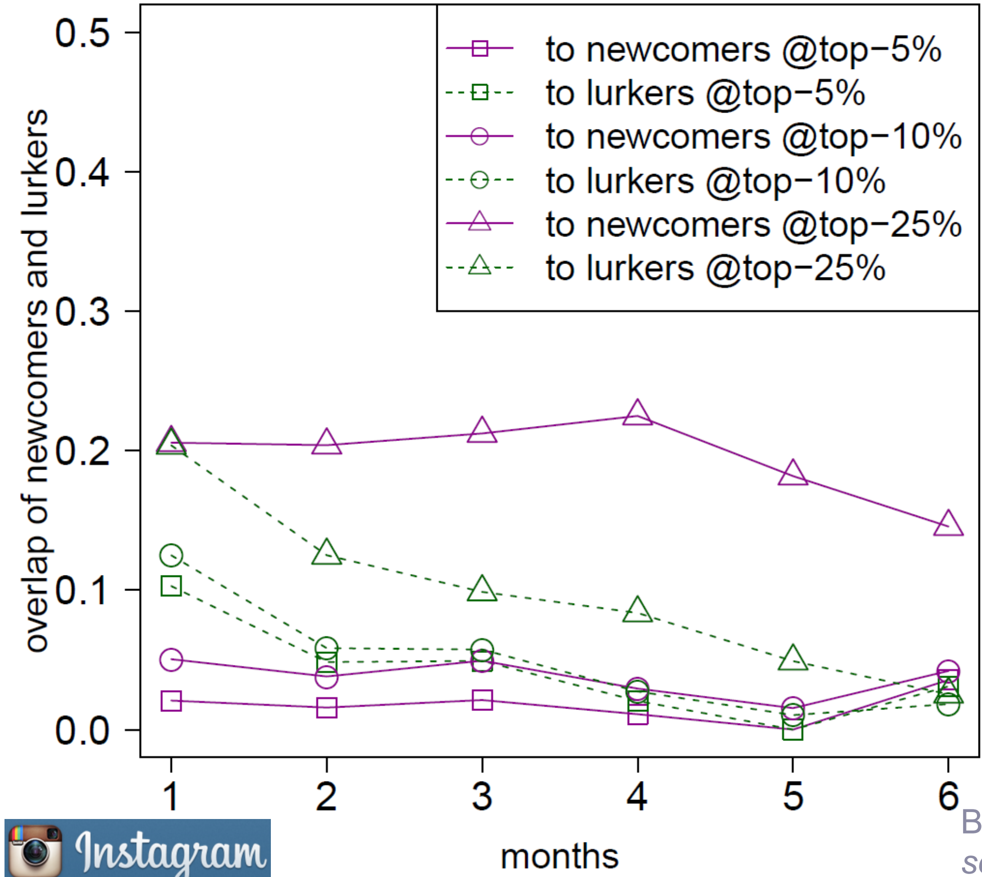
- **Favorite-markings interactions:**
  - *Lurkers matching Newcomers*: 30% down to 20% over time, regardless of the top-%
  - *Newcomers matching Lurkers*: more constant, slightly increasing. Fraction depends on the top-%.

Lurkers vs. newcomers (2/3)



- **Comment-based interactions:**
  - *Lurkers matching Newcomers*: 50% down to 20% over time
  - *Newcomers matching Lurkers*: roughly constant over time
- **Difference in matching:**
  - Inherent characteristics of an OSN
  - Type of interaction

Lurkers vs. newcomers (3/3)



Instagram

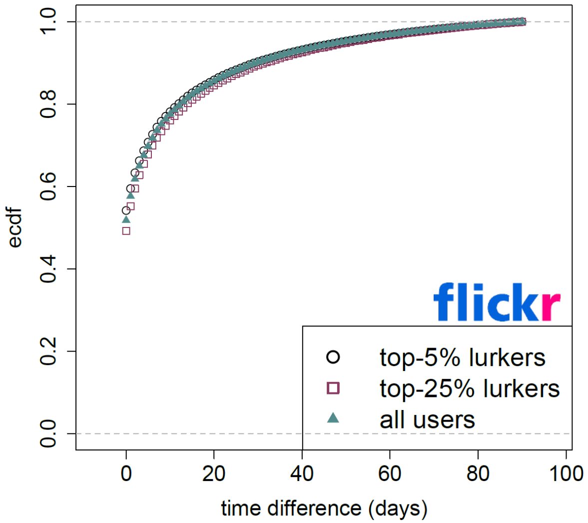
- **Comment-based interactions:**
  - *Lurkers matching Newcomers*: decreasing trend, below 10% on average

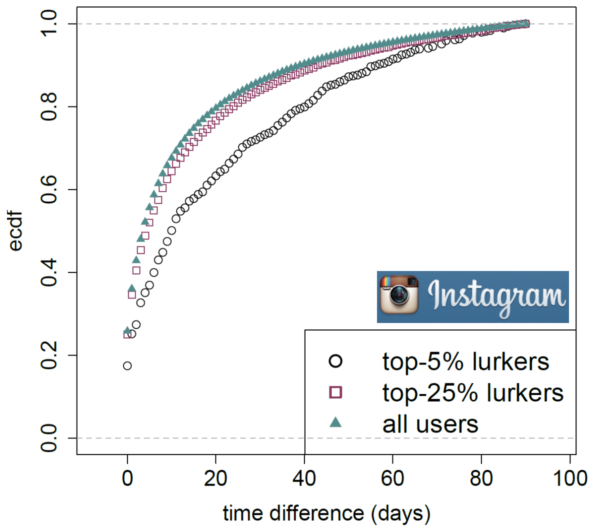
Newcomers' behavior is a form of observational learning [Bandura, 1986]

Observational learning and lurking are related to each other

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Clis, NJ, Prentice Hall.

Responsiveness





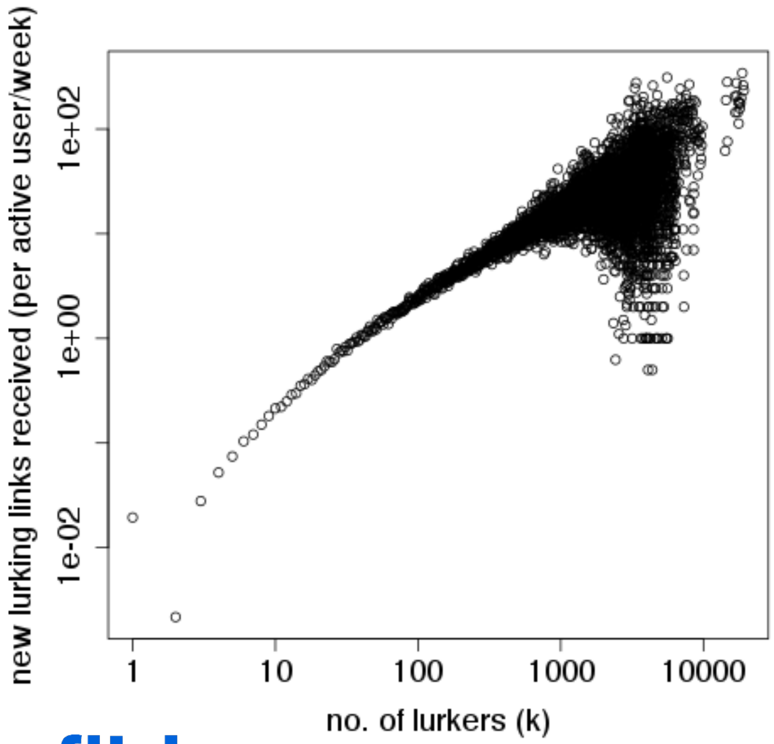
- Distribution of time differences (in days) between any two consecutive responsive actions made by a user w.r.t. a post created by her/his followees
- Timespan: 90 days
- Responses:
  - “favorites” on Flickr, “comments” on Instagram

