Social Interaction in the Flickr Social Network

Joydeep Chandra

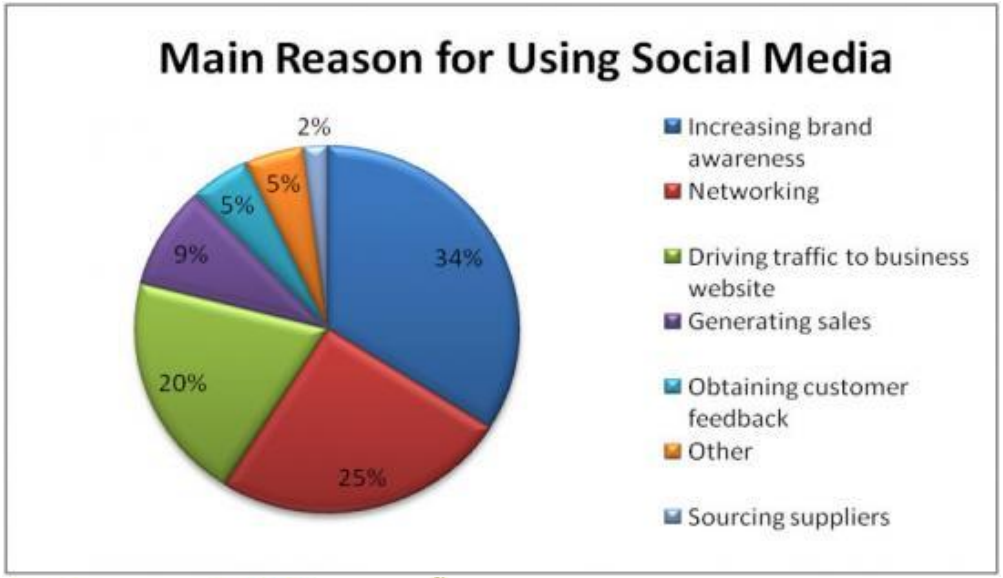
Collaborators: Karthik Gopalakrishnan and Arun Pandey

Department of Computer Science & Engg.,

Indian Institute of Technology, Patna

Social Media dynamics and applications

- Complex Social Media Dynamics through
  - Patterns of social connectivity of users
  - Information exchanged via social ties
- Possible Dynamics
  - Gain connectivity
  - Spreading and Virality
  - Influence



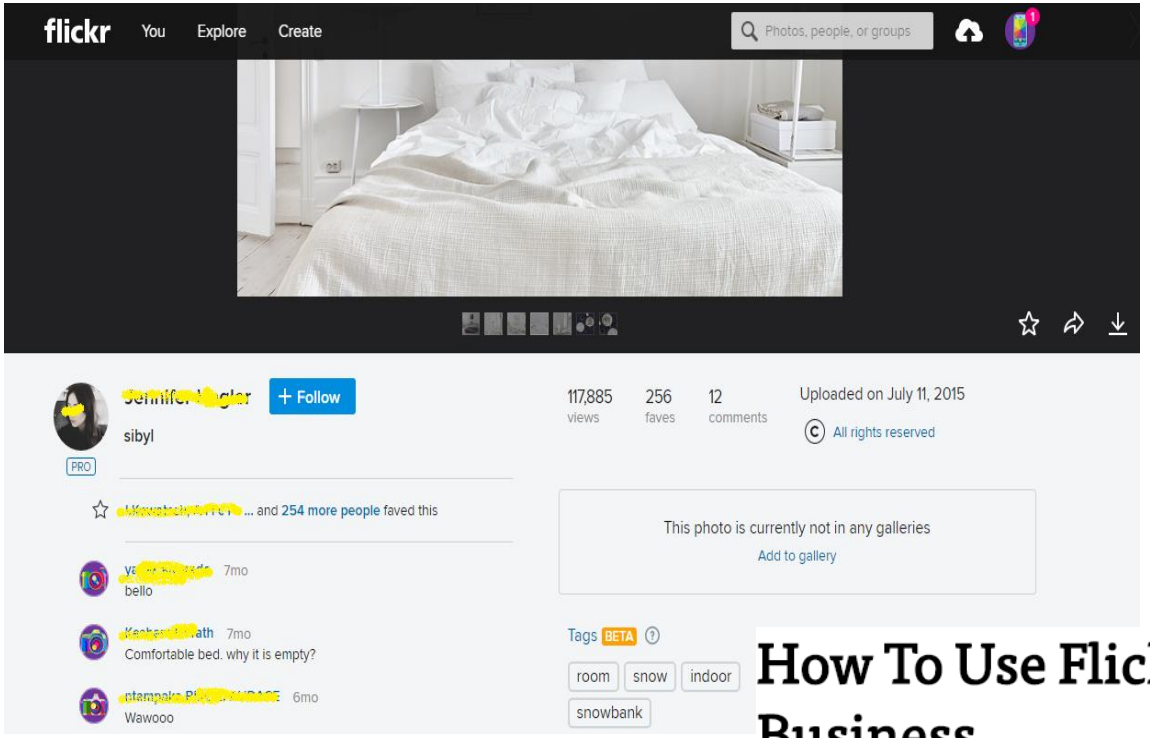
http://www.visibleinteractive.com/

http://rethinkmarketing.com.au/

Motivation

- Excellent opportunity to
  - Analyze interaction dynamics in large-scale social systems
- To characterize human behavior
  - Required- Large-scale fine-grained data traces of human activity

Flickr: The photo sharing network



- Major Features
  - Focus on images
    - Not images as text adjuncts
  - Public Images
    - Opens images to strangers
  - Social Tagging
    - Folksonomies
  - Geo-tagging of photos

How To Use Flickr To Improve Your Small Business

Though not a place to sell products, Flickr is a visual storytelling tool that can help you develop your brand.

