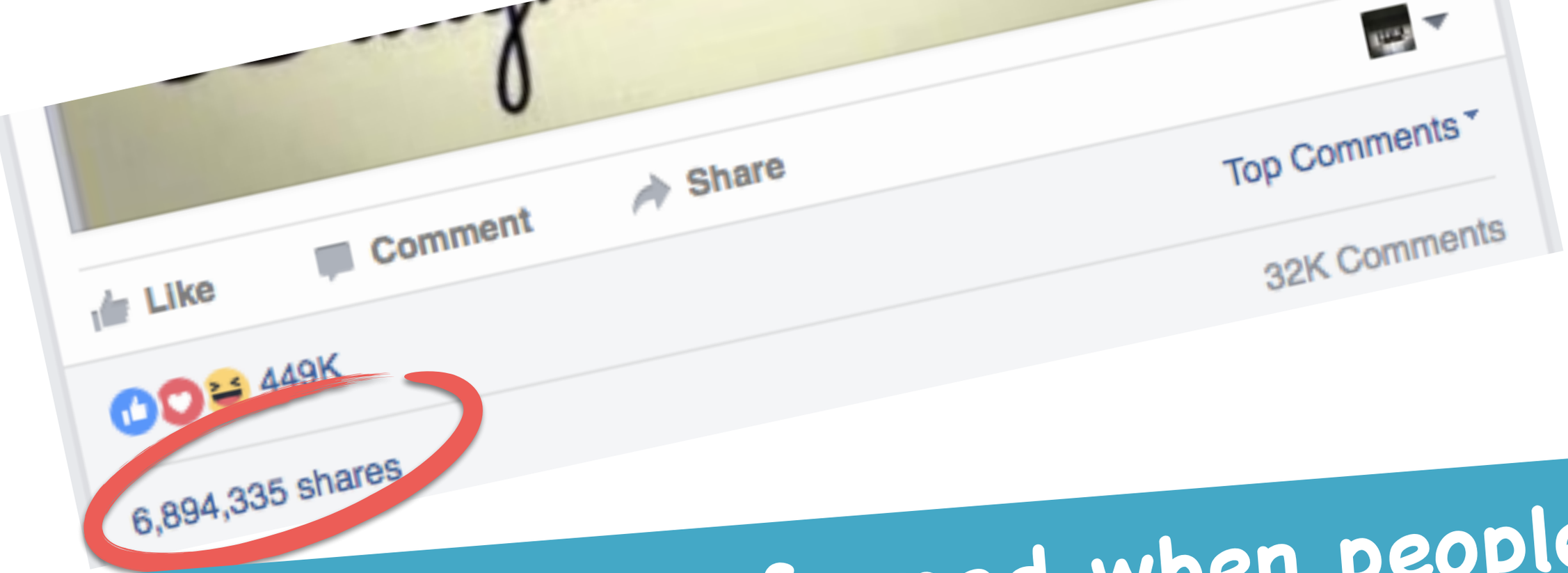
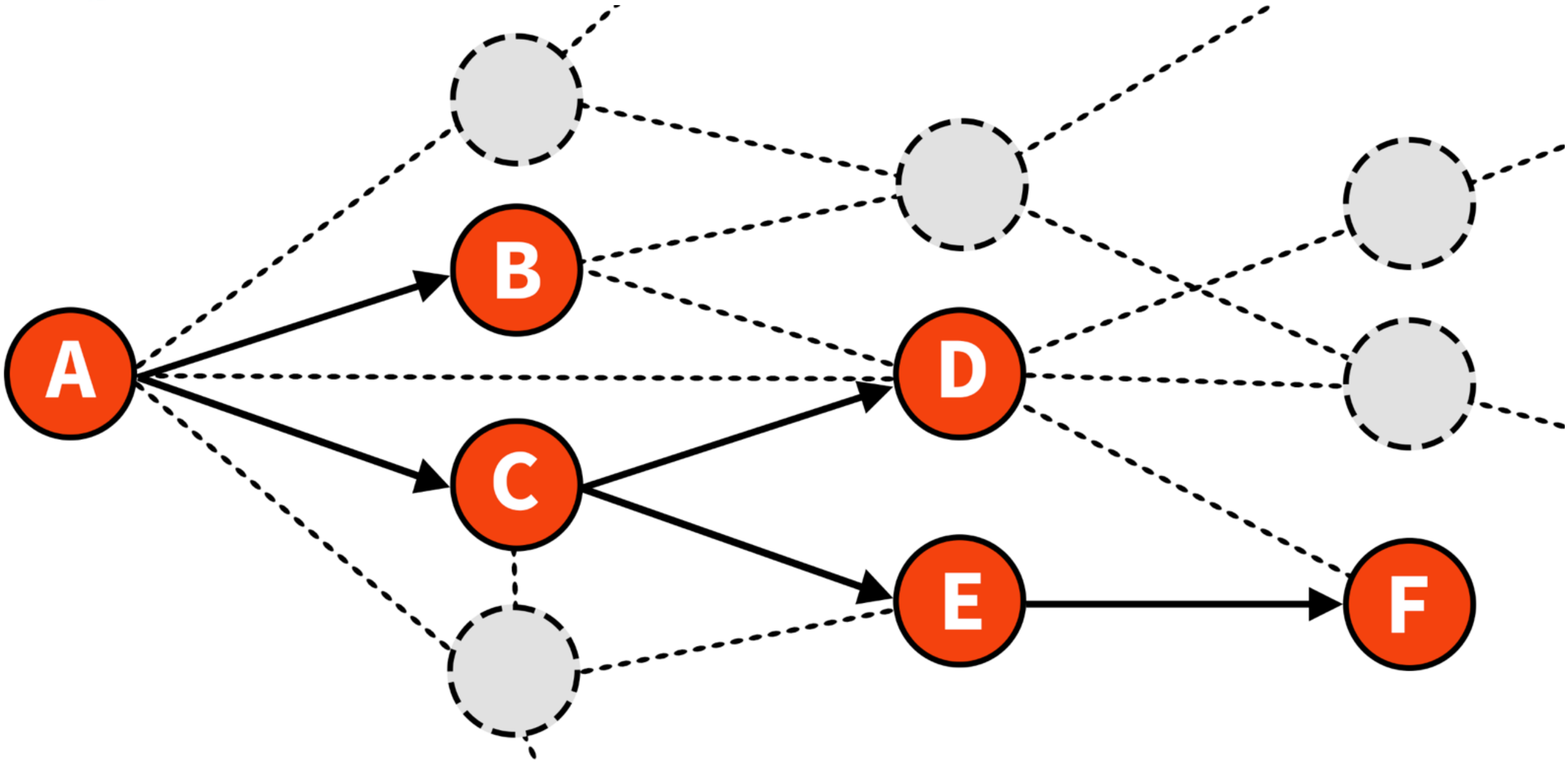
Information Diffusion in Social Networks

A brief overview

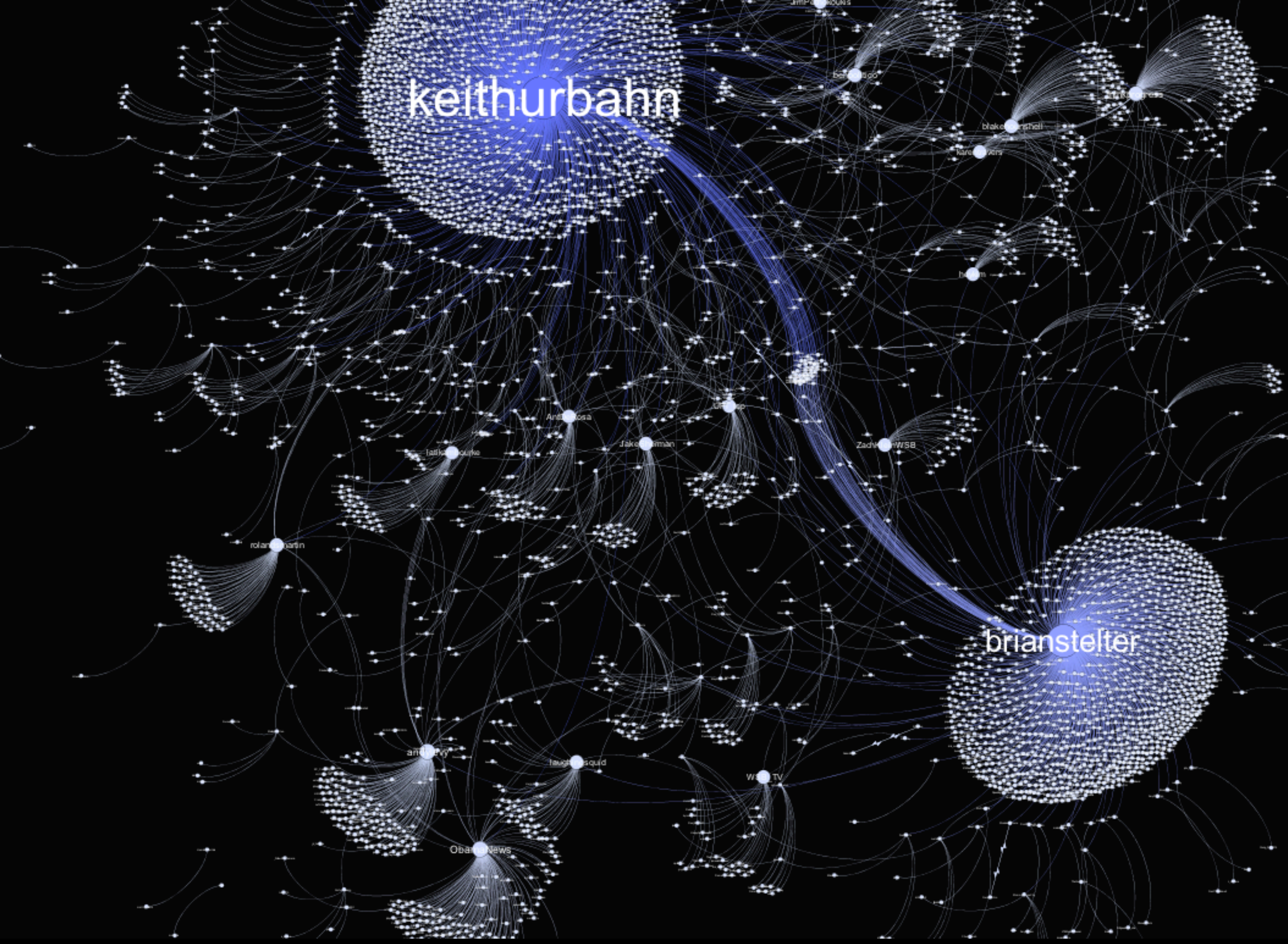


Cascades are formed when people (re)share information with one another

An Information Cascade



J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In WWW '14, 2014.