- On Flickr:
  - about 18 days to observe 80% of responses for the top-ranked lurkers
- On Instagram:
  - about one month to observe the 80% of responses for the top-25% lurkers
  - even longer (more than 40 days) for the top-5% lurkers

Lurkers tend to react more slowly (up to 20 days more in Instagram)

Gap is reduced to a few days when taking into account a larger fraction of lurkers (top-25%)

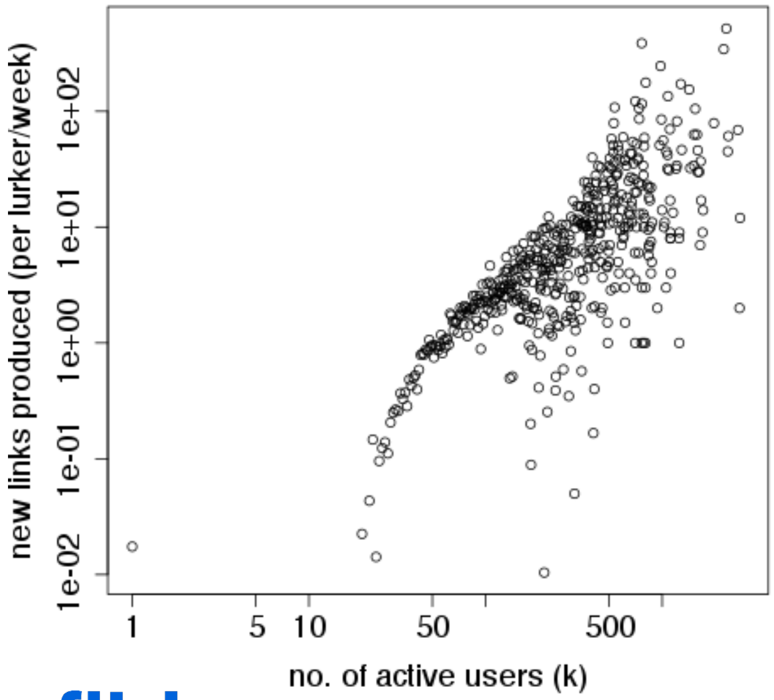
Preferential attachment (1/2)



flickr

- Studying the new connections received by active users for any *k* lurkers (averaged per user and per week)
- The number of lurkers shows a good linear correlation with the average number of new links received by active users
  - i.e., preferential attachment
- Active users receive on average one new connection per week from lurkers for every 120 connections (lurkers) that they already have

Preferential attachment (2/2)



- Studying the new connections produced by lurkers for any *k* active users (averaged per user and per week)

Are lurking connections attached preferentially to active users that already have a large number of connected lurkers?

- No preferential attachment

Lurkers that have a higher number of active users as followees are NOT more likely to create new connections to other active users

Temporal trends and clustering (1/3)

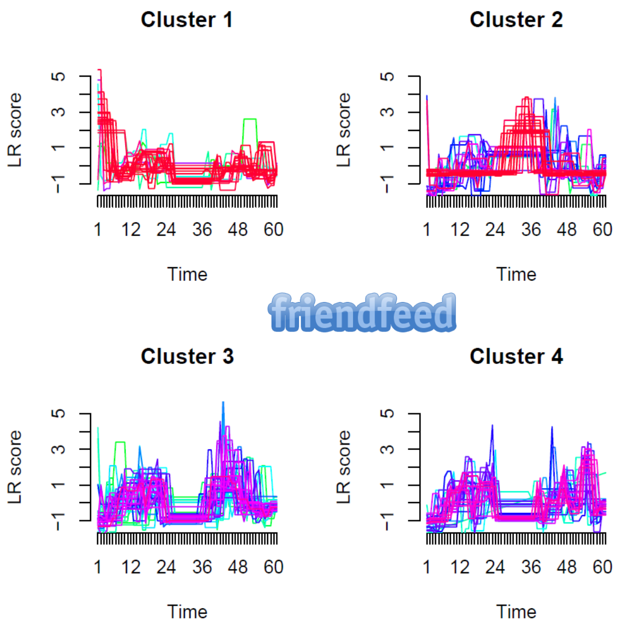
Aim: To detect structures hidden in the lurking trends that vary over time

Task: Clustering of time series representing the users' lurking profiles

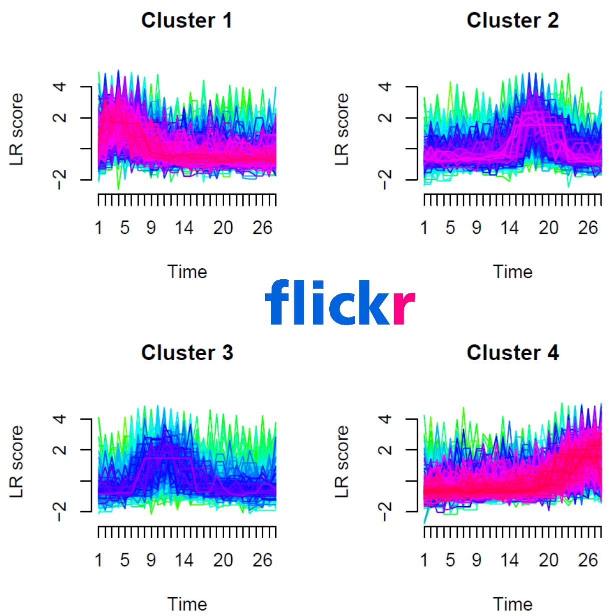
- Repeatedly applying LurkerRank to successive snapshots of a network
- Time series of the normalized LurkerRank scores for every user in the dataset
- Soft clustering over the set of time series using fuzzy c-means clustering
  - For each network, we initially selected the top-25% lurkers at time zero
  - Only users appearing in at least 50% of the subsequent snapshots

Mfuzz R-package tool: http://www.bioconductor.org/packages/release/bioc/html/Mfuzz.html.

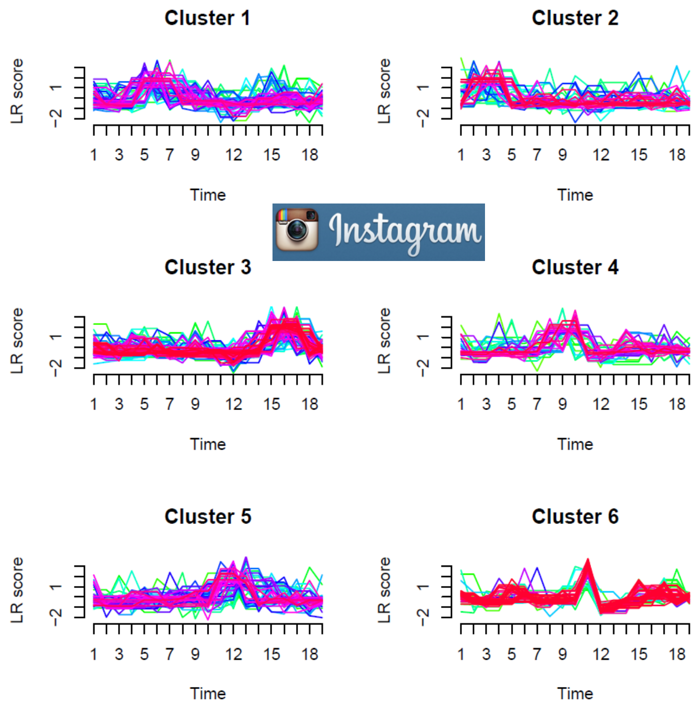
Temporal trends and clustering (2/3)



daily snapshots built on “likes + comments”



weekly snapshots built on “favorites”



monthly snapshots built on “comments”

