

Inferring “*stuff*” from observed networks



16.5.2012
David Stolz

Agenda

- 1 Structure of Approaches
- 2 Recommendation Network
- 3 Blogs
- 4 “Meta-Conclusion”

Structure of Approaches



Understand
Data



Define Goals /
Categorize



Method



Infer

Compare

Add Knowledge

Recommendation Network



Recommendation Network



- 4 Mio. Users
- 16 Mio. Recommendations
only ~3% of purchases associated with recommendation
- 2 Years
- Monetary benefit for recommender ***and*** recommendee

Recommendation Network

- Analyze *cascades*



- Categorize by different product categories
 - Books, DVD, Music, Video

Recommendation Network

- Remove:
 - no-purchase nodes
 - Late recommendations

- Find all local subgraphs
 - ➔ Isomorphism test



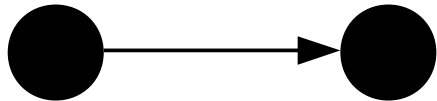
Recommendation Network

- Most frequently observed cascade?



Recommendation Network

- Most frequently observed cascade?

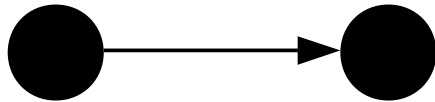


- Differences: Books, DVD, Music, Video?



Recommendation Network

- Most frequently observed cascade?



- Differences:

- Books: 70%
- **DVD: 12%**
- Music: 86.4%
- Video: 74%



Recommendation Network

- Overall: splits = 5 * collisions
- Simple graphs sometimes *more rare* than complex graphs