keithurbahn
brianstelter

diseases

product adoption

Human behaviour

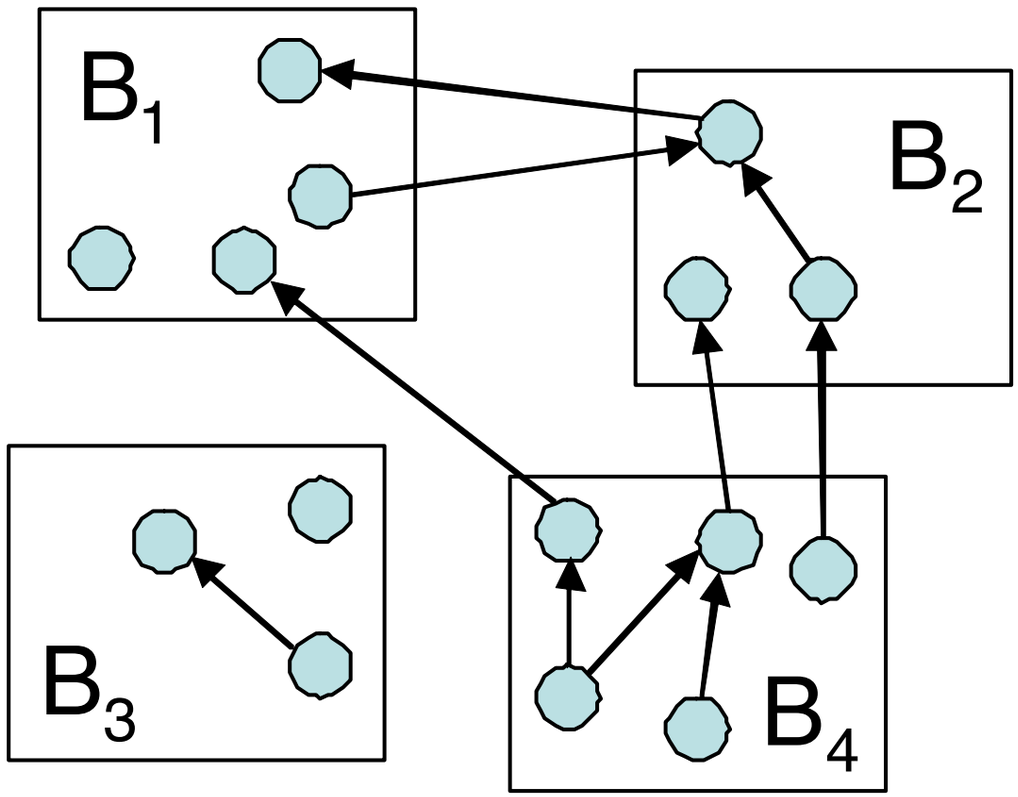
Motivation...??