Here are five ways that you can use Flickr to help improve your business:

Tell your company's story, through pictures

Every brand has a story. Flickr presents an opportunity to share it with your customers, and more importantly, you can give them a chance to be a part of it. Flickr is particularly strong at telling stories visually. Use compelling photos that represent the identity of your company, where its roots are and where it's headed.

And be sure to get your customers directly involved with, and be able to affect, that storyline. Have them submit their pictures interacting with your brand, and encourage them to talk to each other—and you—about it. And remember, in order for people to talk to your brand, you have to make things personal. Your voice must be real, and human.

Always avoid the hard sell

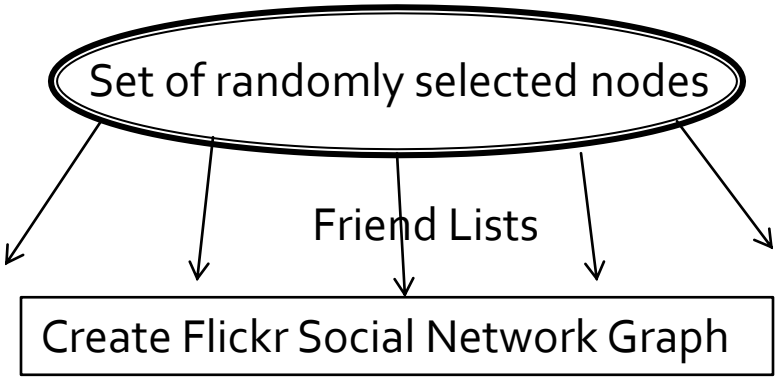
https://www.americanexpress.com

Information Spread through Flickr

- Key questions addressed:
  - How wide is the spread? [Cha et al. www09]
  - Who are the agents of spread? [Lerman ICWSM 07, Karthik et al. COMSNETS 16]
  - How quick is the spread? [Cha et al. Comp Netw. 12]

Data Collection

- Needs to collect
  - The evolving state of the social network
  - Information propagation from one user to another



Dataset Summary

- Time Period: 104 days
- #Links: 33,140,018
- #Users: 2,570,535
- #Photos: 11,195,144
- #Favorites: 34,734,221

For each user in user-list, crawl the social network graph once per day

1. Favorite photos of the users
2. Timestamp of the favorite marking
3. State of the social network at time the favorite marking took place

M. Cha, A. Mislove and K.P. Gummadi. "A Measurement-driven Analysis of Information Propagation in the Flickr Social Network". In the proceedings of the International Conference on the World Wide Web (WWW 2009).

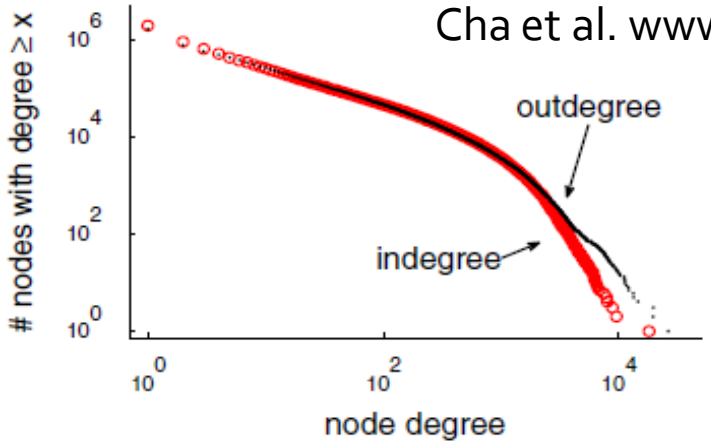
Data Collection Methodology (Contd.)

- Methodology does not take into account
  - deleted favorite markings
- Dataset is not extremely rich with content
  - does not contain information on who commented on a specific photo

Key Findings: Social Network Topology

- Indegree & Outdegree correlated
  - Pearson correlation coefficient– 0.76
- Reciprocity of links
  - 68% links are bidirectional
- # of friends (outdegree) - Heavy tailed
  - 55% users have 1 friend, 90% have less than 10
  - Maximum outdegree = 26,342. Average = 14

Cha et al. www'09

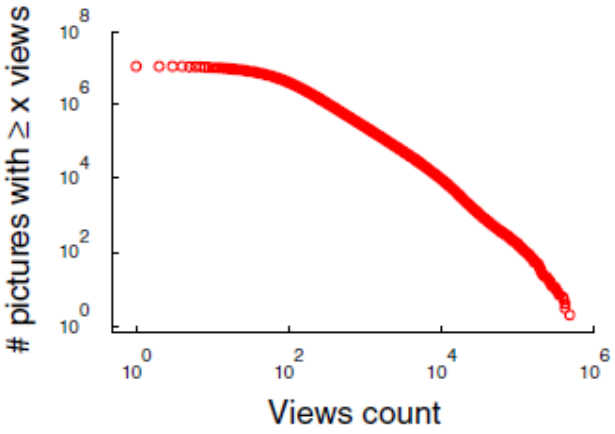


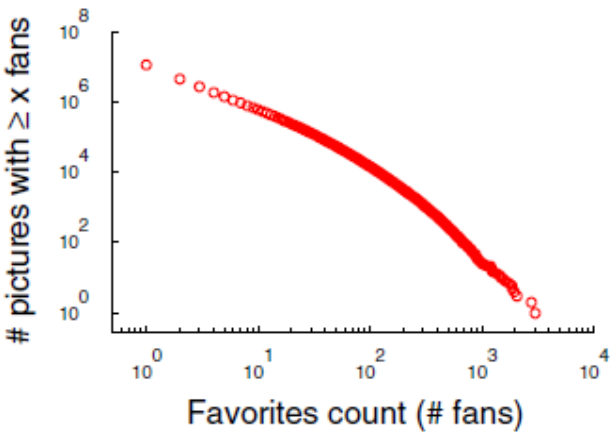
- Maximum path length =27
- Average Path length = 5.67
- Clustering Coefficient
  - High degree nodes 0.05 ~ 0.1
  - Low Degree 0.2 ~ 0.4