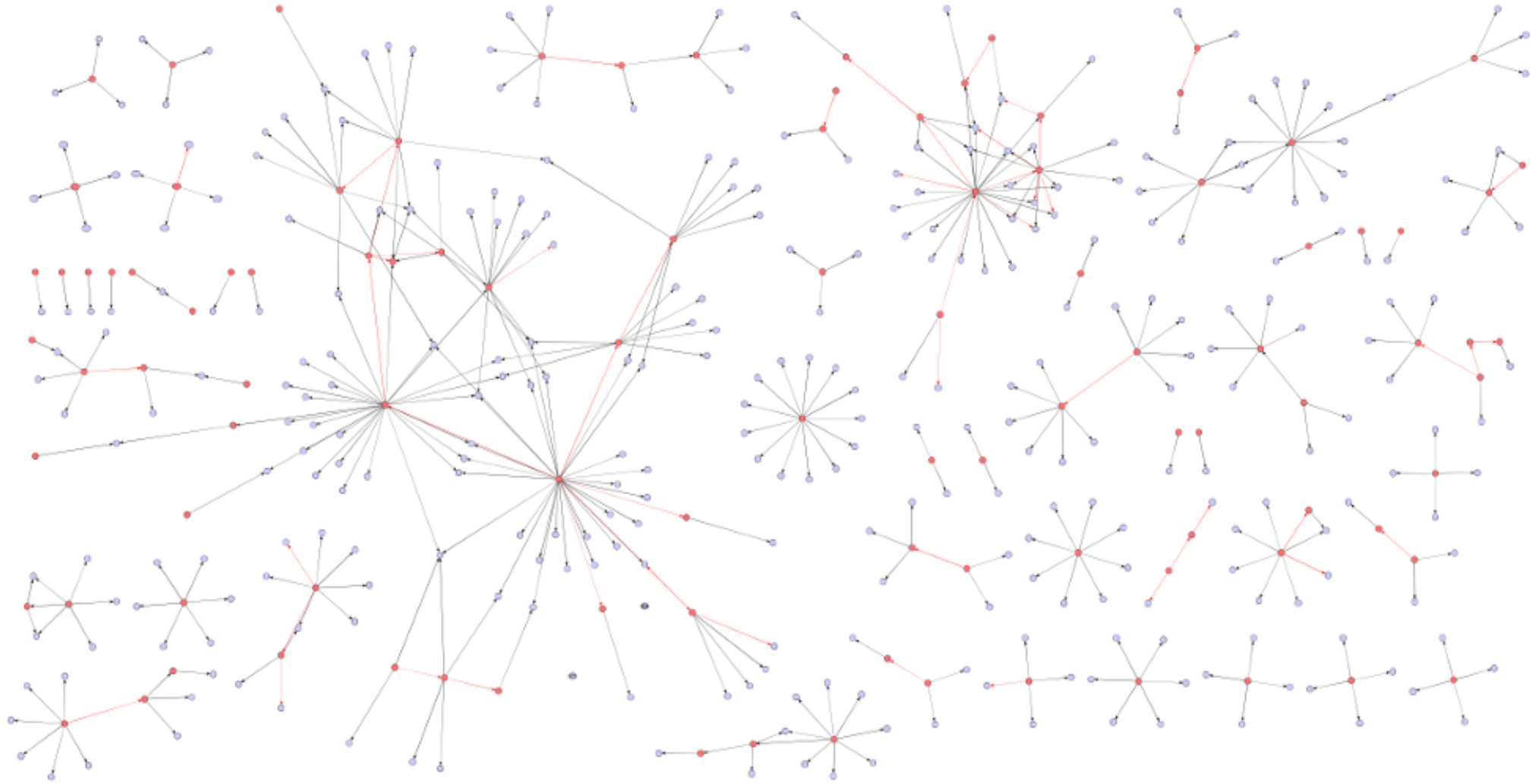
Recommendation Network

Paper Conclusions

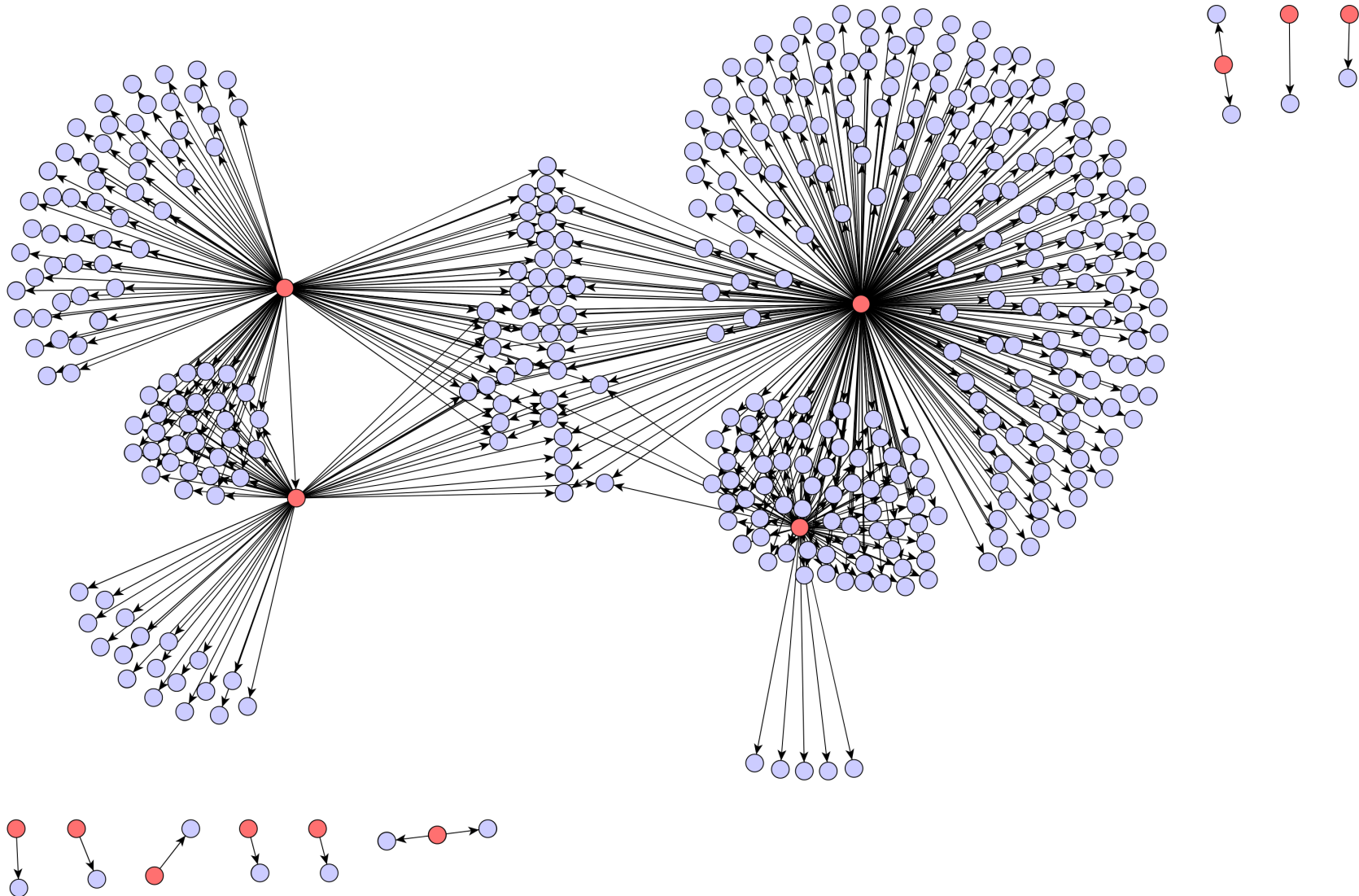
- Most cascades are small
- Underlying social networks lead to (*measurably*) more complex cascades