rumours

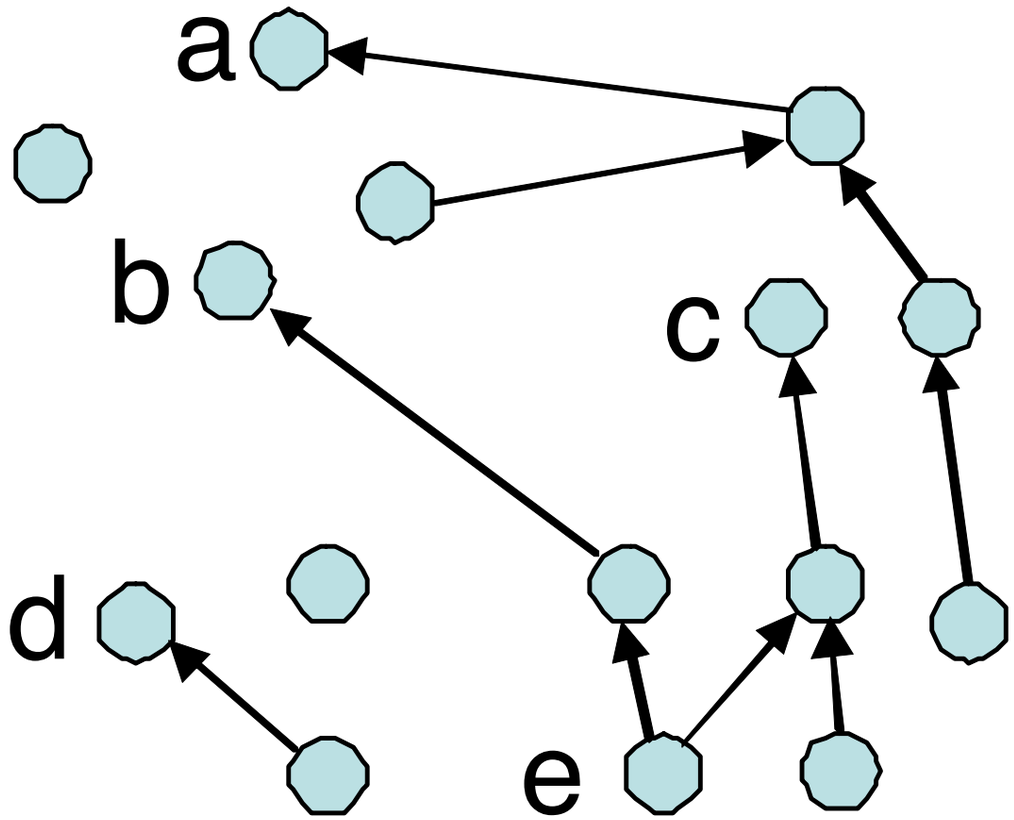
recommendations

news

Cascades in Blogosphere



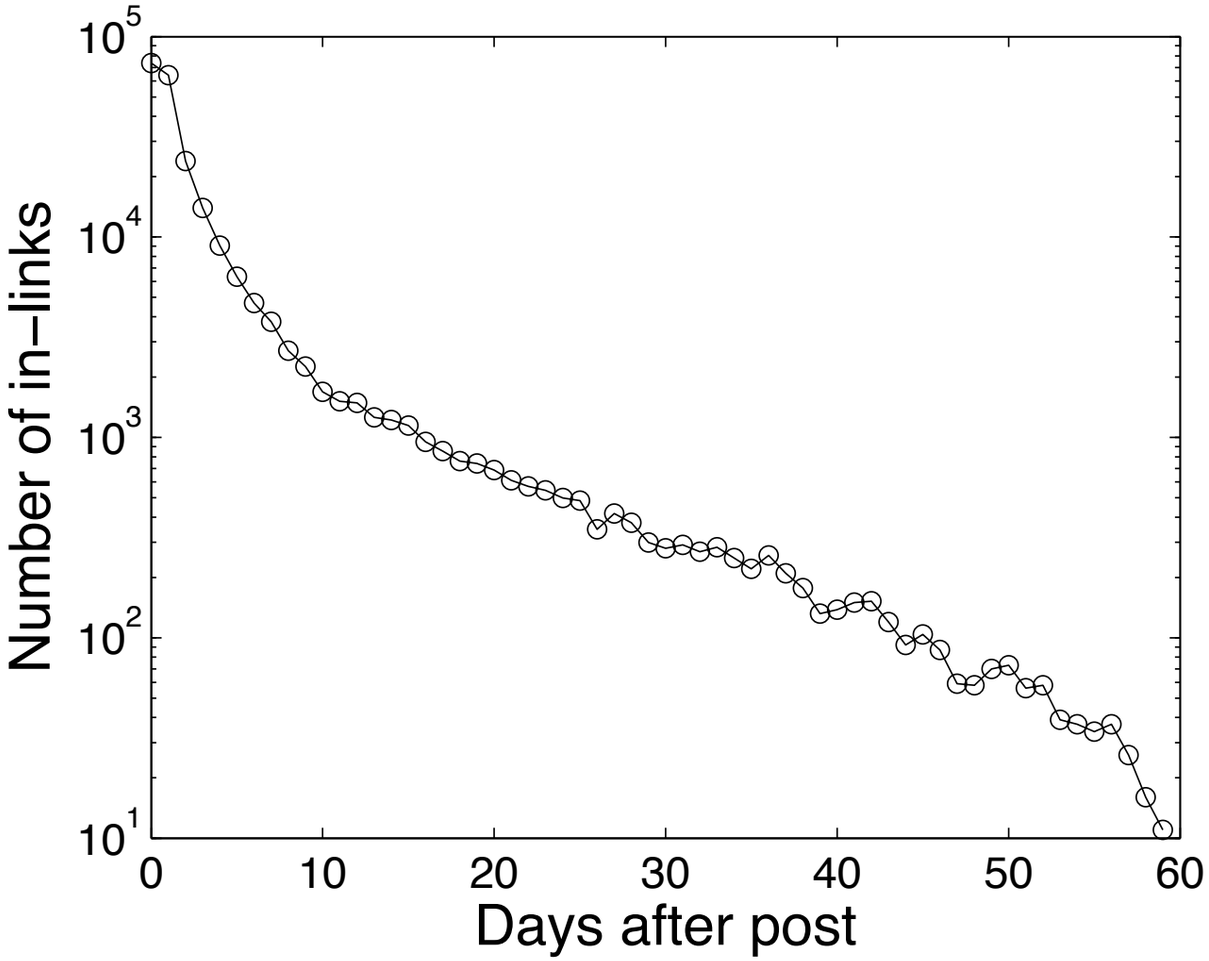
Blogosphere

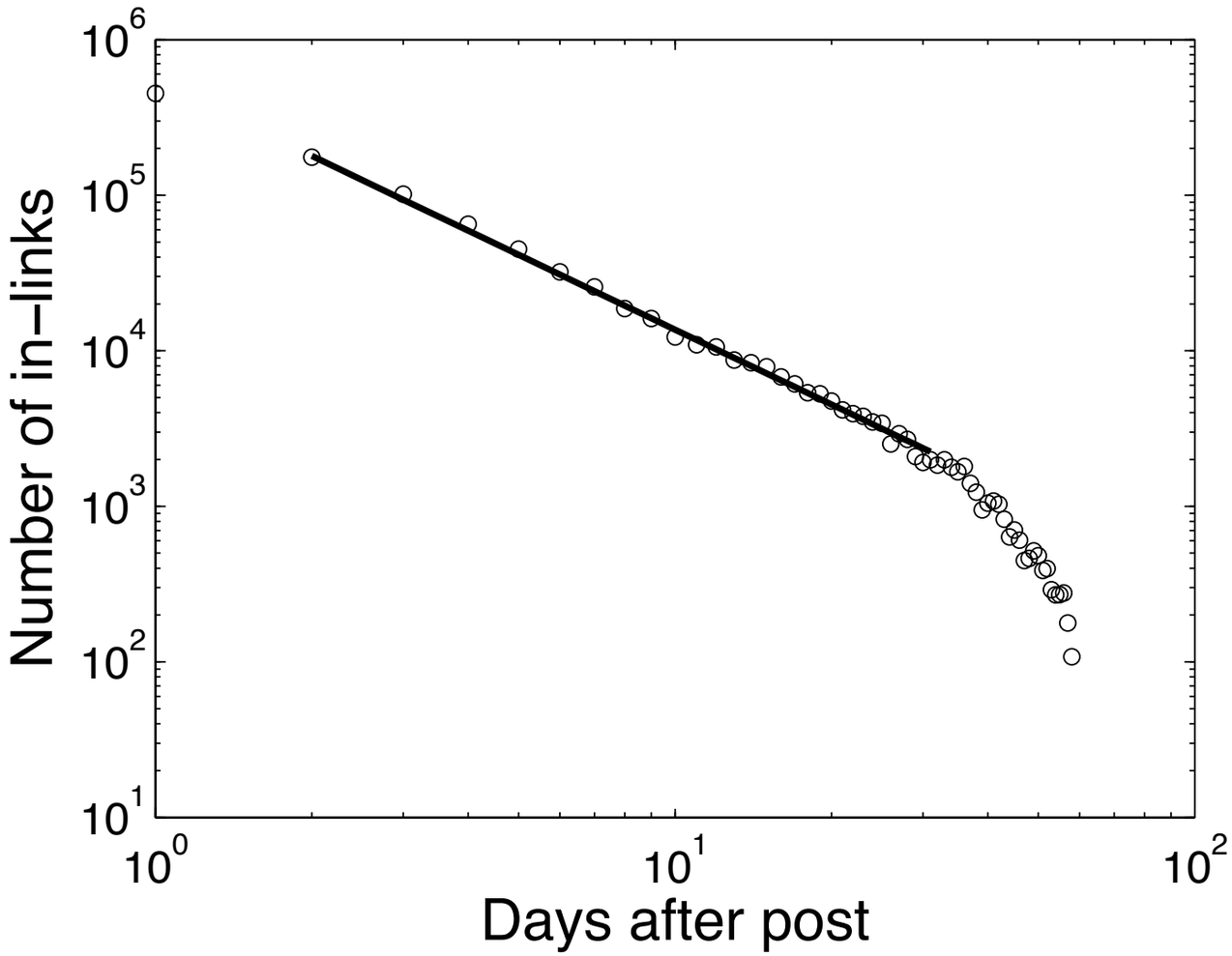


Network of posts in the Blogosphere

J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

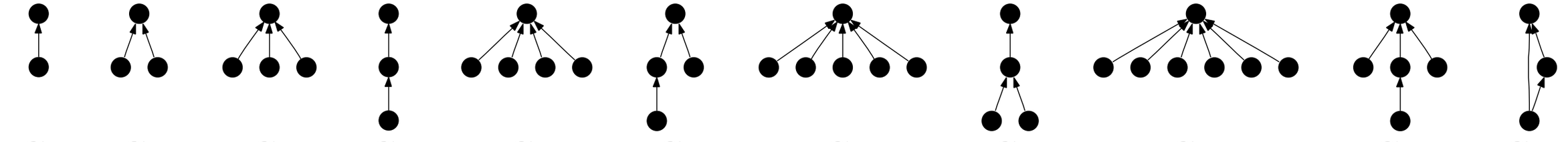
Popularity dies with time!!