Implications

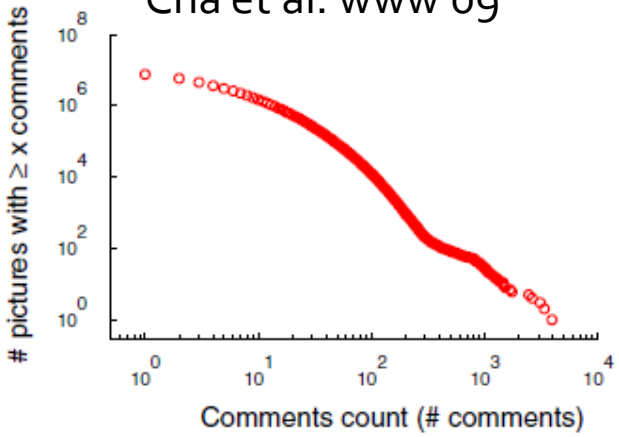
- Exhibits small world properties
- Possibilities of widespread dissemination of popular information throughout the network

Key Findings: Picture Popularity





Cha et al. www'09



Findings:

- View counts are much higher than favs and comments
- High correlation of #favs and #comments
- Low correlation of #views and #favs

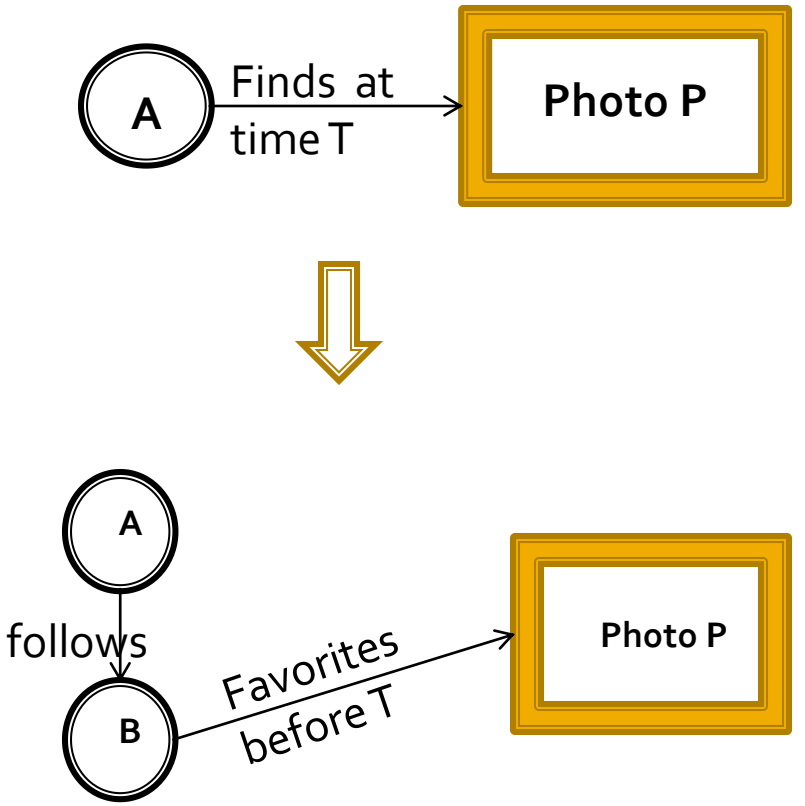
Major implication:

- Users find most pictures uninteresting

Information Propagation through social links

- Dissemination mechanisms
  - Featuring
    - Front & Explore pages
  - Searching
    - Titles, tags and descriptions
  - Links
    - Sets and pools
  - External Links
    - External websites, blogs and emails
  - Social Network
    - Word-of-mouth propagation (**social cascades**)
    - Difficult to find

Major Focus



Key Findings

Cha et al. www'09

Popularity (# Fans)	Total pictures	Total fans	Social cascades				Cascades from uploaders			
			# Photos	Perc.	# Fans	Perc.	# Photos	Perc.	# Fans	Perc.
1-5	2,704,806	4,328,609	1,517,550	56%	2,197,522	51%	1,487,266	55%	2,111,551	49%
6-100	346,870	5,121,820	329,029	95%	2,834,704	55%	306,287	88%	2,307,155	45%
101-300	3,502	499,870	3,502	100%	273,596	55%	3,337	95%	171,085	34%
301-500	154	54,773	154	100%	27,849	51%	147	95%	15,251	28%
501-	29	20,113	29	100%	8,686	43%	28	97%	4,017	20%
Total	3,055,361	10,025,185	1,850,264	61%	5,342,357	53%	1,797,065	59%	4,609,059	46%

- High content locality
  - Most pictures are favorited locally (within 1 hop)
- Social cascades key to propagation
  - Both popular as well as unpopular photos (43~55%)
- For less popular photos
  - Uploaders play an important role in cascade
- For popular photos
  - Nodes beyond 1-hop vicinity play crucial role in cascade

Significance of the observations

- Social network plays an important role in Flickr spreading
  - More than 50% of favoriting is done through social cascades
- Most fans are within a few hops from the uploader
  - High content locality

High content locality: Possible explanations

- Decreasing reproduction rate, R, at each hop
  - Number of infected people when $R<1 \approx N/(1-R)$ [Watts & Peretti '07]
- Homophily
  - Users who like each others pictures tends to become friends
  - People who are friends tends to like each others pictures

