Social Influence and the Diffusion of User-Created Content

Eytan Bakshy Brian Karrer Lada A. Adamic

University of Michigan

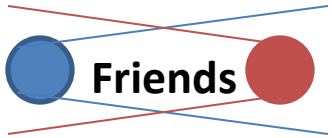
Motivation

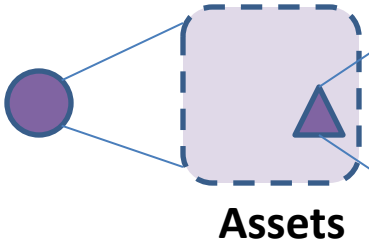
- Direct transfers between users
- More precise identification of heavy influencers
- Simple model of adoption rates
- Importance of network effects

Data Set

- content ownership data (3,409,630 entries)
    - name of content
    - previous owner
    - next owner
    - time stamp
- weekly snapshots of the social graph from Sept. 2008 to Jan. 2009
- user data
    - date joined
    - number of hours played

Definitions





Gesture

- at least 16 unique owners
- user-created only

Active user

- last login after Nov 2008
- play time > 60 hours
- exchanged at least 1 item with another user

Sample

- 100,229 users
- 106,499 assets

Observations and Further Definitions

48% of transfers occurred between friends

users share assets within short time of receiving

For each asset:

Asset size: total number of adopters

% direct: percentage of transfers between friends

% non-leaf: percentage of non-terminating transfers

social influence: a friend adopts before you do

direct influence: obtaining an asset from a friend

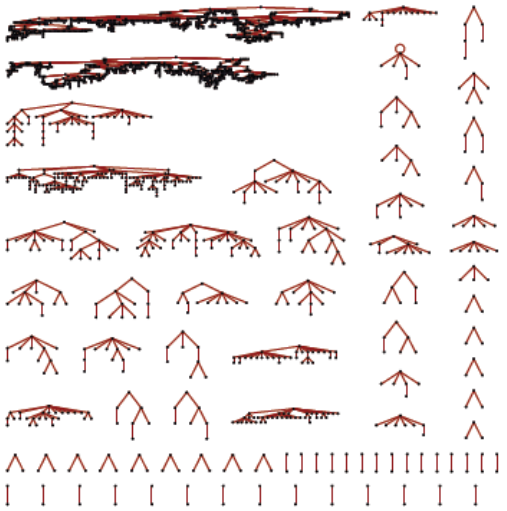
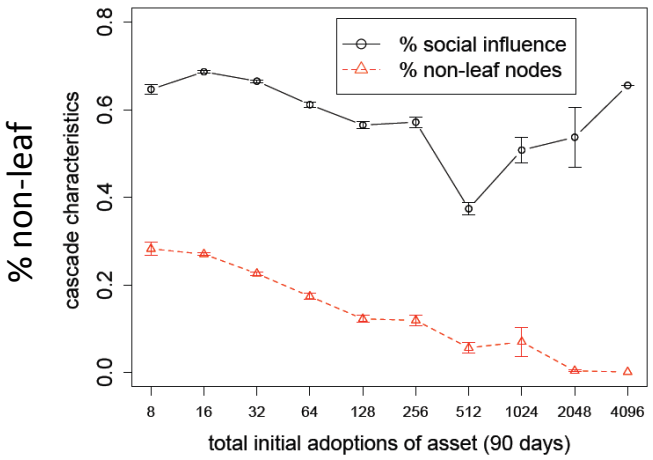


Figure 2: Example of a cascade forest for the Aerosmith(916) gesture. Edges denote transfers of the gesture between users.