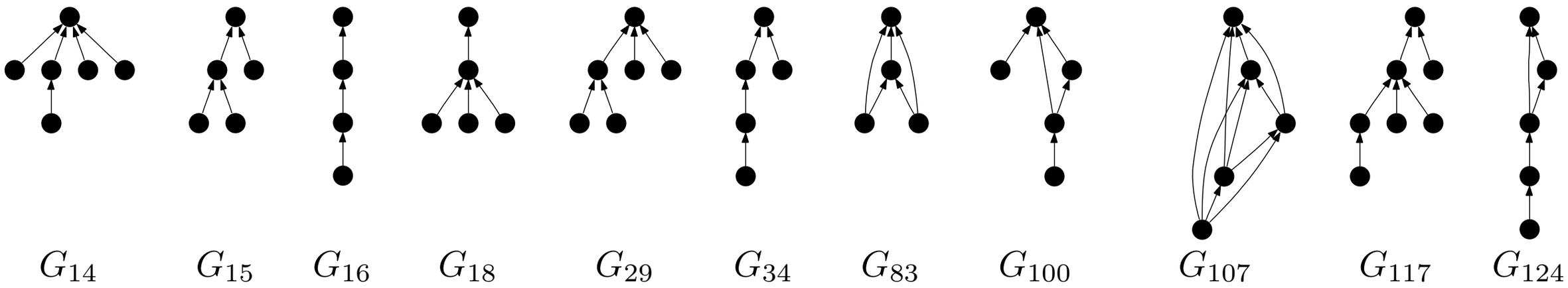




J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

Common Cascade Shapes

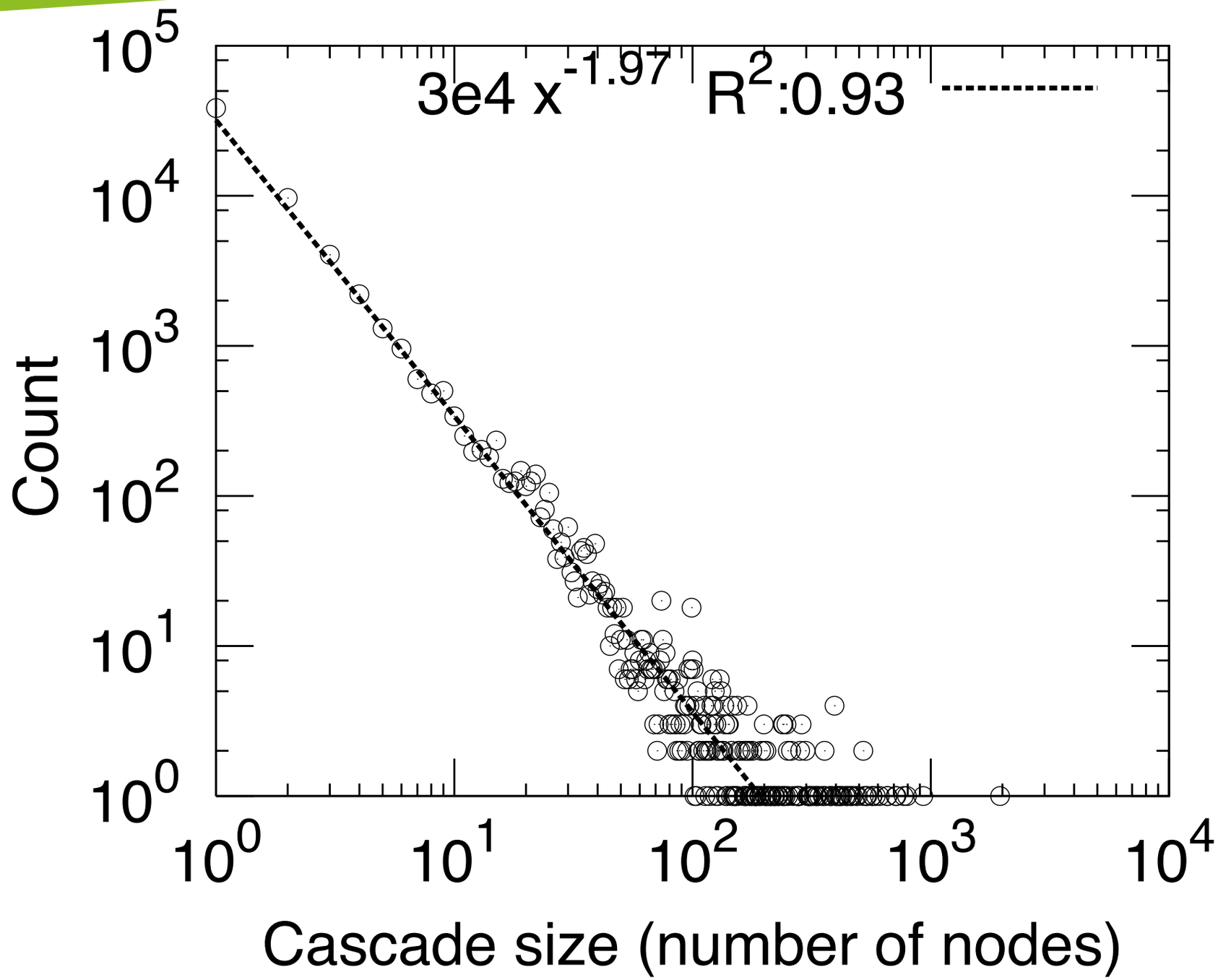




Common Cascade shapes ordered by the frequency

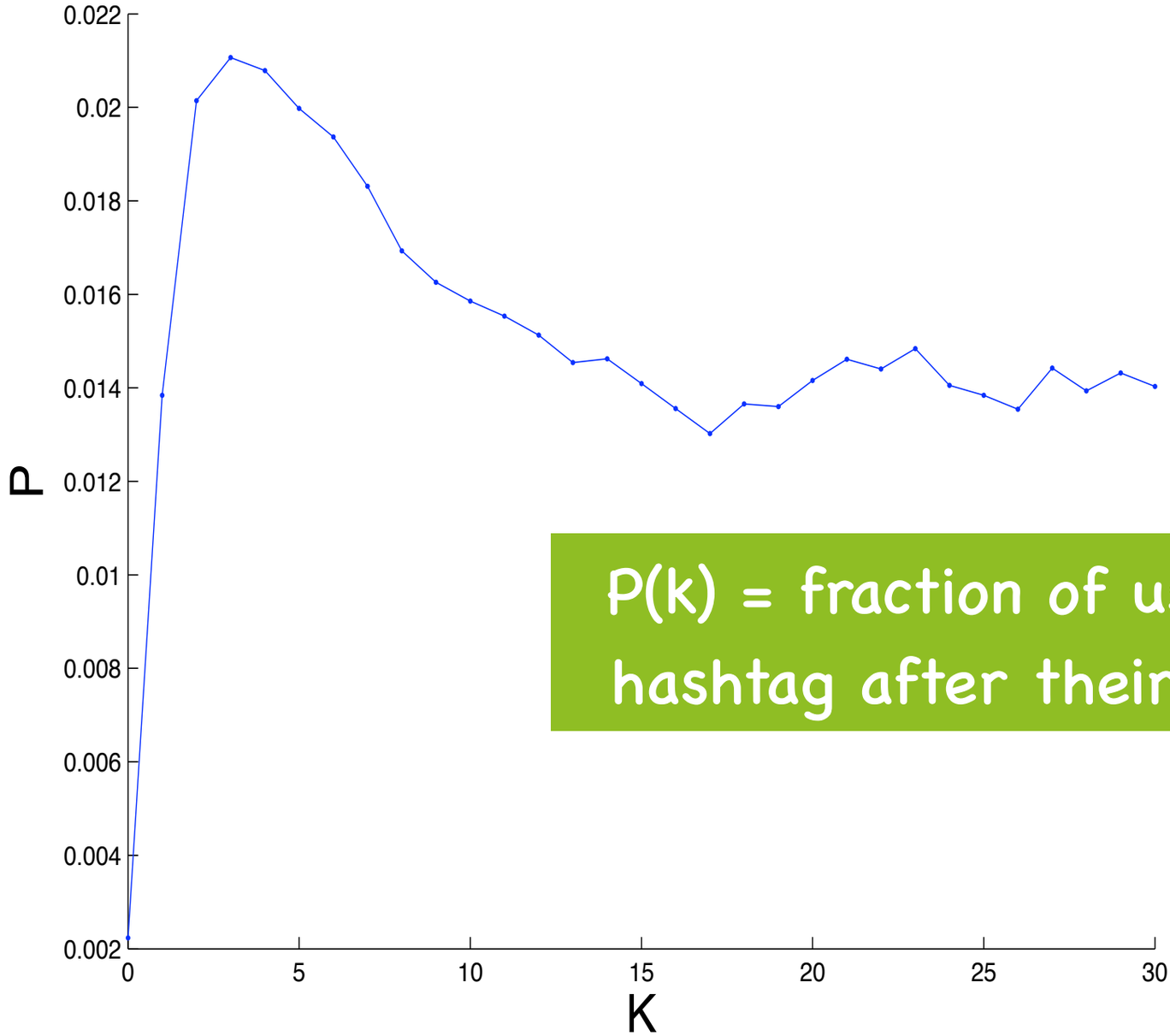
J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

Cascade Size Distribution



J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

How many exposures do you need...?