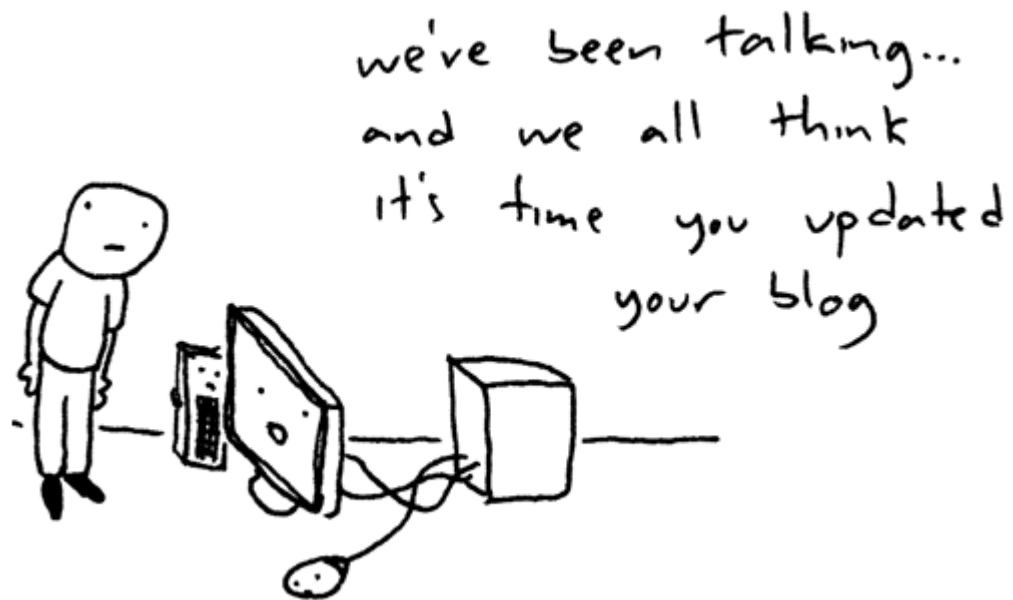
Recommendation Network



Recommendation Network



Blogs



Blogs

- 4 Years (1999 – 2002)
- 25'000 Blogs
- 750'000 Links (*between* blogs)



Blogs

- Exact notion of time
- Only actual entries
 - Filter out “Side-bars”



Blogs

- Time characteristics
- Community structure
- Bursts



Blogs

Time Graph:

- Label Edges with time
- Label Nodes with time interval



- Prefix Graph \mathbf{G}_t :