The agents of spread

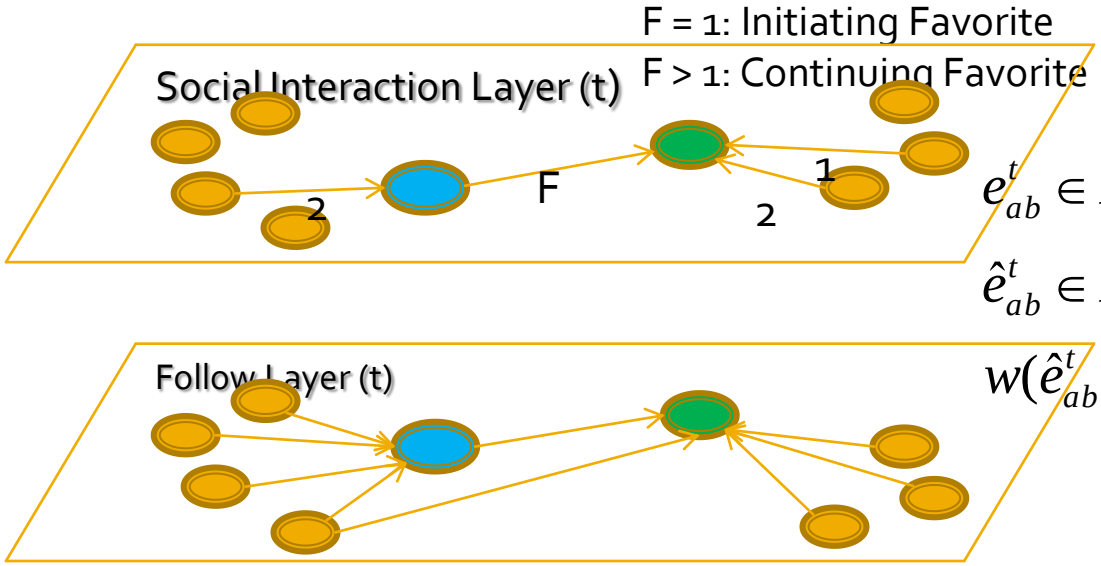
- Examine the growth of the interaction network in Flickr
  - Preferential attachment, reciprocation
- Examine link formation after an interaction occurs
  - Multiplex triangle closure

Karthik Gopalakrishnan, Arun Pandey, Joydeep Chandra, Social interaction in the Flickr social network. COMSNETS 2016

Flickr as a Multiplex Network

- Multiplex networks: Networks with fixed set of nodes and each node pair have different type of relations
  - Layers of network represent different relations
- Flickr as a two-layer temporal multiplex network:

$$G^t = \left(G_f^t, G_{si}^t\right) \qquad G_f^t = \left(V, E_f^t\right) \qquad G_{si}^t = \left(V, E_{si}^t, w\right)$$



$$e_{ab}^t \in E_f^t \Rightarrow a \text{ follows } b$$

$$\hat{e}_{ab}^t \in E_{si}^t \Rightarrow a \text{ has favorited photo of } b$$

$$w(\hat{e}_{ab}^t) : \text{No. of photos of } b \text{ favorited by } a$$

Flickr as a Multiplex Network (Contd.)

- Since the data collection methodology does not take into account the possible deletion of a favorite marking:

$$E_{si}^{t_0} \subseteq E_{si}^{t_1} \subseteq E_{si}^{t_2} \ldots \subseteq E_{si}^{t_n} \text{ for } t_0 \leq t_1 \leq t_2 \ldots t_n$$

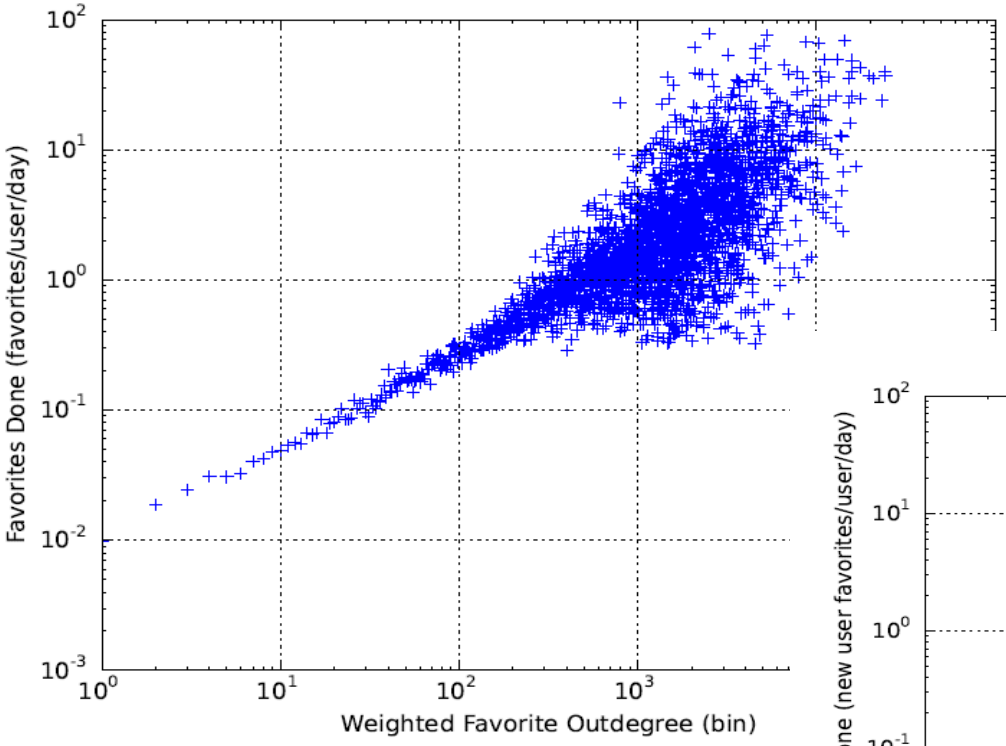
- But the same cannot be said for the follow edge sets since edge deletion is taken into account