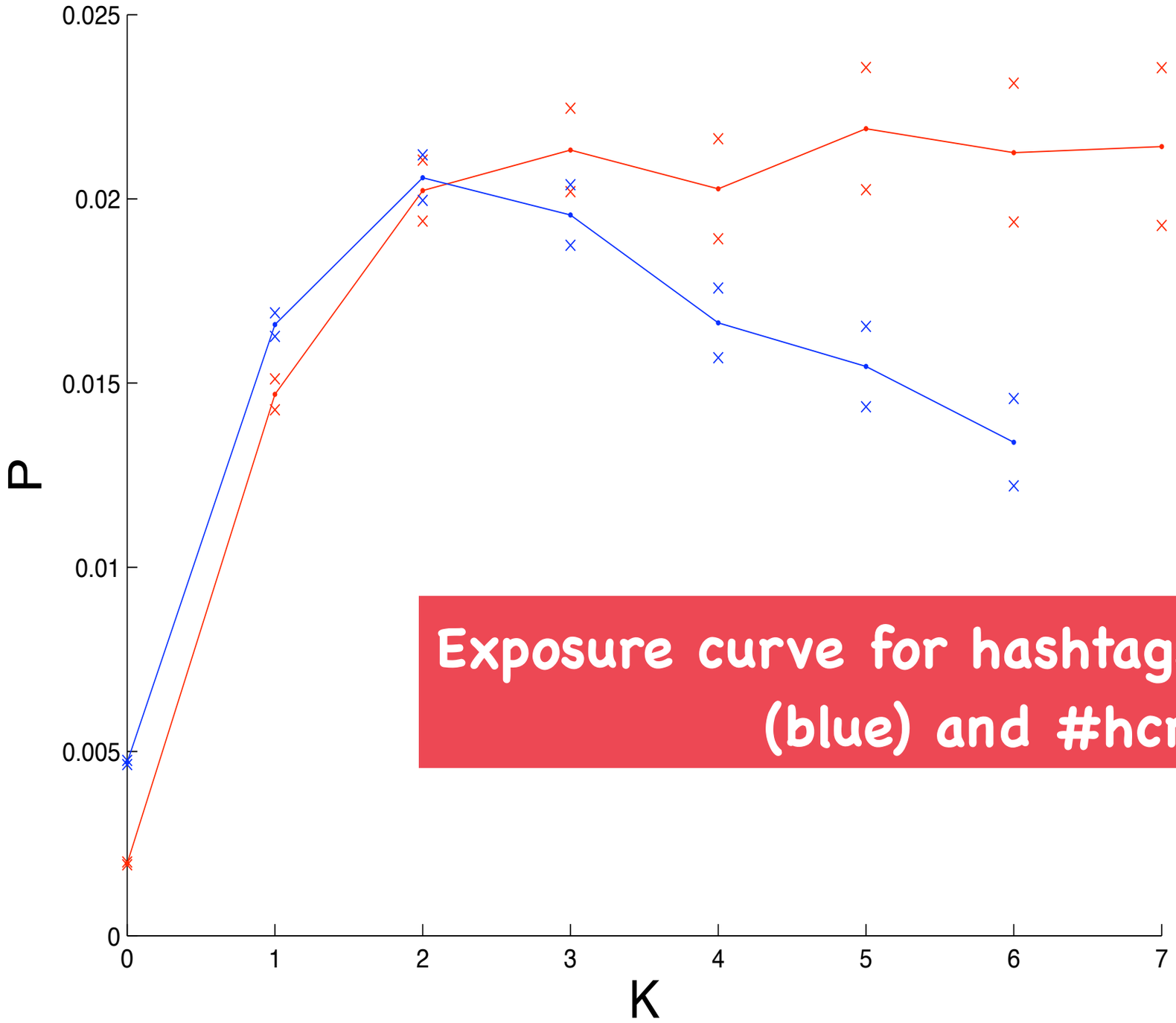


P(k) = fraction of users who adopt hashtag after their k-th exposure

Average exposure curve for top 500 hashtags

Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. WWW11

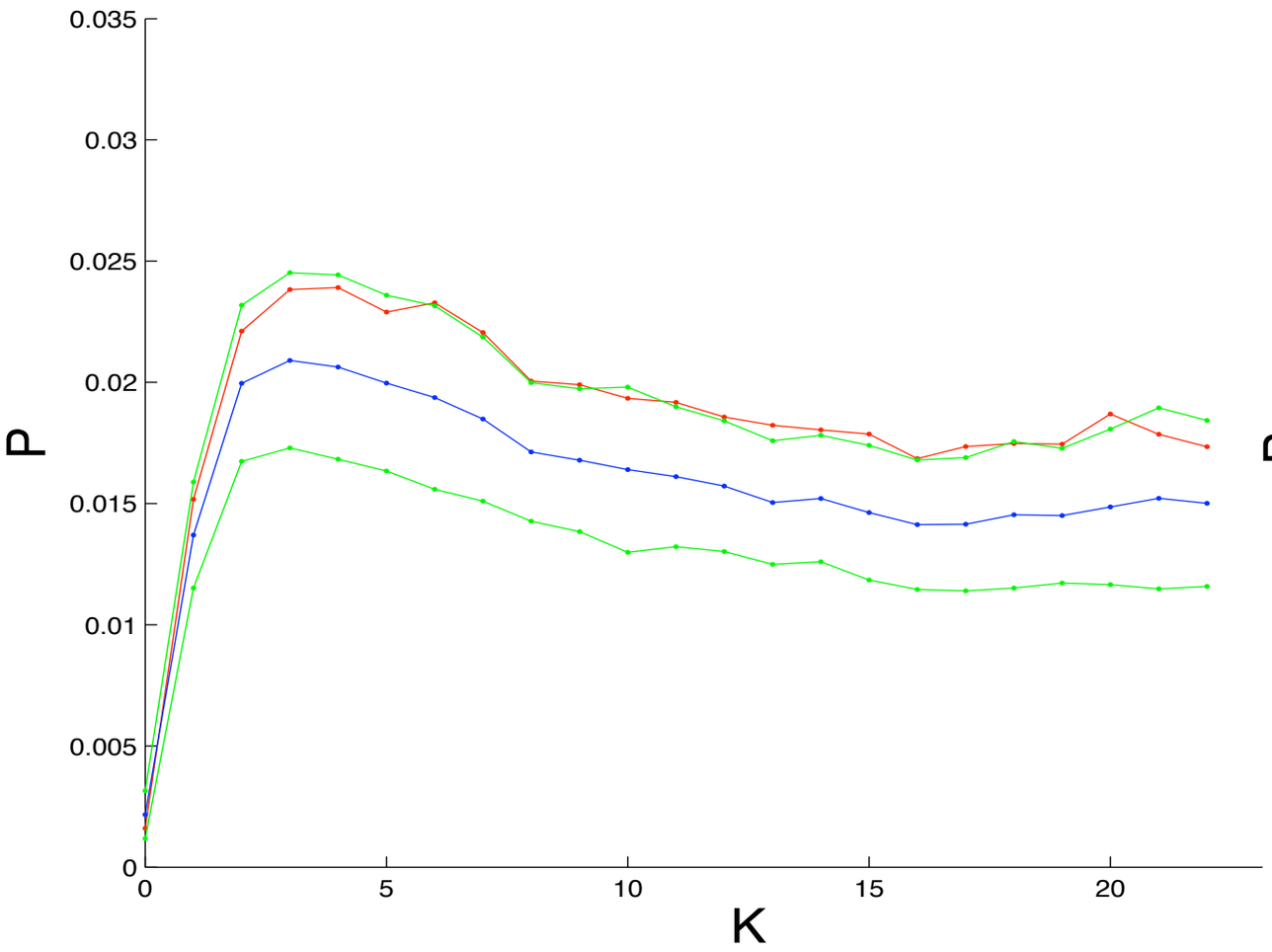
Example: Exposure v/s Adoption



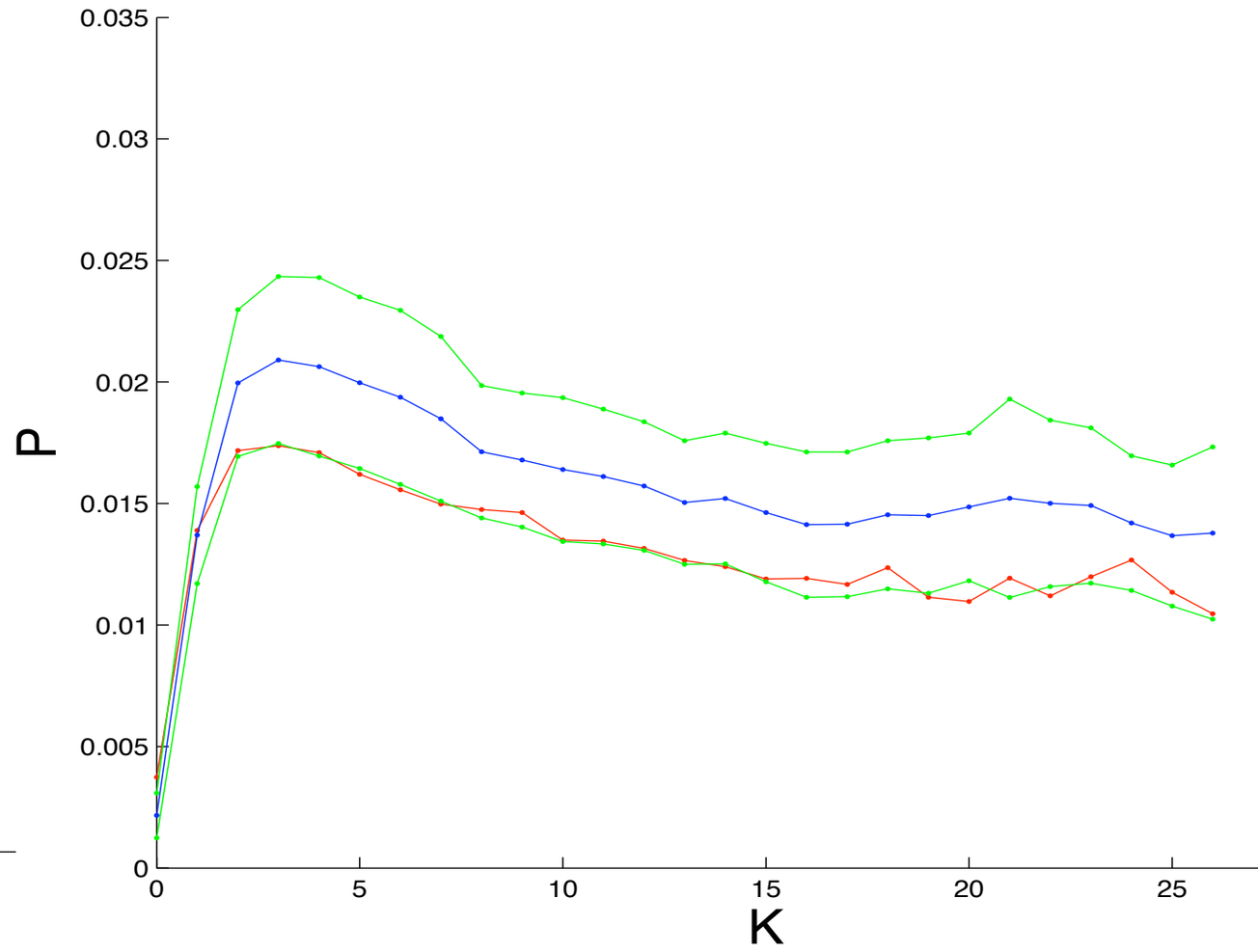
Exposure curve for hashtags #cantlivewithout (blue) and #hcr (red)

Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. WWW11

"Complex Contagion" phenomena



Politics, Celebrity, Idioms, and Music



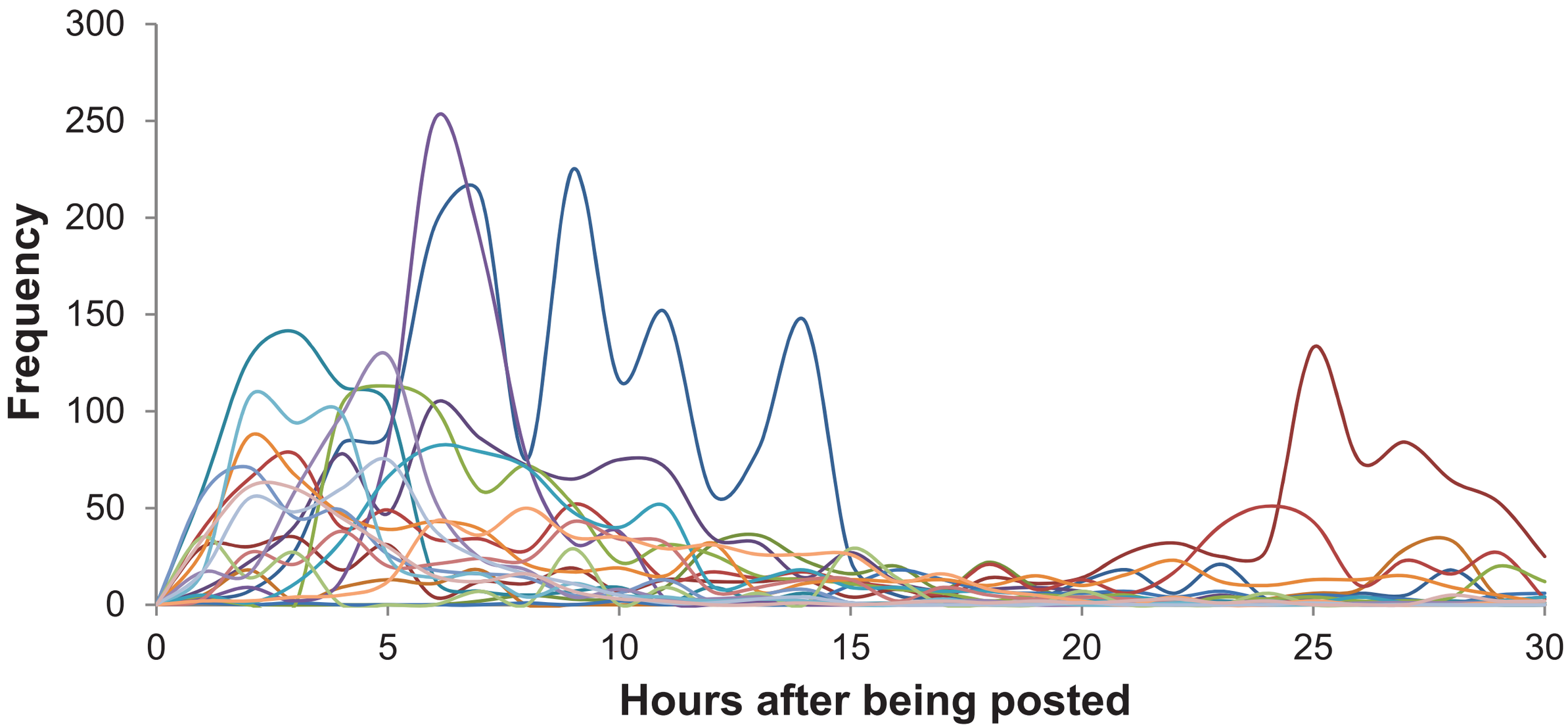
Technology, Movies, and Sports

Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. WWW11

"Adoption of ***politically controversial*** hashtags are ***affected*** by **multiple exposures**, while repeated exposures have ***less effect*** on adoption of ***conversational idioms***…"

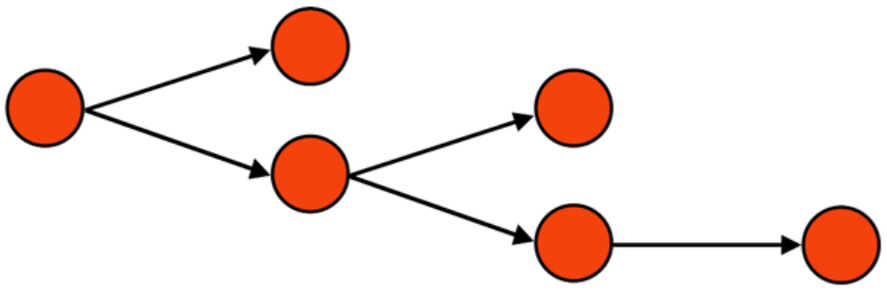
Can cascades be predicted...??

Change in frequency of hashtags with time

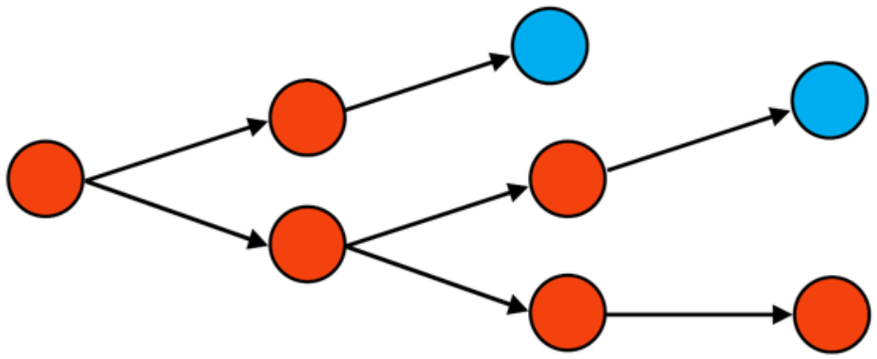


J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

Will cascade reach median size?

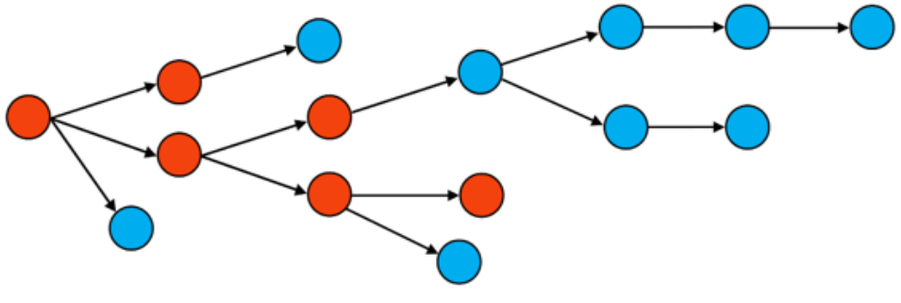


k reshares



?

less than the median *f(k)*



?

more than the median *f(k)*

J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In WWW '14, 2014.

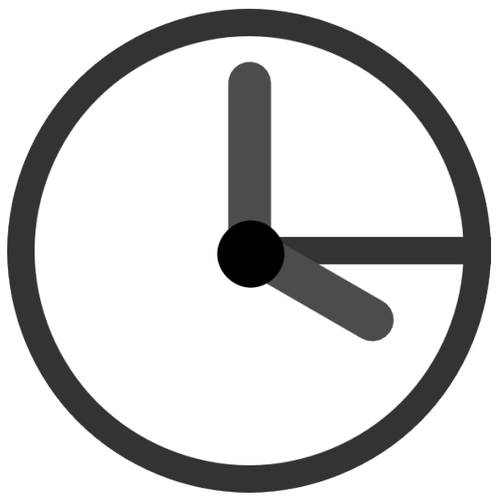
What factors affect predictability??

Content

e.g. +ve or -ve emotion

User

e.g.topicality of user



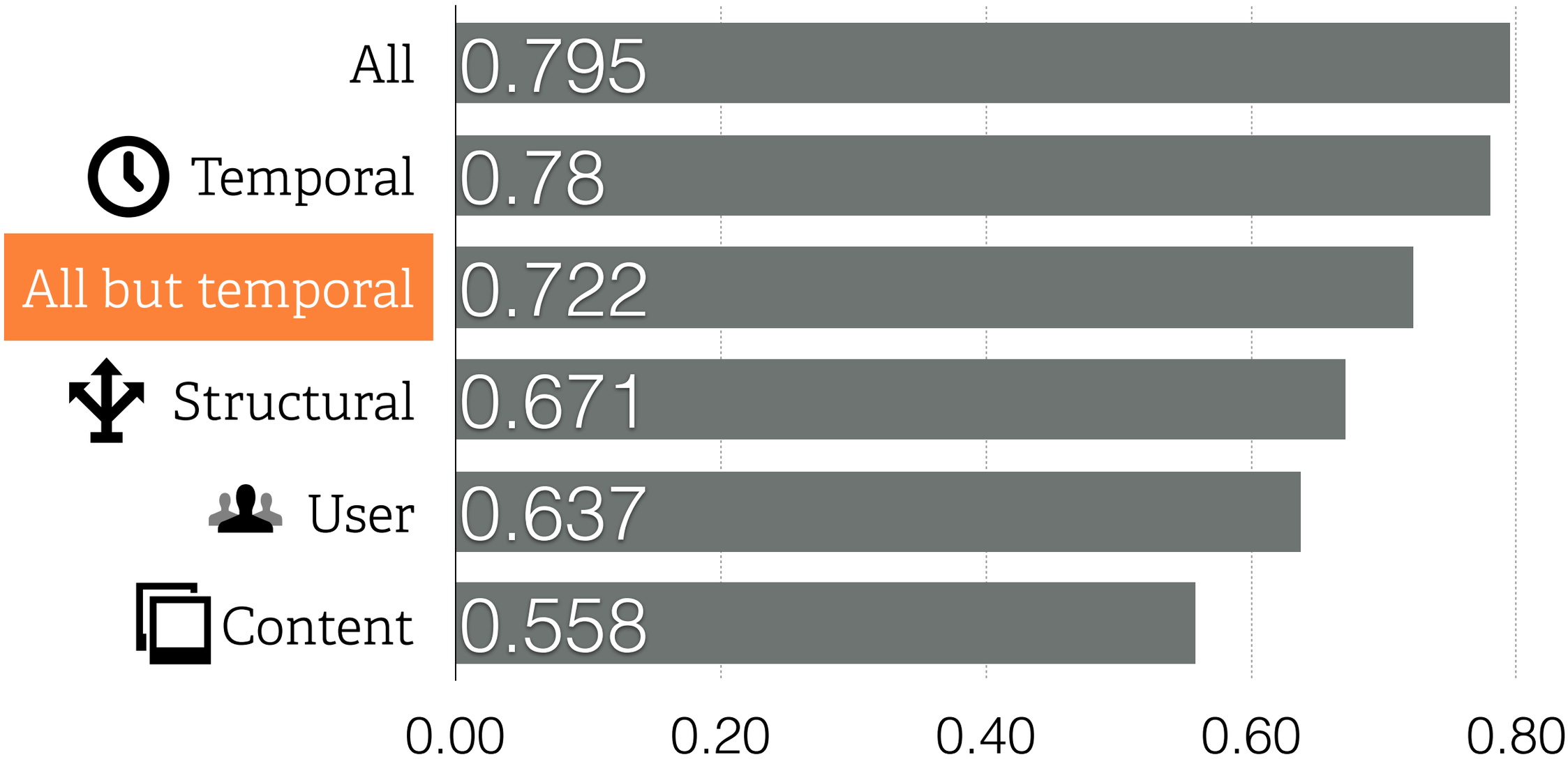
Temporal

e.g. time between re-shares

Structural

e.g. weiner index

Which features perform better for prediction task??