- Subgraph of \mathbf{G} up to time t

Blogs

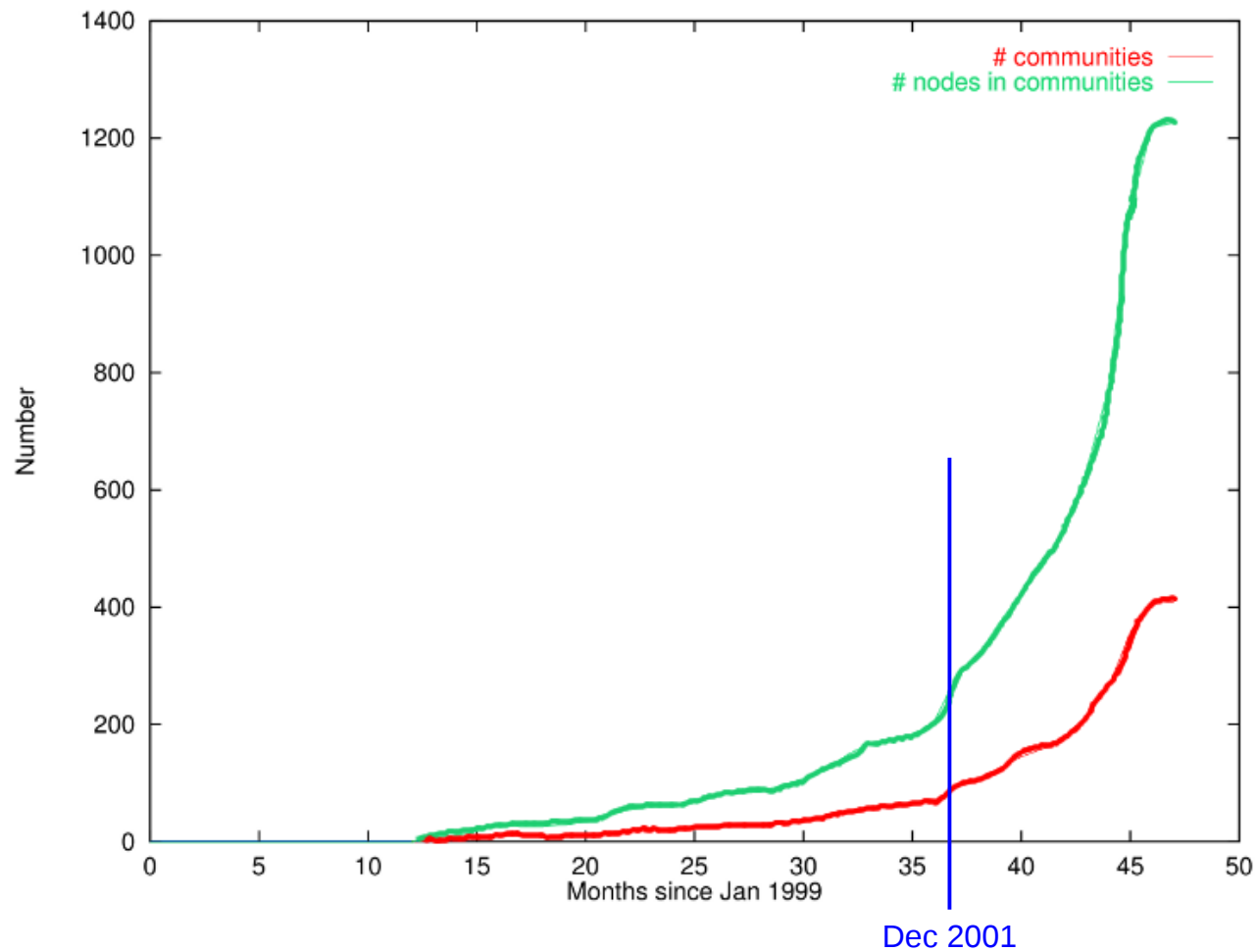
Community Extraction

- Two step algorithm:
 - Find new community