Regression



- **asset size or popularity**: total number of adopters
- **% social:** adoptions occurred after at least one friend adopted
- **% non-leaf**: percentage of non-terminating transfers

↑ % social
↓ number of additional adoptions

speed of individual transfers ↑ between friends

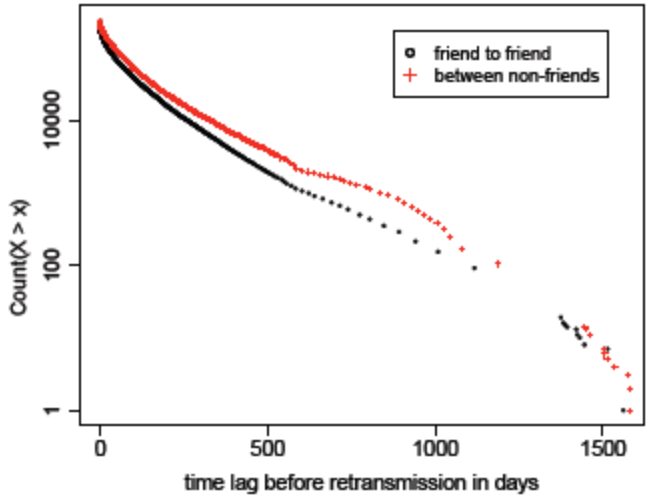
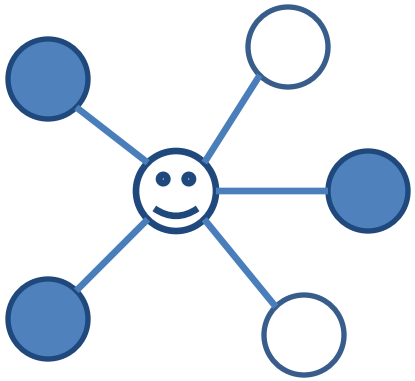


Figure 4: Delays between a users' adoption and re-transmission times, for assets with 100-200 adopters.

Rate of Adoption for an Asset



$state\ k$

k friends have adopted the asset

T_k exp dist. w/ mean $\frac{1}{\lambda_k}$

Time until adoption

Rate of adoption for state k

$$M_k = \sum_i t_k^i$$

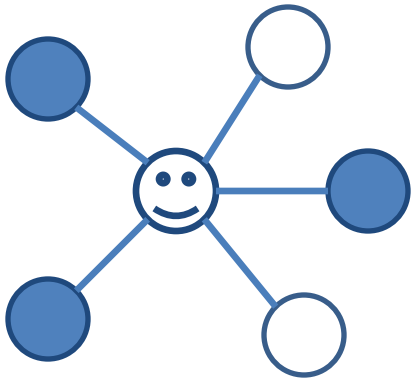
Total time spent in state k over all users i

A_k Number of users who adopted at state k

$$\lambda_k = \frac{A_k}{M_k}$$

Probability Density F(k) $$\prod_k \lambda_k^{A_k} e^{-\lambda_k M_k}$$

Rate of Adoption for an Asset



$state\ k$

k friends have adopted the asset

T_k exp dist. w/ mean $\frac{1}{\lambda_k}$

Time until adoption

Rate of adoption for state k

$$M_k = \sum_i t_k^i$$

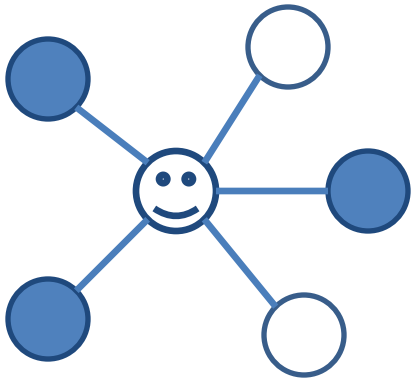
Total time spent in state k
over users who adopted the asset