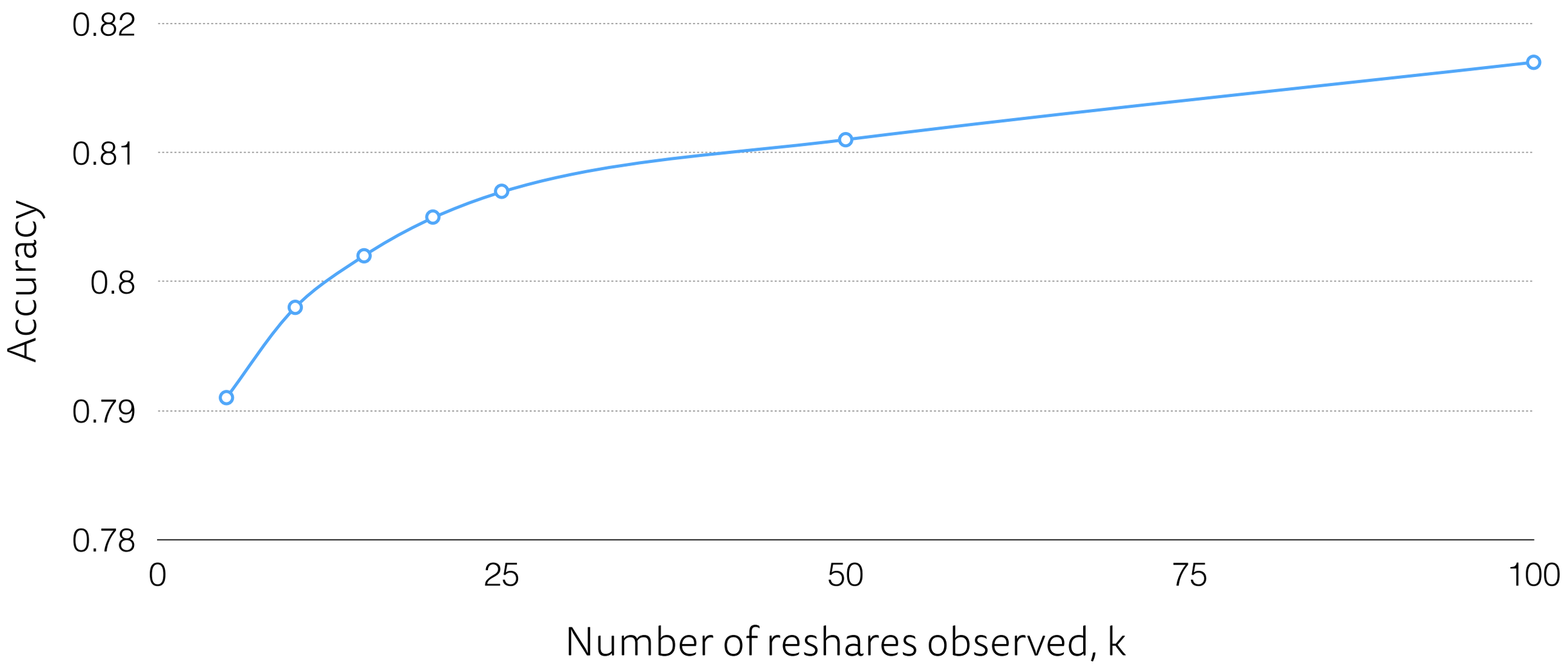


Logistic Regression

Accuracy for k=5

J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In WWW '14, 2014.

How does the performance change with k…???



J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In WWW '14, 2014.

Can we come up with a model for cascade prediction??

Cascade Generation Model

Susceptible Infected Susceptible

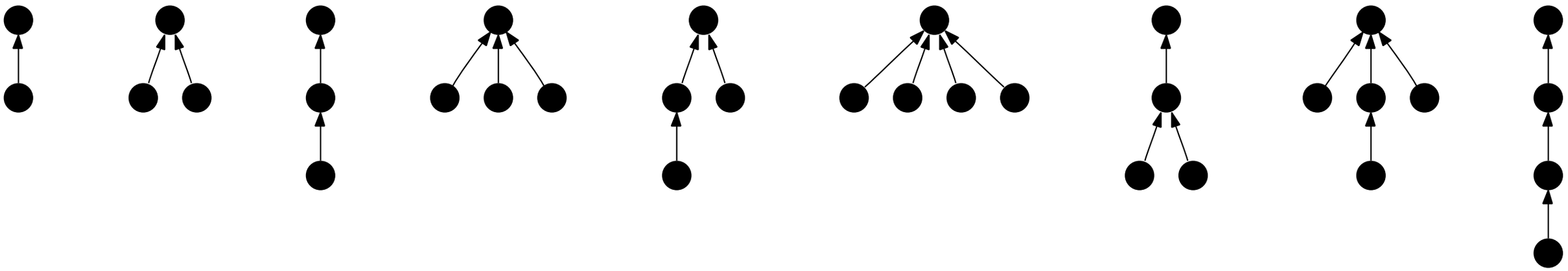
1. Uniformly at random pick a blog u, mark it as infected and add to the cascade graph

2. Infect each of its directed neighbours with probability β

3. Add the newly infected nodes $\{v_1, v_2, \ldots, v_n\}$ to the cascade

4. Set the state of node u as not infected or susceptible. Continue recursively from step 2, until no nodes are infected.

Frequent Cascade shapes



Top 10 most frequent cascades generated by Cascade Generation Model

J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

SEISMIC:

A Self Excitation Model for Information Cascades

Goal

Given tweet and retweets upto time "T", can we predict its final popularity…?

Defining Intensity for SEISMIC

Poisson Process: $\lambda_t = \lambda$

In SEISMIC

Infectiousness: "probability" of retweeting

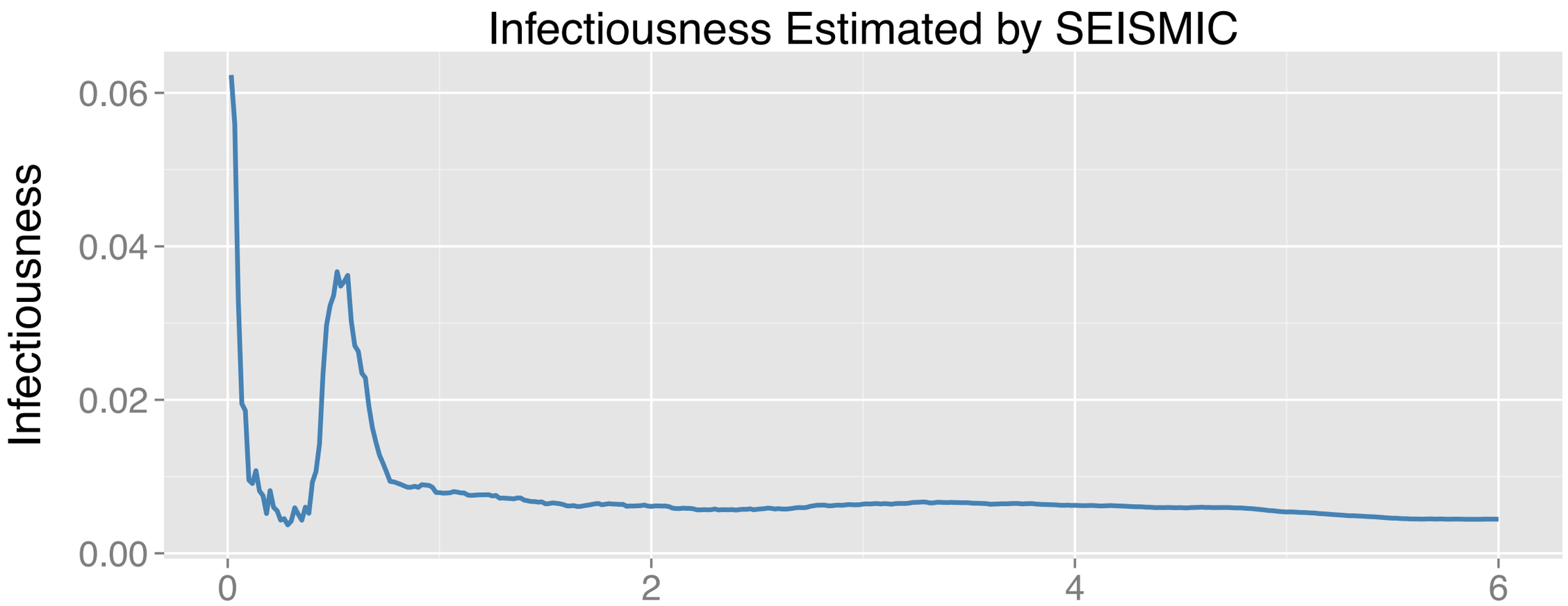
$$\lambda_t = p_t \cdot \sum_{t_i \leq t} n_i \phi(t - t)$$