 - Expand it



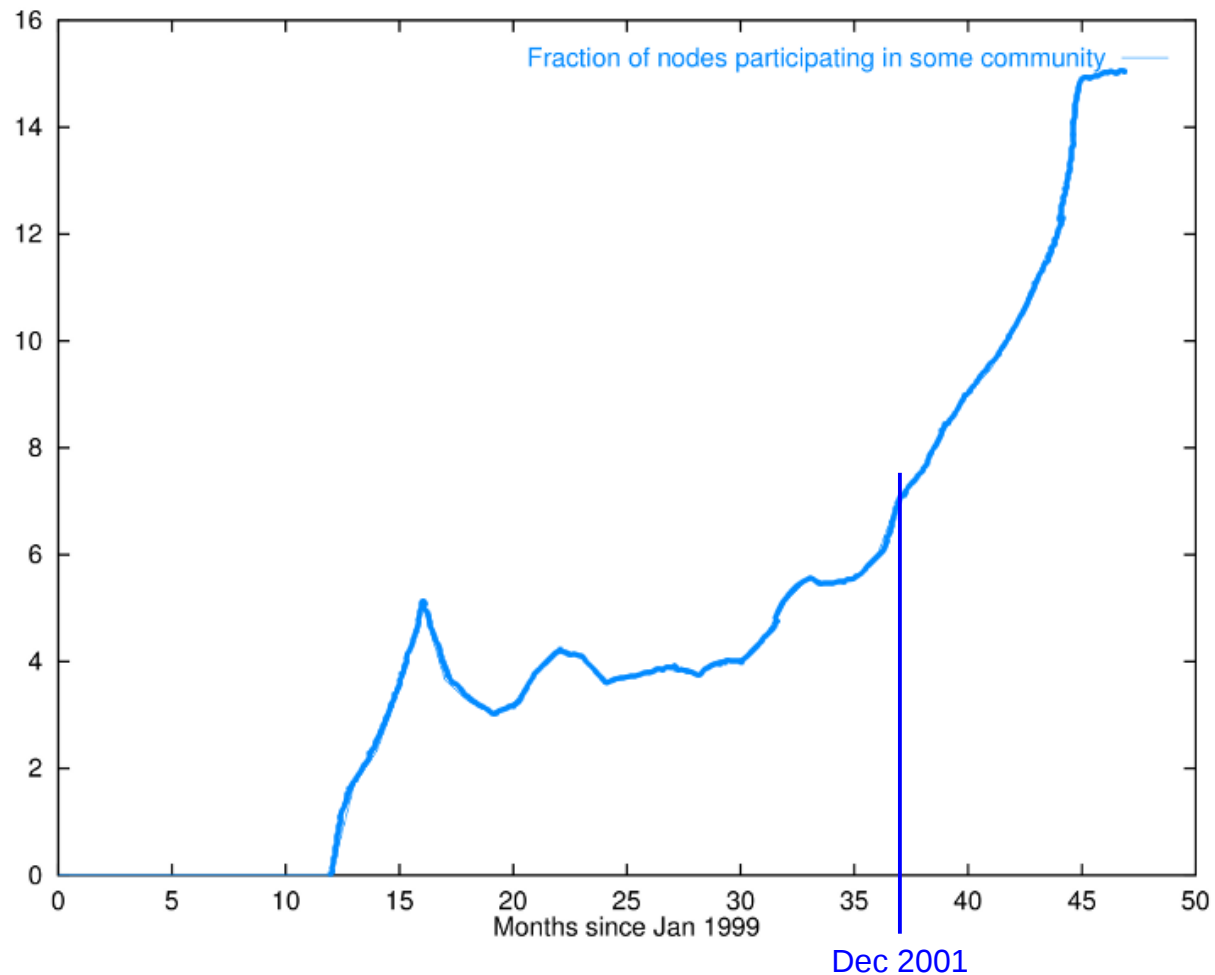
Blogs

Communities (based on Prefix Graphs)



Blogs

Communities (based on Prefix Graphs)