Goals

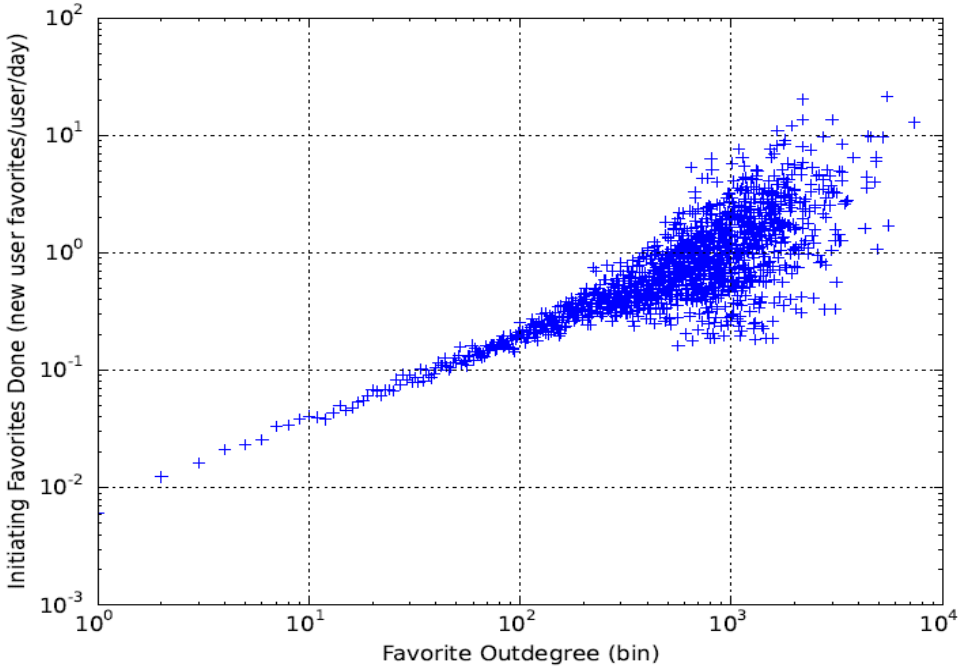
- Observe
  - Link creation properties in both layers
    - Of nodes with different degrees
  - Reciprocity

- As per our formulation, construct the graphs as on first and last days of the crawl period
  - **Favoriting of photos**: Weighted favorite out/in-degree vs. no. of favorites created/received per day
  - **Favoriting of users**: Favorite out/in-degree vs. no. of initiating favorites created/received per day