Temporal trends and clustering (3/3)

Clearer trends, more homogeneous clusters according to least-effort interactions (e.g., “likes”/”favorites”)

More noisy clusters according to time-consuming interactions

Lurking series do not tend to group into decreasing trends

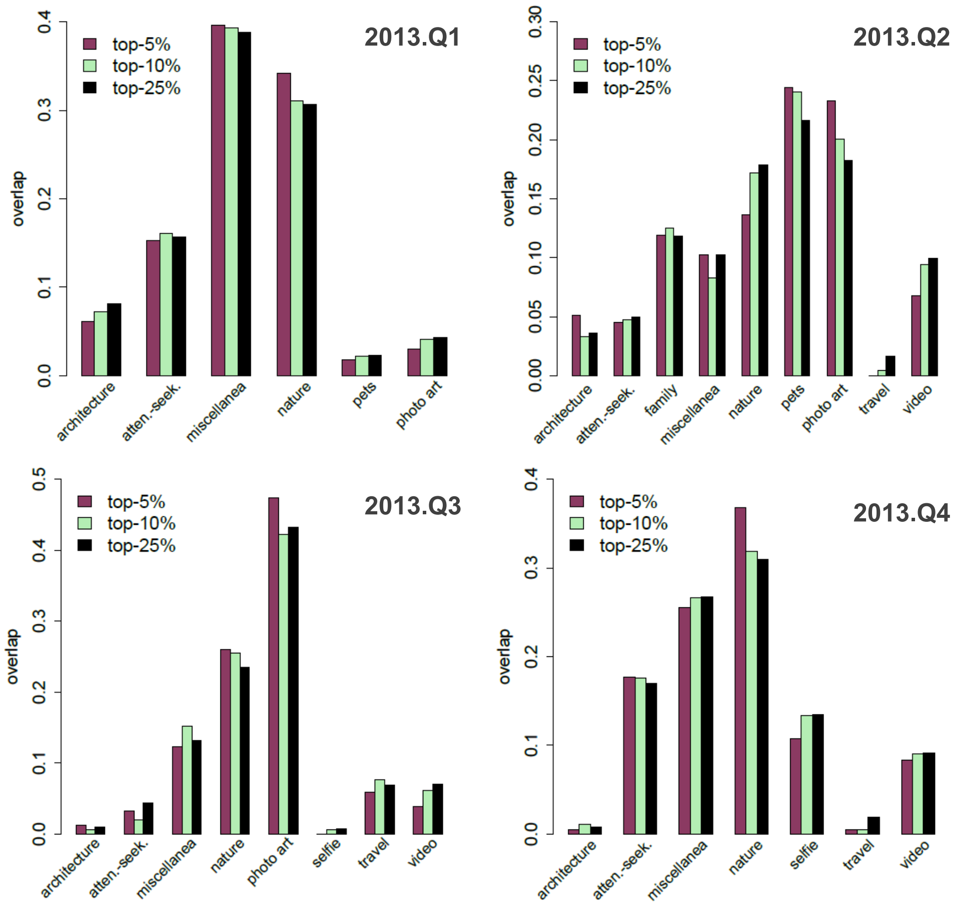
i.e., lurkers are not likely to spontaneously “de-lurk” themselves

Topical evolution: LDA-learned topics

LDA topic ids	topic-set label	main descriptors (i.e., media tags) of topic-set	subnetwork-induced size
0, 6, 10	nature	sky, sunset, whpflowerpower, whpsignsoftheseason, clouds, nature, landscape, sea, beach, flowers, water, trees, hinking, summer, fall, autumn	8,185
12, 14	architecture	whpstraightfacades, architecture, building, instaworld˙shots, streetphotography, spain, madrid, paris, france, london, sicily, design, arquitectura, youmustsee	2,884
13	fun	love, me, swag, lol, fun, like, awesome, cool, happy, food	1,314
16	pets	whppetportraits, cats, caturday, catstagram, dog, cute, pets, kitty, catsofinstagram, petsofinstagram	3,124
19	video	whpmovingphotos, whpreplacemyface, whpbigreveal, whpfilmedfromabove, instavideo, video, whpmovingportrait, movies, videogram, instagramvideo	3,062
1, 2, 7	miscellanea	whpthroughthetrees, ig˙captures, whpmyhometown, whpliquidlandscape, whpemptyspaces, whpmotherlylove, whpthanksdad, whpstraightfacades, whpmyfavoriteplace, whpfirstphotoredo, whpstrideby	16,573
8, 18	travel	worldunion, whpmyfavoriteplace, travel, world_shotz, worldcaptures, worldplaces, igworldclub	1,200
3, 4, 5, 17, 11	attention-seeking	instagood, instamood, photooftheday, pleasecomment, pleaseshoutout, teamfollowback, igers, picoftheday, instadaily, bestoftheday, webstagram, iphonesia, igdaily	5,794
9, 15	photo art	whpsilhouettes, whpselfportrait, whplookingup, whpreflectagram, , selfie, blackandwhite, whpbehindthelens, whpstilllife, silhouette, bnw, monochrome	11,882

- Statistical topic modeling: Latent Dirichlet Allocation (LDA)
- Tags occurring in less than 5 documents or in more than 75% of the documents were filtered out
- Among models with 5<*topics*<50, 20 topics model was the most *interpretable* one
- Finer-grain topics learned by LDA were aggregated in thematically-cohesive *topic-sets*

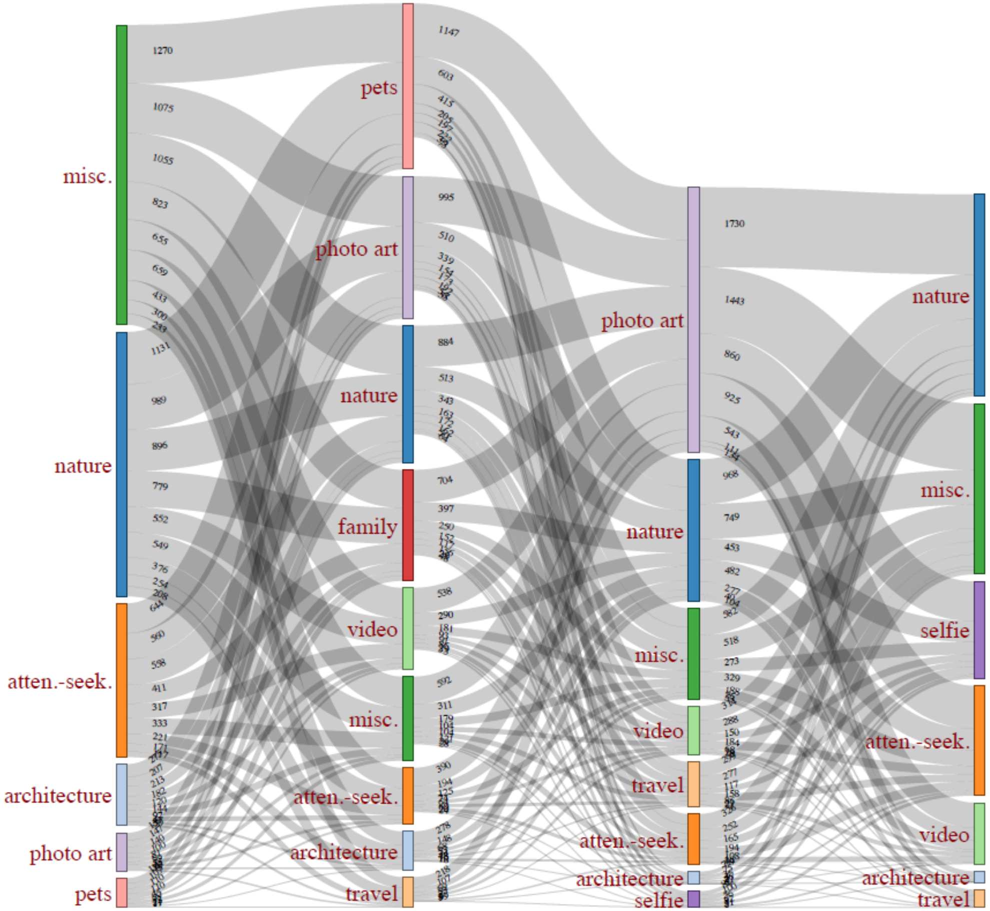
Topical evolution: Topic-specific subgraphs