"Rate of viewing"

(Intensity of arrival of new newly exposed nodes)

Intuition for Infectiousness

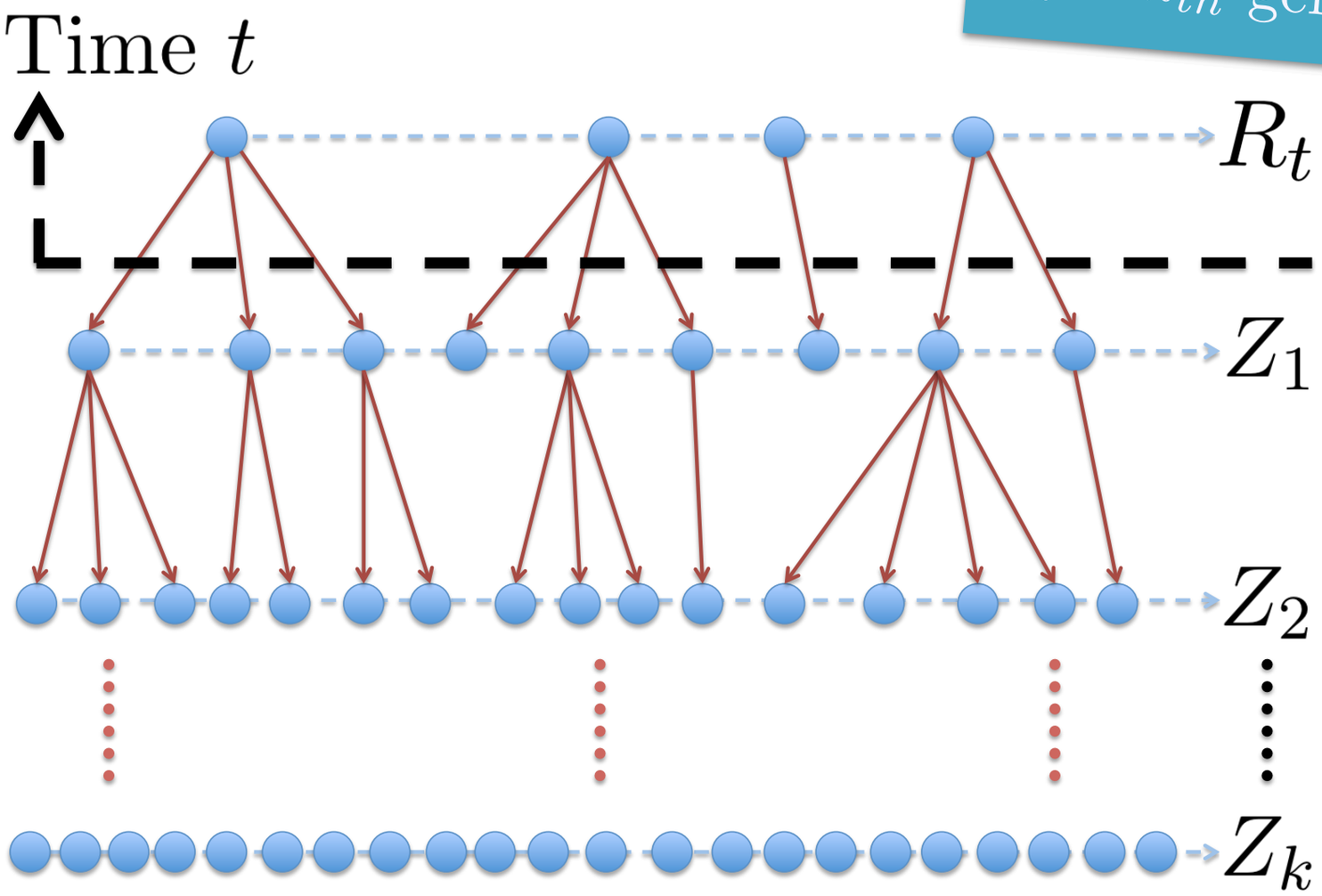
$$\hat{p}_t = \frac{\textbf{Number of Retweets}}{\textbf{Number of views}}$$



Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. KDD'15

Predicting final popularity

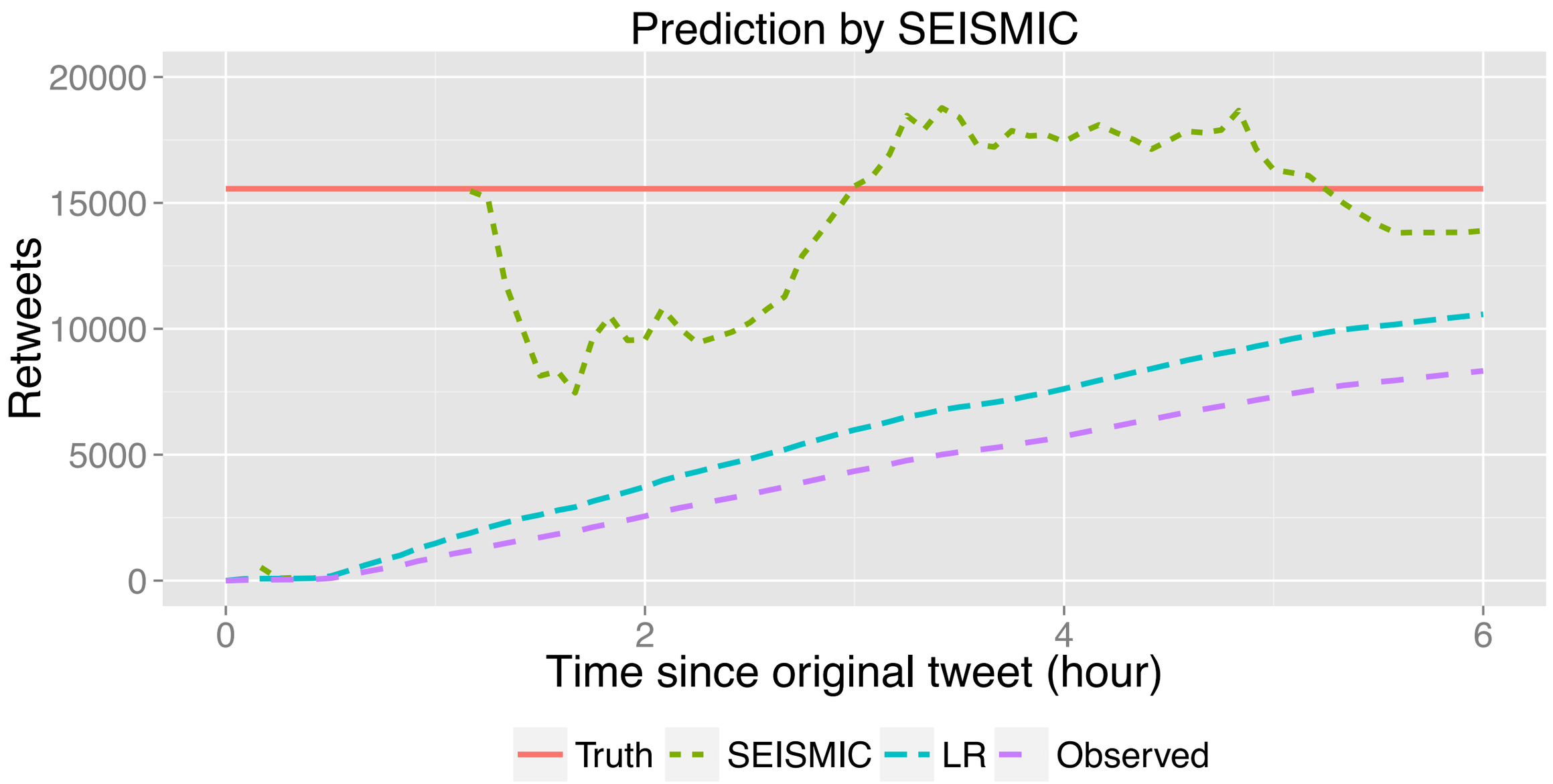
$Z_k = k_{th}$ generation descendants



Final popularity: $R_\infty = R_t + \sum_{k=1}^{\infty} Z_k$

Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. KDD'15

Prediction by SEISMIC



Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. KDD'15

Other Similar Problems!!!

Network Inference

Influence Maximization

Cascades Detection

Firefighting

References

J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M.Hurst. Cascading behavior in large blog graphs. In Proc. ICDM, 2007

H.W. Shen, D. Wang, C. Song, and A.-L. Barabási. Modeling and predicting popularity dynamics via reinforced poisson processes. arXiv:1401.0778, 2014.

J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In WWW ’14, 2014.

Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In Proceedings of the 20th international conference on World wide web, WWW ’11, 695–704. New York, NY, USA:ACM.

Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’15, pages 1513–1522, 2015.

Thank You!!!