Fraction

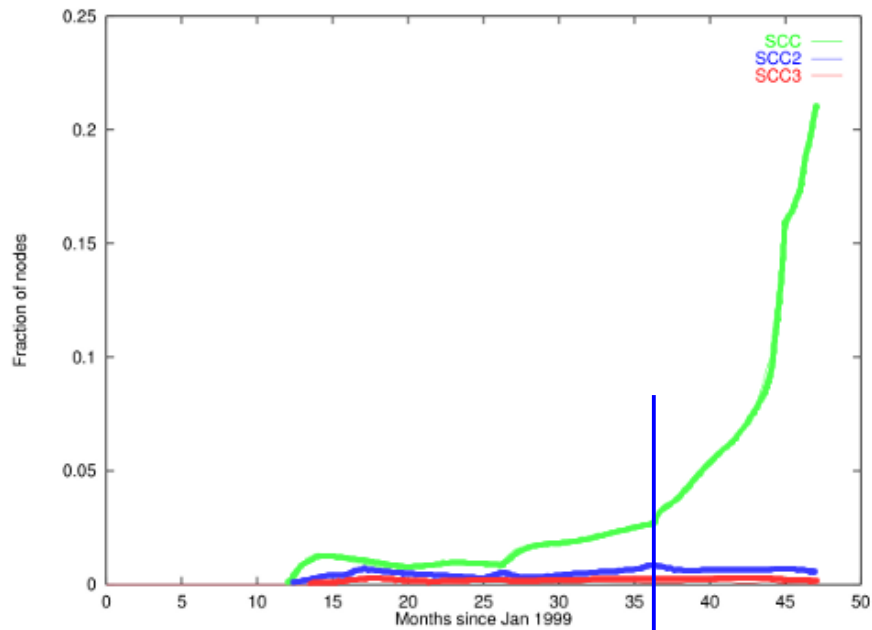
Fraction $\in [0,16]$?



Blogs

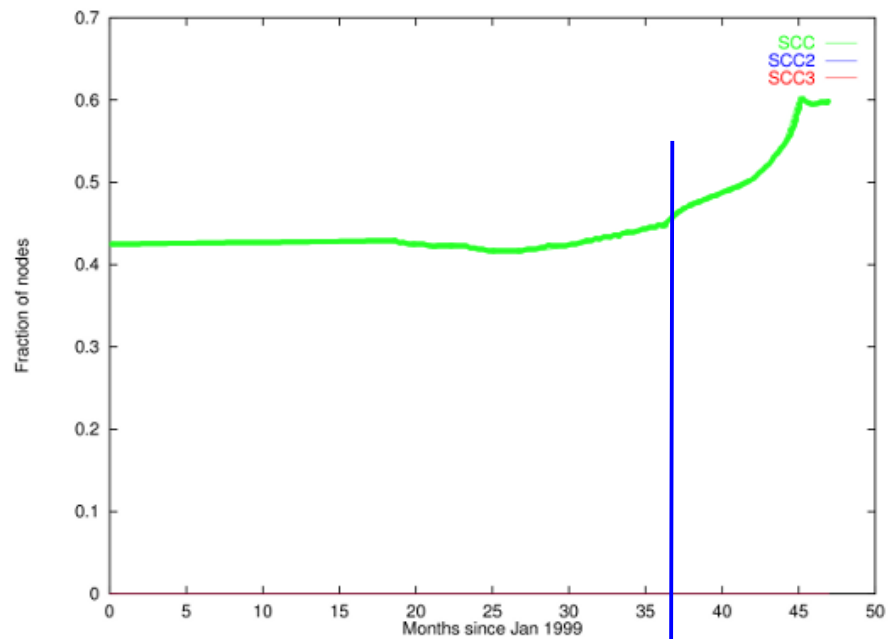
SCC Comparison against “Random” Graph

Observed



Dec
2001

“Random”