- Top-ranked lurkers in snapshot graphs vs Top-ranked lurkers in topic-specific subgraphs
- **Overlap score**: intersection of top-ranked lurkers normalized over the sum of intersection values obtained over all topics
- Full graph: relatively good matching between *generic* and *topic-specific* lurkers

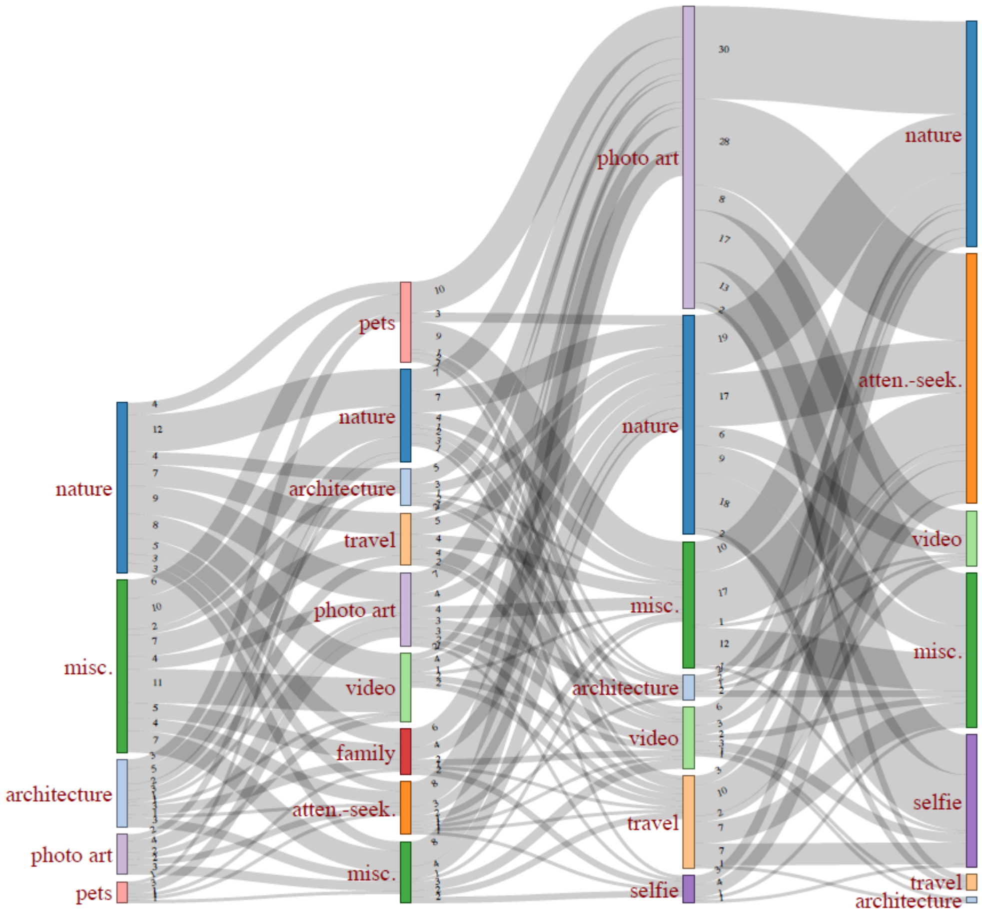
Lurkers are more likely to focus on well categorized contents

Topical evolution: Transition diagrams (1/2)



- **All user** transitions from one *topic-set* to another during the quarters of year 2013
- A core of *topic-sets* is always present over time (with varying proportions)
- Topical usage patterns **continuously change over time**

Topical evolution: Transition diagrams (2/2)



- **Top-25% ranked lurkers** transitions from one *topic-set* to another during the quarters of year 2013
- Lurkers tend to show patterns of topical interests that **do not significantly differ** from the ones of all users
- **Newcomers behavior:** higher flow in the outgoing transitions

R. Interdonato, C. Pulice, A. Tagarelli (2015)
«Got to have faith!»: The DEvOTION algorithm for delurking in social networks
In *Proc. ASONAM'15*

DELURKING-ORIENTED TARGETED INFLUENCE MAXIMIZATION

The DEvOTION algorithm for delurking in social networks

The challenge of delurking

- Lurkers are **social capital holders**:
  - they gain benefit from others' information without significantly giving back to the community
- A major goal is to **delurk** such users
  - **Delurking:** to develop a mix of strategies aimed at encouraging lurkers to return their acquired social capital, through a more active participation to the community life.

Delurking strategies have been conceptualized in social science and human-computer interaction research:
- Reward based external stimuli
- Providing encouragement information
- Improvement of the usability of the system
- Guidance from elders/master users

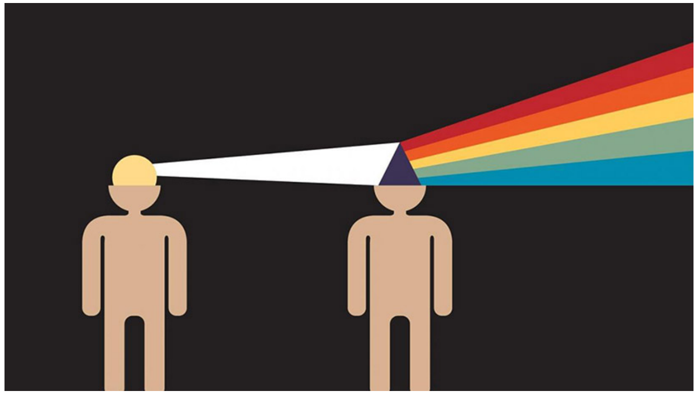
But **no computational approach** has been so far defined to turn lurkers into active participants in the social network

Information Diffusion

- **Influence diffusion process**
  - **Seed set S:** initial set of nodes selected to start the diffusion
  - **Node activations:** Nodes are activated starting from the seed nodes, in discrete steps and following certain rules
  - **Influence spread $\sigma(S)$:** *expected* number of activated nodes when the diffusion process started from the seed set S ends