A_k Number of users who adopted at state k

$$\lambda_k = \frac{A_k}{M_k}$$

Probability Density F(k) $\prod_k \lambda_k^{A_k} e^{-\lambda_k M_k}$

Rate of Adoption for a User



$state\ k$

k friends have an asset i

T_k exp dist. w/ mean $\frac{1}{\lambda_k}$

Time until adoption

Rate of adoption for state k

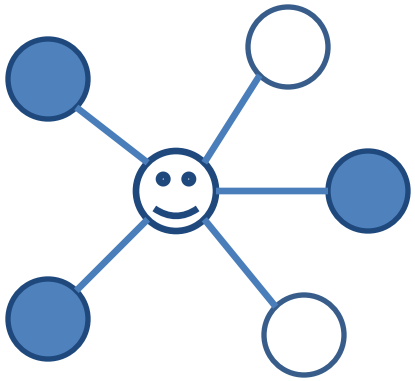
$M_k = \sum_i t_k^i$ Total time the user spent in state k for the asset i

A_k Number of assets adopted from state k

$$\lambda_k = \frac{A_k}{M_k}$$

Simplified Probability Density $\prod_k \lambda_k^{A_k} e^{-\lambda_k M_k}$

Rate of Adoption for a User



$state\ k$

k friends have an asset i

T_k exp dist. w/ mean $\frac{1}{\lambda_k}$

Time until adoption

Rate of adoption for state k

$$M_k = \sum_i t_k^i$$

Total time the user spent in state k
over all assets the user was observed to adopt

A_k Number of assets adopted from state k

$$\lambda_k = \frac{A_k}{M_k}$$

Simplified Probability Density $\prod_k \lambda_k^{A_k} e^{-\lambda_k M_k}$

Submodularity

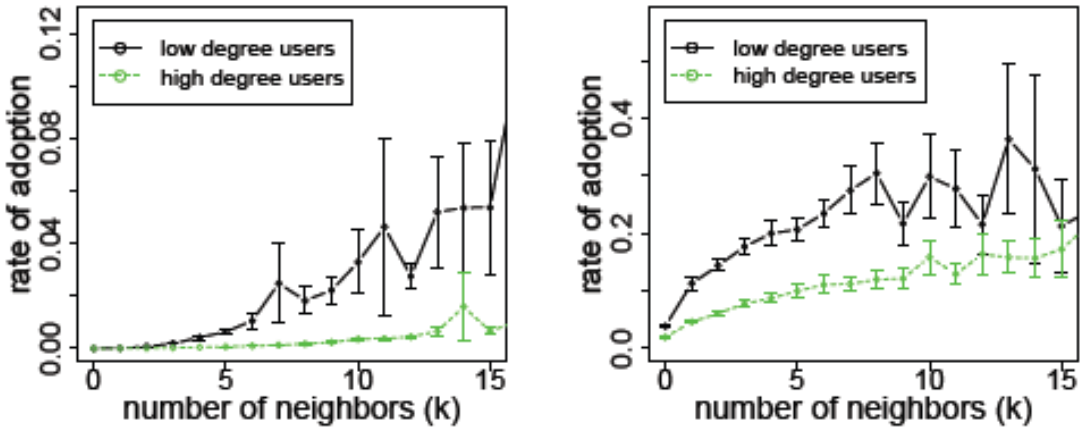


Figure 6: The average rate of adoption for users as function of adopting neighbors, k. The black curve corresponds to users of low degree that have 15-100 friends, and the green curve corresponds to users with 100-1000 friends. Left: entire population of assets. Right: adopted population of assets. The rates are in units of inverse days.

Comparison with Cox Model of Adoption

↑ # adopting friends

↑ chance of own adoption

↑ average degree

↓ adoption rate

↑ total # of adopters

↑↓ adoption rate

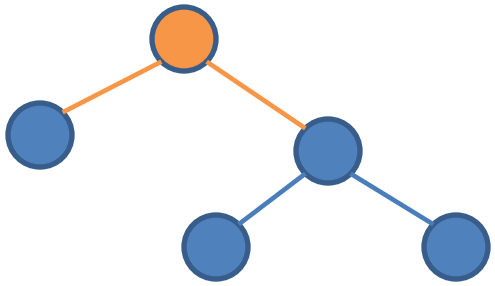
↑ user experience (time)