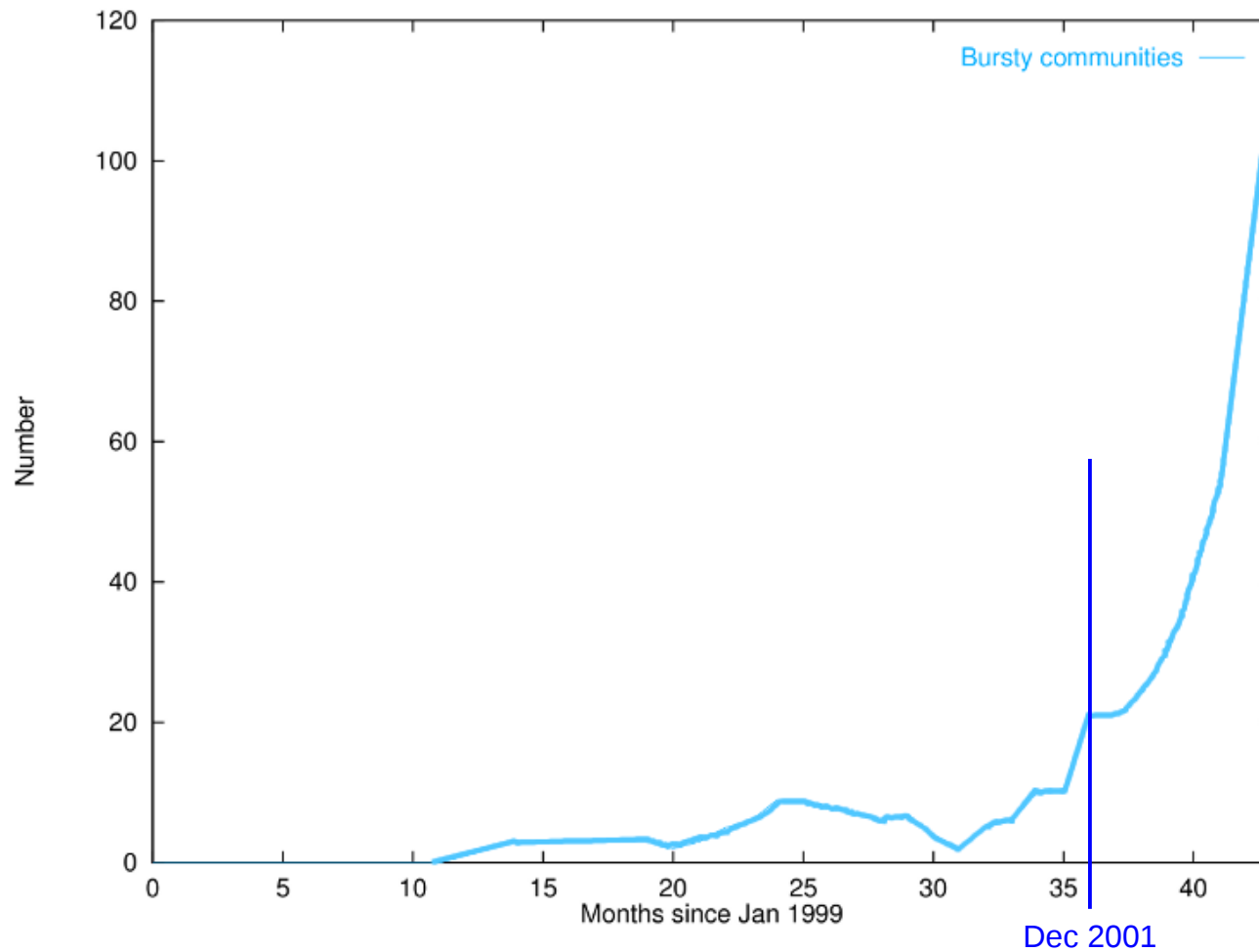


Dec 2001



Blogs

Bursts



Blogs

Paper Conclusions

- End of 2001:
 - #Communities: increased
 - Connectedness: increased
 - Burstyness: increased
- ➔ User behavior has changed



Blogs

In another community, a blogger Dawn hosts a poll to determine the funniest and sexiest blogger. She conducts interviews with other bloggers in the community, of course listing their sites. She then becomes obsessed with one of the other bloggers Jim, which spurs comments by many others in the community.

Blogs

In another community, a blogger Dawn hosts a poll to determine the funniest and sexiest blogger. She conducts interviews with other bloggers in the community, of course listing their sites. She then becomes obsessed with one of the other bloggers Jim, which spurs comments by many others in the community.



“Meta-Conclusion”

- Empirical results matter, even if they don't astonish
- **Every** step of the 4 step approach influences the result!
- Talk is silver, silence is golden.
(= don't publish papers just for the sake of publishing them)

DIABLO

Battle.net Account Name

Password

LOGIN

Remember Account

CREDITS

CINEMATICS

MANAGE ACCOUNT

OPTIONS

EXIT

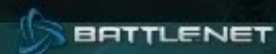
BREAKING NEWS

We're aware that intermittent failures are occurring during character and game creation and are in the process of investigating. Please feel free to keep trying if you're receiving an error at character or game creation as they should eventually complete successfully.

If you're experiencing a different issue please visit the beta forums at www.Battle.net/D3



Diablo III © 2011 Blizzard Entertainment. All rights reserved.



Discussion