Independent Cascade

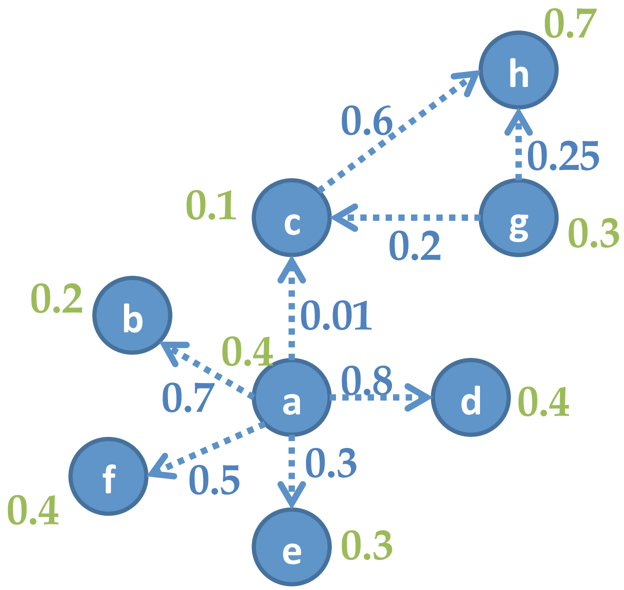
- **Contagion propagation model**
- **Sender-centric**



- **Linear Threshold:**
  - **Exposure to multiple sources is needed for a user before taking a decision**
  - **Receiver-centric**

Delurking-oriented Targeted Influence Maximization

- **Target of the diffusion process:** a set of top lurkers
- **Goal:** find a set of nodes capable of maximizing the likelihood of "activating" the target lurkers



- $LS \in [0,1]$: minimum lurking score that a node in the network must have in order to be regarded as a target node
- $\mu(S)$: *final active set* obtained using a seed set S
- $DC(\mu(S))$: *delurking capital* associated with the final active set $\mu(S)$, and defined as:

$$DC(\mu(S)) = \sum_{v \in \mu(S) \setminus S \;\wedge\; l(v) > LS} l(v)$$

Delurking-oriented Targeted Influence Maximization

- ***Objective Function:***

Given a graph $G = < V, E, b, l >$, a diffusion model on G, a budget k, and a lurking threshold LS, find a seed set $S \subseteq V$ of nodes such that, by activating them, the final active set $\mu(S) \subseteq V$ will have the maximum delurking capital:

$$S = \underset{S' \subseteq V s.t. |S'| \leq k}{argmax} DC(\mu(S'))$$

The function is defined in terms of the cumulative amount of the scores associated with the activated (target) nodes

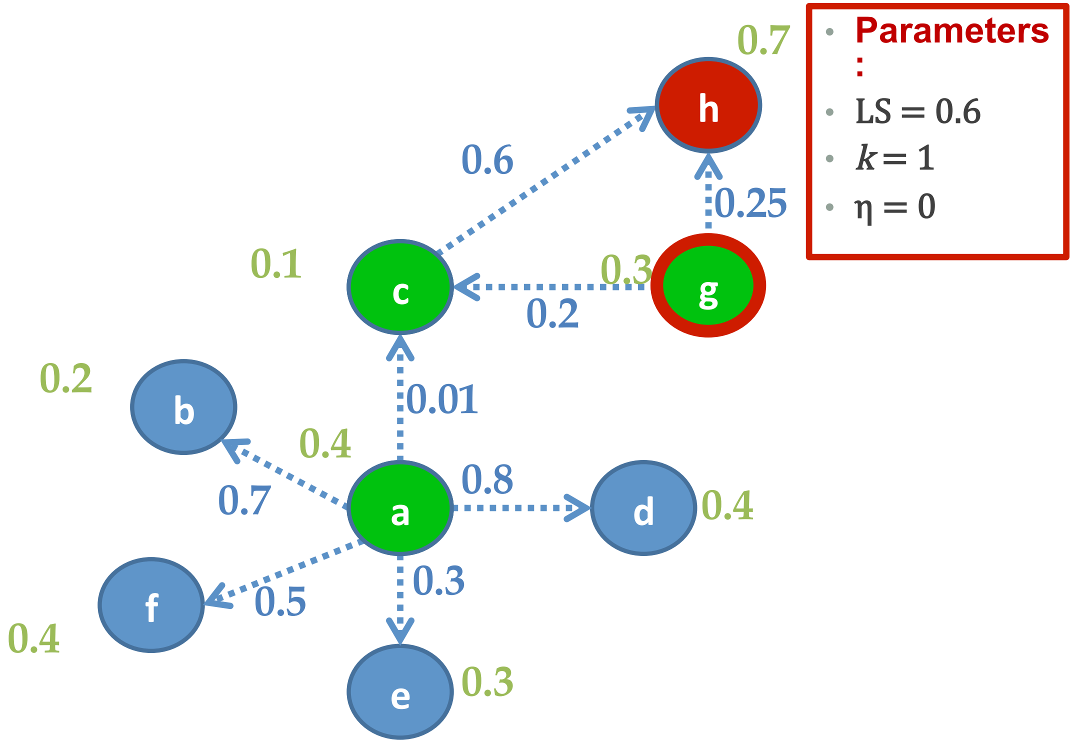
The DEvOTION algorithm

- The delurking capital function defined is **monotone and submodular** under the LT model
- **NP-Hard problem**: can be addresses using a **greedy solution**

DEvOTION (DElurking Oriented Targeted Influence maximizatiON)

- **Greedy method** designed to address the delurking-oriented targeted IM problem
- Exploits the search of shortest paths in the diffusion graph in a ***backward*** fashion
- Allows **path pruning** within a certain neighborhood

The DEvOTION algorithm

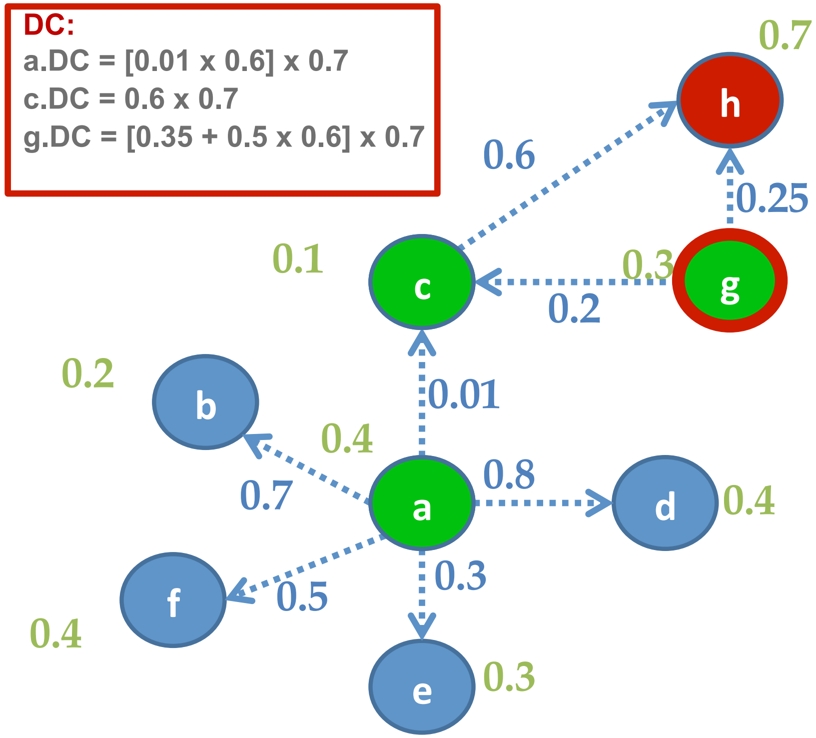


1. Compute the target set
2. Compute the set *T* of nodes that reach the target ones
3. Keep track of the best seed as the node in *T* with the highest **marginal gain** (i.e., Delurking Capital DC)