↓ adoption rate of friends' content

Influence

Measuring Entropy of Users Responsible for Transfers

Entropy = 1 bit



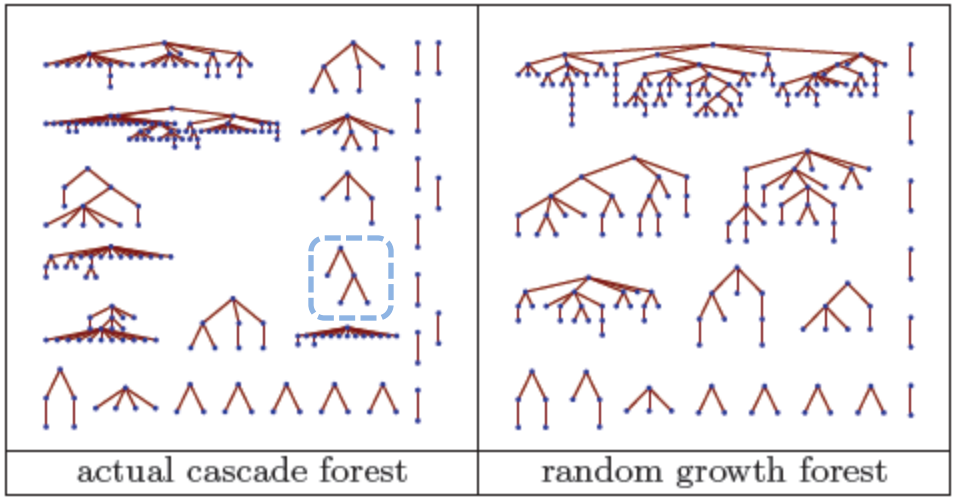


Figure 8: Comparison between actual growth of cascades and a null model where each previous adopter is equally likely to be sharing assets.

Our Model:
2.72 bits

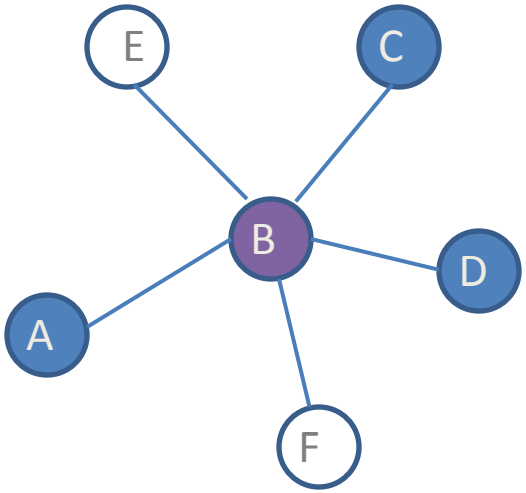
Null Model:
3.48 bits

Proportion of missing edges in the cascade

$$Null\ Model(n, p)$$

Total number of owners of the asset

Strength of Influence



Transfers : # of A to B transfers

Expected: E[A to B transfers]
+ Pr[A was the one who transferred asset to B]

$$\gamma = \frac{transfers - expected}{expected}$$

Mean: -0.286

↑ number of friends ↓ γ score

↑ number of assets ↓ γ score

↑ number of transfers to friends / assets owned ↑ γ score

– Earlier asset acquisition – γ score

Early Adopters vs. Later Adopters

>= 20 gestures or >= 40 gestures
Among first 5% or 10 %of adopters for all owned assets

>= 20 gestures
Latter half of adopters

68 days earlier ← Joined Second Life

↓ time spent playing ↑

8 ↓ number of friends

-0.08 Gamma scores -0.22

↑ number of transfers ↓

↑ Future popularity of assets adopted

↑ Pr[you adopt before your friends] ↓ Gamma score ↓ number of transfers

Conclusions

↑ rate of adoption ↑ friends adopt ↓ asset popularity

↑ friends ↓ likely to be influenced by any one of them

↑ friends ↓ one's own direct influence

↑ strength of friendship ↑ number of assets transferred

early adopters have no greater influence

Additional Equations