Favoriting of Photos/Users: Preferential Creation

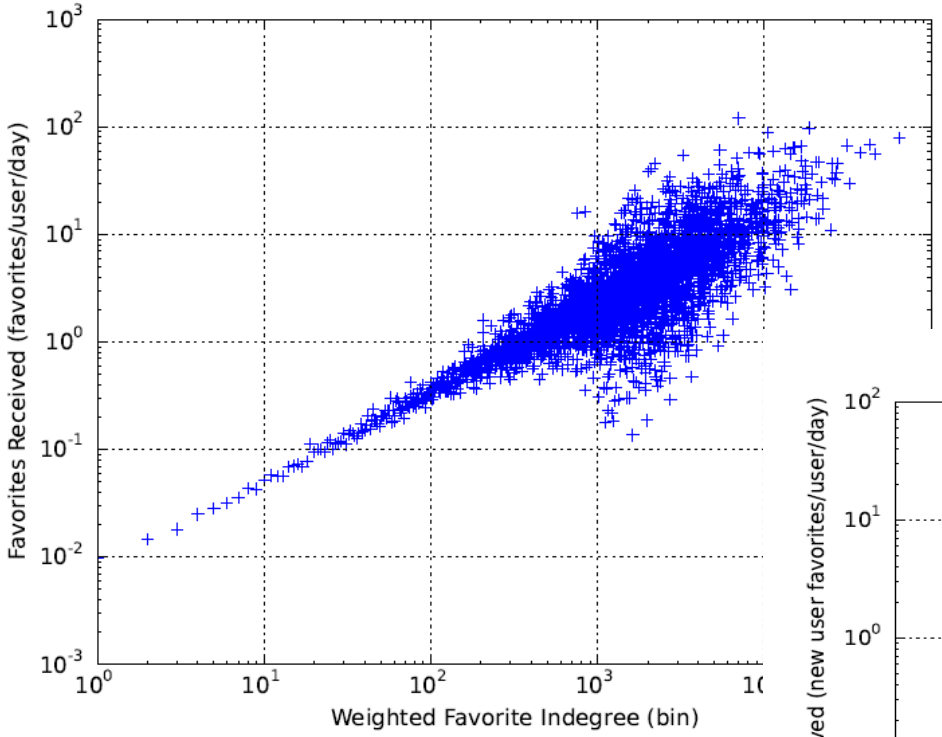


Users who favorite a lot of photos continue to favorite a lot of photos

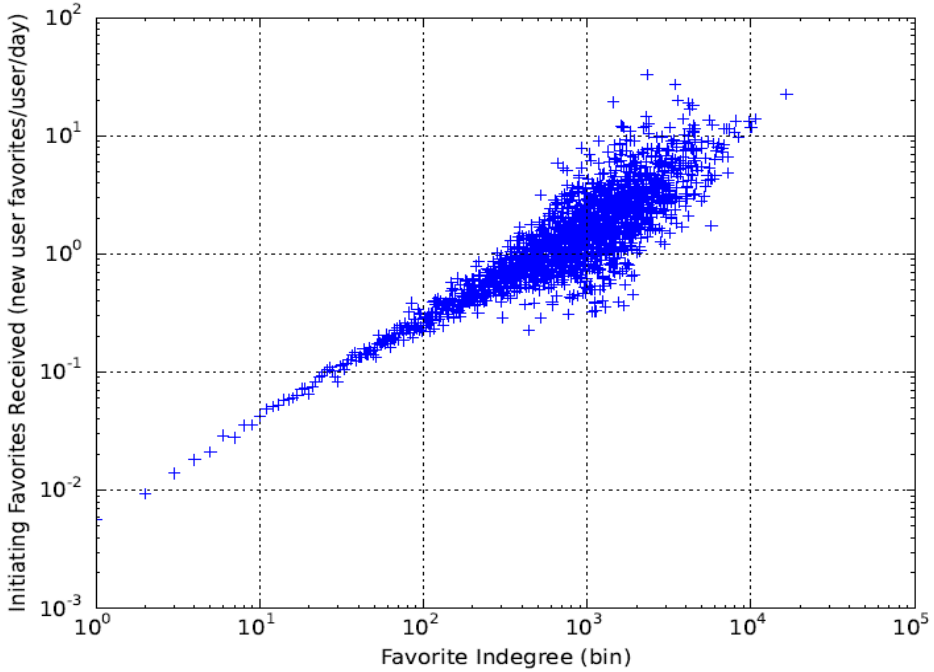


Users who explore the profiles and favorite the photos of many users continue to be exploratory and favorite the photos of many more users

Favoriting of Photos/Users: Preferential Reception



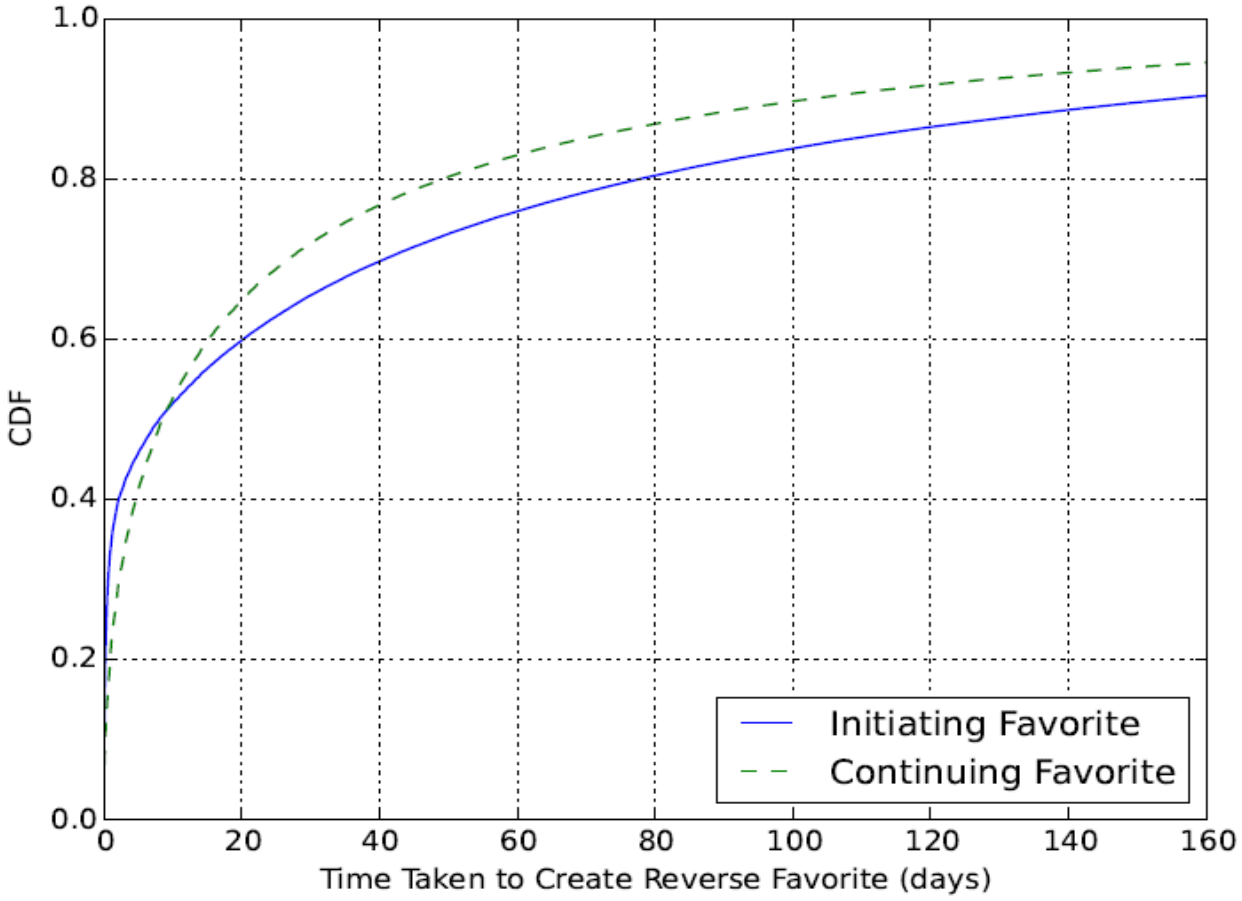
Similar trends in link reception



Reciprocation

- Reciprocation
  - Creation of a link from a node to another causes the creation of a link in the opposite/reverse direction
- If user A favorites a photo uploaded by user B, does that favorite cause user B to later favorite a photo uploaded by user A?