➢ Steps 2 and 3 are repeated until *k* seeds are chosen

The DEvOTION algorithm

DC:
a.DC = [0.01 x 0.6] x 0.7
c.DC = 0.6 x 0.7
g.DC = [0.35 + 0.5 x 0.6] x 0.7



- **Parameters:**
- $LS = 0.6$
- $k = 1$
- $\eta = 0$

Marginal gain computation:

- Backward procedure over all nodes in the target set
- Compute a set of paths and their probability exploring the graph backward
- At each iteration an unxeplored neighbor is added to the path in a depth-first fashion
- Paths with probability lower than η are pruned

A. Tagarelli, R. Interdonato (2013)
Ranking vicarious learners in research collaboration networks.
In *Proc. ICADL'13*

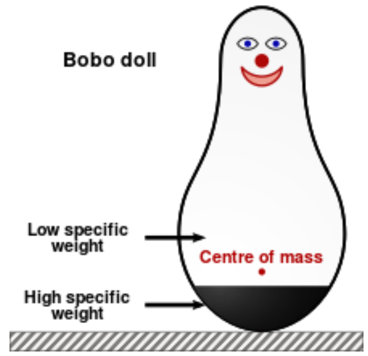
APPLICATIONS TO OTHER DOMAINS

Vicariously learning in collaboration networks

"Lurking" scenarios in information networks

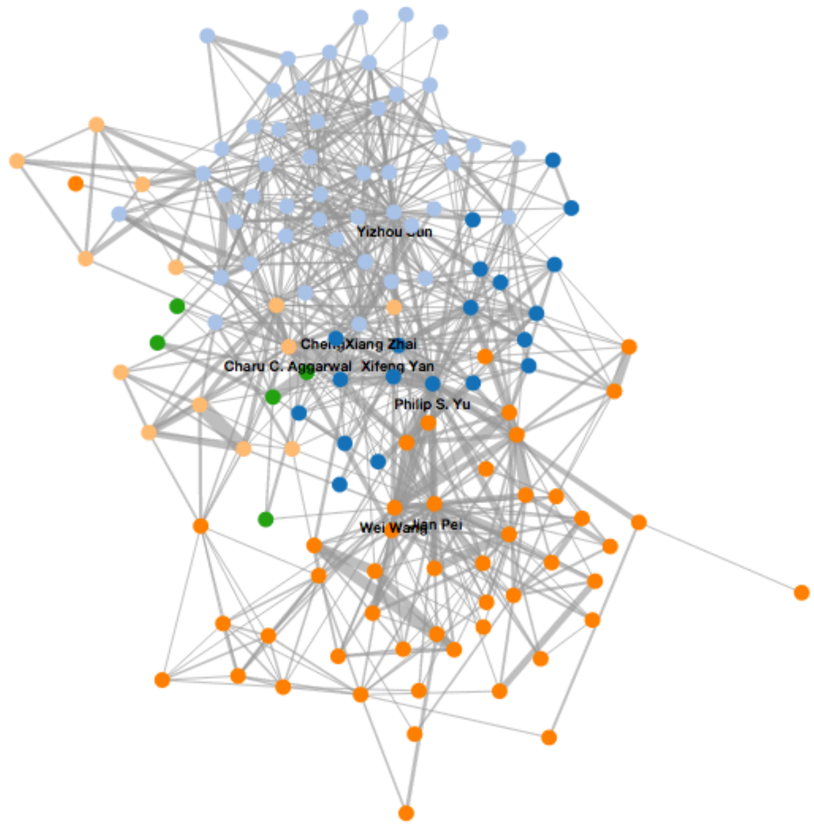
- **Leeching** (a.k.a. free loading)
  - Greedy (or even illegal) use of computer resources
  - Examples:
    - Downloading in P2P networks
    - Direct linking
    - Wi-Fi leeching

- **Vicariously learning**
  - Occurs in *observational learning* contexts:
    - learning through being given access to the learning experiences of others
  - Focus: (research) collaboration networks



Research collaboration networks (1/2)

- Formed on top of digital libraries
- Common assumption:
  - two researchers are regarded as connected to each other if they have co-authored a paper
- Typical tasks:
  - expert finding
  - community discovery
  - relation prediction



Research collaboration networks (2/2)

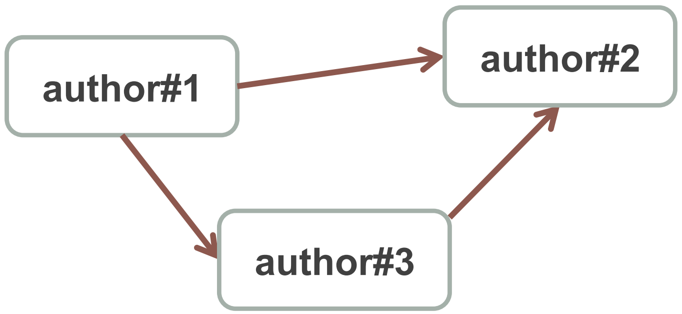
- Mining hidden **expert-apprentice** or **advisor-advisee** relationships to understand:
  - Research community formation in a particular institutional context
  - Evolution of research themes over time
  - Predicting influence of a research study on a community
  - How to foster several experts on specific topics

- Current trend: expert-oriented investigation of co-authorships
- However, many members in a RCN are more likely to be **apprentice**:
  - in the initial stage of a researcher lifetime (early career)
  - w.r.t. all topics that at a particular time do not represent a researcher's main research interests

Vicariously learning

- Learning through being given access to the learning experiences of others
- In a publication context:
  - **people who marginally contribute to the research activity?**

- Vicariously-learning-oriented RCN:
  - Directed weighted graph model
    - Basic model for edge orientation: comparison of relative amount of publications



- **pubs(author#1) > pubs(author#2)**
- **pubs(author#1) > pubs(author#3)**
- **pubs(author#3) > pubs(author#2)**

Vicariously learning oriented RCN

- Interactions among authors expressed through edge weights based on:
  - **number of co-authorships**
    - to express the strength of collaboration
  - **number of *advisees* for each *advisor***
    - an advisor tends to divide her attention over all incoming stimuli that come from her advisees

$$w_t(i,j) = coPubs(i,j,t)\left(1 - \frac{\sum_{k \in advisees(i,t) \setminus \{j\}} coPubs(i,k,t)}{\sum_{k \in advisees(i,t)} coPubs(i,k,t)}\right)$$

$coPubs(i,j,t)$: number of papers coauthored by authors i and j at time t

$advisees(i,t)$: number of advisees (i.e., out-neighbors) of author i at time t

Vicarious learner ranking (1/2)

- **VLRank** algorithm:
  - Adaptation of LurkerRank such that
    - The lurking-oriented graph model is replaced with the VL-oriented weighted graph model
    - Advisees act as lurkers w.r.t. advisors (i.e., active users)
- Evaluation on the DBLP dataset
  - Static analysis
    - Full dataset (about 1.2M nodes, 4.7 links)
  - Evolution of vicarious learners
    - 3-year snapshots

time interval	*# nodes*	*# links*	*avg in-degree*	*avg path length*	*# source nodes # sink nodes*	*avg in/out-degree **	*clustering coefficient*
-2013	1,191,619	4,712,489	3.95	7.50	54,647; 533,101	1.75	0.18
2004-2006	341,282	957,922	2.81	7.61	32,511; 139,016	1.33	0.44
2007-2009	469,345	1,412,556	3.01	7.16	40,021; 188,166	1.41	0.32
2010-2013	582,206	1,926,184	3.31	6.82	45,916; 227,990	1.50	0.28

Vicarious learner ranking (2/2)

- Issue: lack of ground-truth
- **ArnetMiner** based ranking
  - Expert's activity score: used to rank the researchers based on the cumulated weighted impact factor of one's papers published in the last years

$$r_i^* = \frac{1 + \sum_{j \in H^+(i)} AS(j)}{1 + \sum_{j \in H^-(i)} AS(j)}$$

  - $H^+(i)$ is the set of authors with h-index greater than i
  - $H^-(i)$ is the set of authors with h-index lower than i
  - $AS(j)$ is the activity score of author j provided by ArnetMiner

Quantitative analysis: Kendall correlation

	-2013	*2004-06*	*2007-09*	*2010-13*
VLRank vs. InOut	.153	.249	.259	.256
VLRank vs. DDRank	.284	.283	.295	.298
PageRank vs. InOut	-.097	.182	.177	.177
PageRank vs. DDRank	.133	.246	.246	.255
VLRank vs. AMRank	.115	–	–	.148
PageRank vs. AMRank	.043	–	–	.083
VLRank vs. PageRank	.422	.424	.410	.407

VLRank always obtains positive Kendall scores

VLRank always achieved higher correlation with InOut, DDRank and AMRank than PageRank, with gains up to 21.7% for InOut, 11.8% for DDRank, and 6.5% for AMRank.

Quantitative analysis: Bpref

	-2013			*2004-06*			*2007-09*			*2010-13*		
	$p = 10$	$p = 25$	$p = 50$	$p = 10$	$p = 25$	$p = 50$	$p = 10$	$p = 25$	$p = 50$	$p = 10$	$p = 25$	$p = 50$
VLRank vs. InOut	.336	.584	.664	.362	.561	.702	.409	.583	.713	.415	.580	.714
VLRank vs. DDRank	.687	.784	.744	.605	.701	.706	.644	.726	.717	.667	.730	.720
PageRank vs. InOut	.099	.328	.467	.204	.449	.666	.211	.460	.670	.219	.469	.676
PageRank vs. DDRank	.481	.626	.592	.528	.639	.648	.544	.658	.654	.580	.671	.668
VLRank vs. AMRank	.191	.448	.645	–	–	–	–	–	–	.264	.508	.663
PageRank vs. AMRank	.131	.338	.573	–	–	–	–	–	–	.166	.385	.603
VLRank vs. PageRank	.804	.853	.857	.620	.726	.815	.656	.754	.834	.650	.735	.817

VLRank always outperforms PageRank also in terms of Bpref

Bpref scores generally increase with the *p*% of relevant candidates

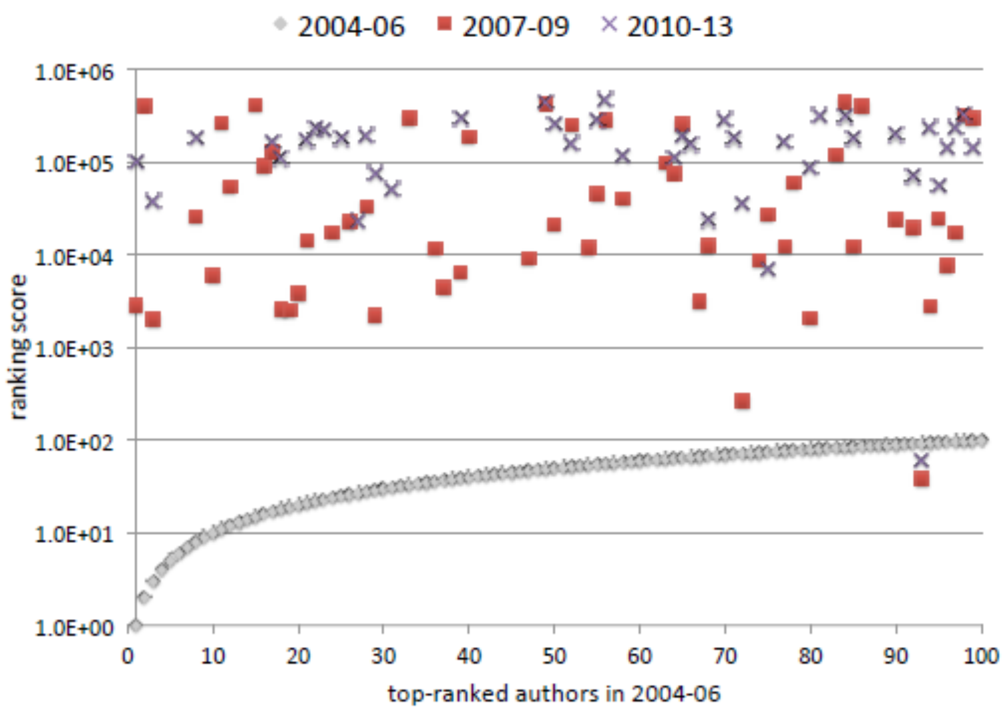
Qualitative analysis

- Comparison between the top-100 ranked lists produced by VLRank and PageRank on the whole DBLP network (-2013)

- VLRank detected and assigned highest scores to authors whose status can be tagged as vicarious learner with a certain objectivity
  - e.g., short career always within a research team, long career but with many co-authors, etc.

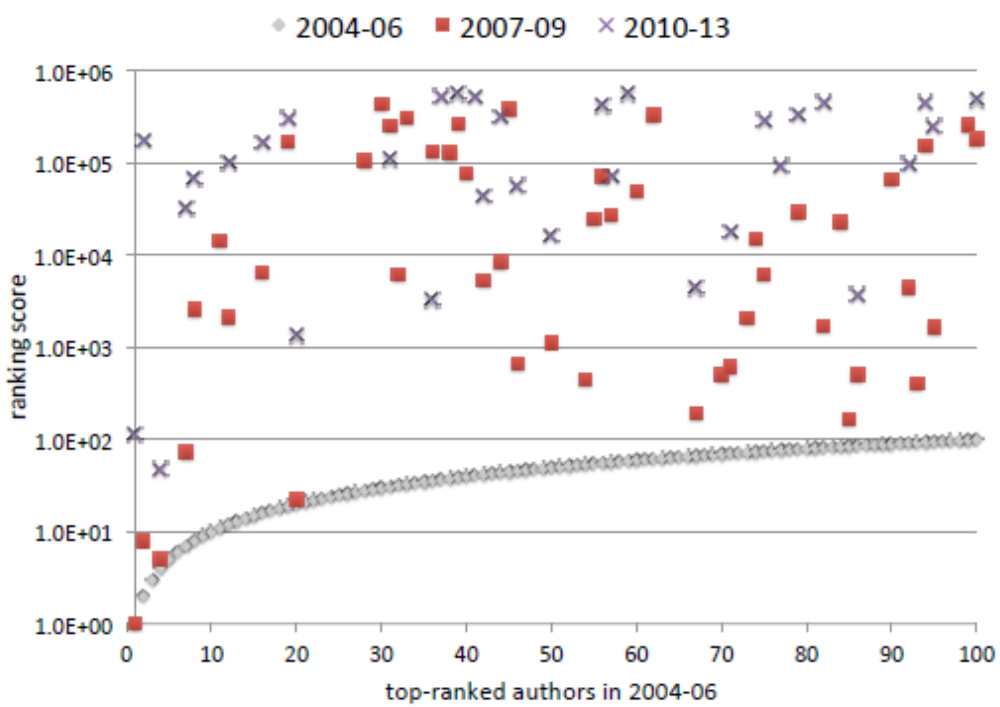
- Several authors in the PageRank top-ranked list should be considered as team leaders, or at least active contributors
  - e.g., many publications with few co-authors

Temporal evolution of Vicarious Learners: VLRank



A large number of top-100 authors by VLRank in 2004-06 were also present in the subsequent periods but with much lower ranks

Temporal evolution of Vicarious Learners: PageRank



PageRank failed to effectively capture the temporal evolution of vicarious learners

A. Tagarelli, R. Interdonato (2014)
Lurking in Social Networks: Topology-based Analysis and Ranking Methods.
Soc. Netw. Analys. Mining (SNAM)

APPLICATIONS TO OTHER DOMAINS

Lurking in Social Trust contexts

Social trust and lurking (1/2)

- Measuring trust behaviors has long been an important topic in psychology and social science
- Computer science perspective: trust based on **active behaviors** shown by the users in an online community

Trustworthy users:
influential users, verified profiles

Untrustworthy users:
spammers, trolls, fake profiles

What about lurkers?