Reciprocation



Over 50% of both kind of favorites reversed within 10 days (more likely the cause)

Within 10 days: Reversal slightly faster in response to initiating favorites
Hypothesis: exploratory nature + homophily

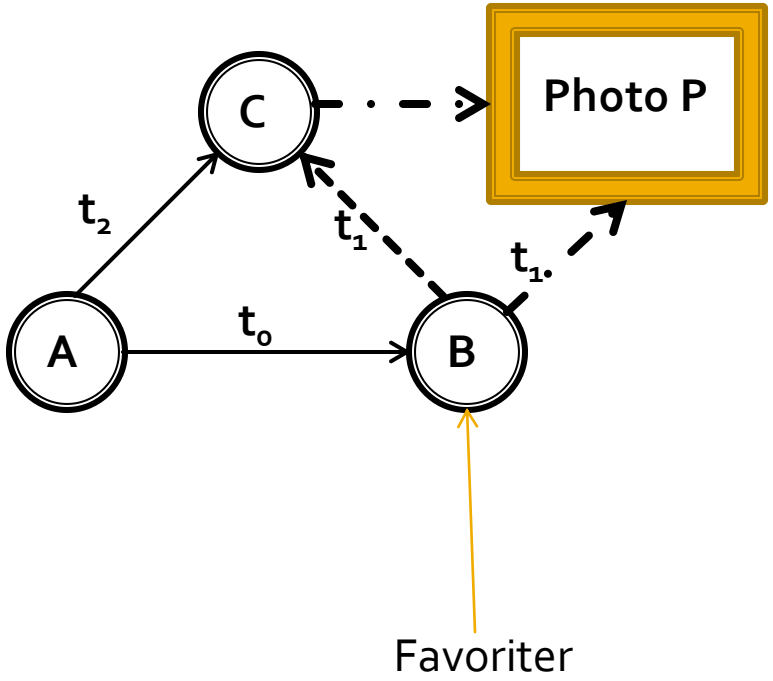
After 10 days: Reversal faster in response to continuing favorites
Hypothesis: homophily

Link Formation after a Favorite

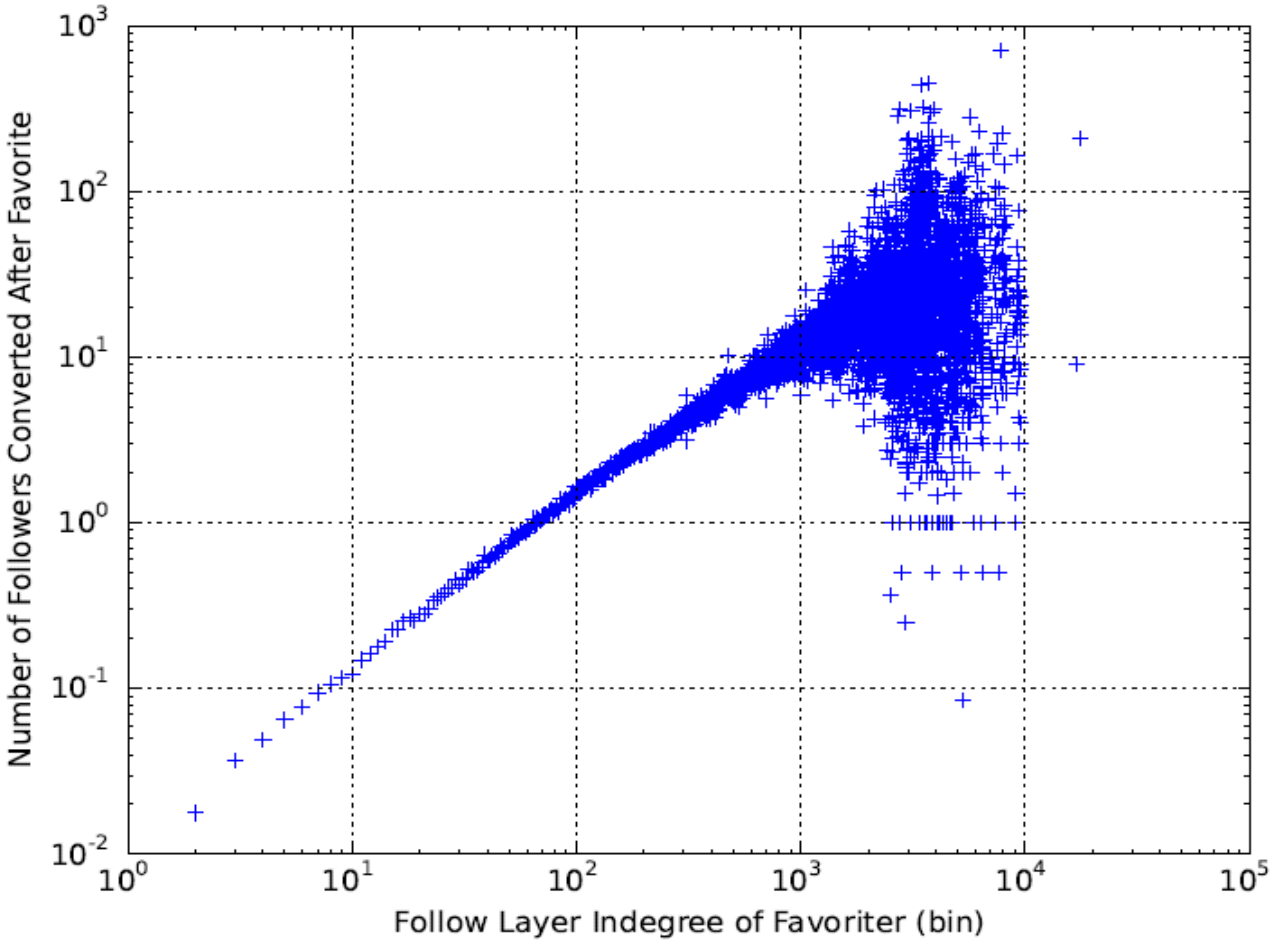
- Lerman and Jones – social browsing an important way by which users browse Flickr
- Mislove et al. – proximity bias in link creation in Flickr
  - Probably few global discovery mechanisms and users could primarily explore their neighborhoods only
- Browsing + link creation in the neighborhood after a social interaction?

Multiplex triangle closure

- For a 2 layer network
  - 8 different types of triangles (Mathematically possible)
  - Not all configurations will make practical sense
    - Given the discovery mechanisms



Multiplex Triangle Closure (Contd.)



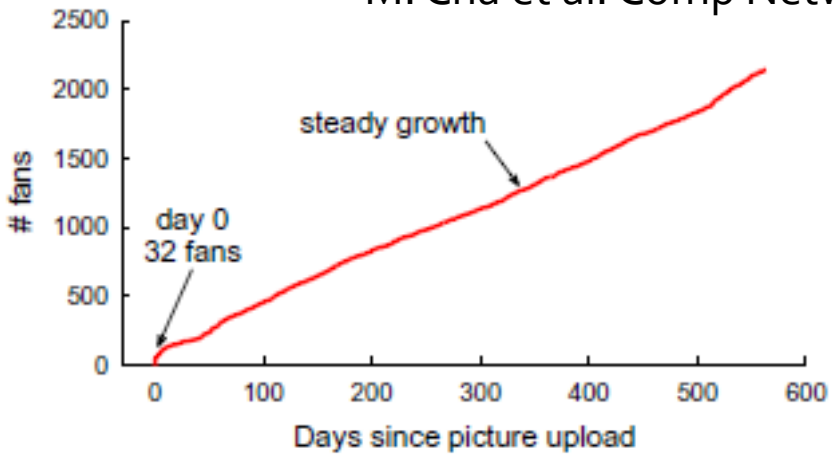
After a favorite is created, foll-fav-foll triangles are closed in proportion to the number of followers of the favoriting user

Summary

- Favoriting of both kinds: photos and users, well-described by preferential creation and preferential reception
- Most favorites reciprocated within 10 days if at all they are reciprocated
  - Observed a difference in reciprocation times for initiating and continuing favorites
- Examined link formation after a favorite via multiplex triangle closure

How quick is the spread?

M. Cha et al. Comp Netw. 12

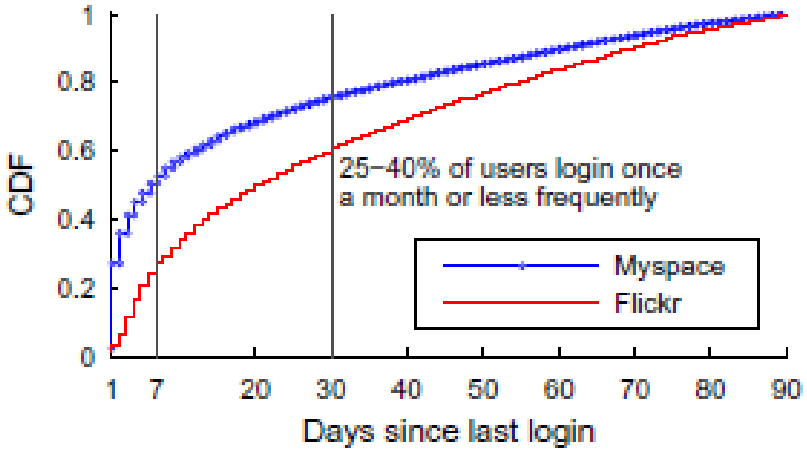


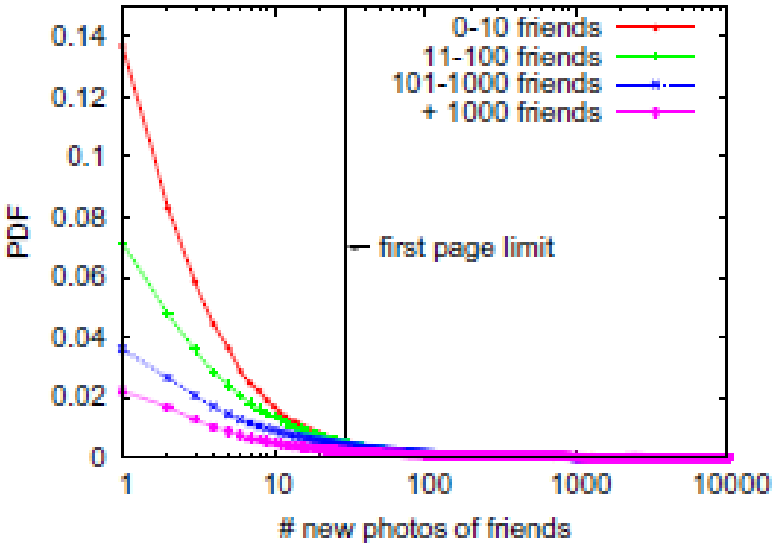
- Strong content locality
  - About 90% of the fav-markings come from 2-hop neighborhood of uploaders
- Steady increase in popularity over a long period of time

Meeyoung Cha, Fabrício Benevenuto, Yong-Yeol Ahn, Krishna P. Gummadi, Delayed information cascades in Flickr: Measurement, analysis, and modeling, Computer Networks, Volume 56, Issue 3, 23 February 2012, Pages 1066-1076

Possible reasons

- Bursty login behavior
  - Large fraction of user logs in less frequently
  - Adds to the cascading delay
- Content aging
  - Probability of checking photos
    - drops rapidly with increasing number of web clicks





Summary

- High content locality
  - Content spreads mostly up to 2 hops
- High interactivity and social cascades
  - Popular users
  - Reciprocation
- High delay in spreading
  - Bursty login behavior
  - Content aging

Acknowledgement

Alan Mislove – for graciously sharing the data set

Organizers – For providing the opportunity to present

References

- **Z. Yang, J. Xue, X. Zhao, X. Wang, B. Y. Zhao and Y. Dai.** "Unfolding Dynamics in a Social Network: Co-evolution of Link Formation and User Interaction". WWW '13 Companion.
- **B. Viswanath, A. Mislove, M. Cha and K. P. Gummadi.** "On the Evolution of User Interaction in Facebook". WOSN '09.
- **J. Leskovec, L. Backstrom, R. Kumar and A. Tomkins.** "Microscopic Evolution of Social Networks". KDD '08.
- **A. Mislove, H. S. Koppula, K. P. Gummadi, P. Druschel and B. Bhattacharjee.** "Growth of the Flickr Social Network". WOSN '08.
- **K. Lerman and L. Jones.** "Social Browsing on Flickr". ICWSM '07.