Social trust and lurking (2/2)

- (Active) users tend to avoid wasting their time with people who show null or slow responsiveness – like lurkers do

Should lurkers be treated as *untrustworthy* users?

- Preliminary insight into understanding relations between lurkers and trustworthy/untrustworthy users:
  - Comparison between LurkerRank and *TrustRank* [Gyongyi, et al., 2004]

Goal: To improve the trustworthiness of the lurkers to be detected

Gyongyi, Z., Garcia-Molina, H., Pedersen, J. O. (2004) Combating Web Spam with TrustRank. In *Proc. VLDB*.

TrustRank-biased LurkerRank (1/2)

- Definition of **TrustRank-biased LurkerRank** methods
- TrustRank in a nutshell
  - A biased PageRank in which the teleportation set corresponds to the “good part" of an a priori selected *seed set*
  - The seed set is a relatively small subset of nodes in the graph, each of which is labeled as either trustworthy or untrustworthy by some *oracle* function

- **Issue in OSNs**: inferring trust from user interactions
  - Number of received likes, favorites, or comments as implicit trust statements

- Assumption: the higher the number of users that indicate trust in a user, the more likely is the trustworthiness of that user
- **Trust-Entropy-based oracle** function: $H(i) = -\frac{1}{\log |V_i|} \sum_{j \in V_i} p_j \log p_j$ $\quad p_j = \frac{ET(j,i)}{\sum_{k \in V_i} ET(k,i)}$

TrustRank-biased LurkerRank (2/2)

- A user i is regarded as “good" if the corresponding $H(i)$ belongs to the 3rd quartile of the distribution of H values over all users
- Note that: if user i likes a post by j, then
  - edge j -> i is created in the LurkerRank graph
  - edge i -> j is created in the TrustRank graph

flickr	LR vs. TrustRank	trust-LR vs. TrustRank	trust-LR vs. LR
LRin	.393	.436	.639
LRout	**.562**	.556	**.980**
LRin-out	.441	.640	.688
ac-LRin	.445	.434	.728
ac-LRout	.561	.559	.945
ac-LRin-out	.402	**.724**	.498

Kendall correlation

- All LurkerRank methods show positive correlation with TrustRank
- Higher correlation when using TrustRank-biased LR
- TrustRank-biased LR have still strong correlation with their respective LR methods

Trust-oriented bias in LurkerRank would not significantly decrease lurker ranking effectiveness while also accounting for the user trustworthiness

CONCLUSION AND FUTURE WORK

What we have done …

Modeling lurking behaviors

Topology-driven definition of lurking

In-, Out-, and InOut-neighbors driven ranking methods

Vicariously Learning on RCNs

- VLRank methods

Lurking and Social Trust

- Trust-biased LurkerRank methods

Delurking via Targeted Influence Maximization

- The DEvOTION method

Evaluation on Twitter, FriendFeed, Flickr, Google+, and Instagram

- Reciprocity, preferential attachment
- Delurking-oriented randomization model
- Percolation/resilience analysis

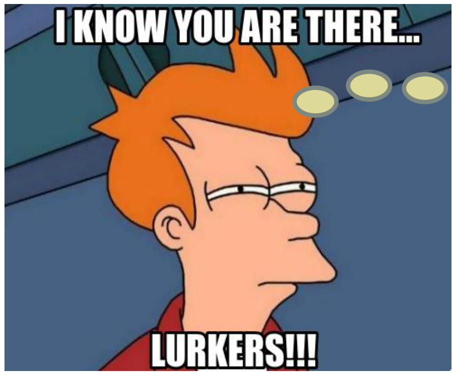
Lurking over time

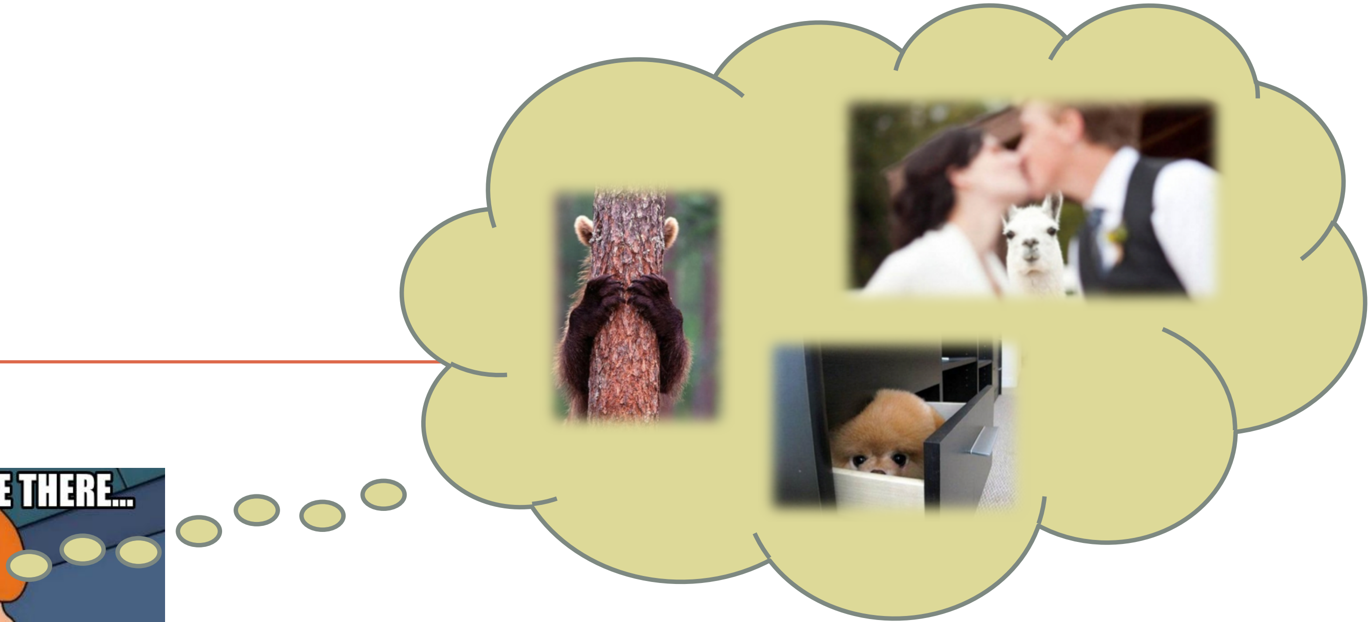
- Lurkers vs. inactive users , and newcomers
- Responsiveness
- Preferential attachment
- Temporal trends and clustering
- Topic evolution

…and what we would like to do

- Extensions of the lurking concept
  - Context-biased lurking
  - Boundary-spanning lurking
  - Multi-layer networks
- Integration with
  - Community detection algorithms
  - Trust/Distrust ranking algorithms
- *… any other idea is welcomed!*

THANKS





Contact us at:
http://uweb.dimes.unical.it/tagarelli/