

# On the impact of social cost in opinion dynamics

Panagiotis Liakos

**Katia Papakonstantinou**

University of Athens

Algorithmic Game Theory Athens  
NTUA, July 20<sup>th</sup>, 2016

# Formation of opinions in a social context

intrinsic belief  
+  
friends' expressed  
opinions

expressed  
opinion



# Motivation

exponential growth of **online social networks**



ever-increasing amount of social activity **information** available



ability to **analyze** user behavior and **interpret** sociological phenomena  
at a large scale [AKM08]



**Investigating game theoretic models of networks against real data**

We consider the phenomenon of **opinion formation under social influence**. Given a network dataset, we want to be able to:

- verify the existence of influence among users
- build a model that describes user behavior in the network.

## Our contributions:

- ① We analyze user activity in **digg** and verify that social interaction results in influence on opinions among the participants.
- ② We initialize a sociological model using real data. Based on the Game Theory framework, we experimentally show that the *repeated averaging* process results to *Nash equilibria* which are illustrative of how users really behave.

# What is digg?

**digg** is a news aggregator with a curated front page, aiming to select stories specifically for the Internet audience such as science, trending political issues, and viral Internet issues.

The screenshot shows the Digg website interface. At the top, there's a navigation bar with 'digg' logo, 'My Profile', 'Friends' Activity 47', 'Submit New', 'Logout', and a search bar. Below that is a secondary navigation bar with categories: 'All', 'News', 'Videos', 'Images', 'Podcasts', and 'Customize'. A third bar shows sub-categories: 'Technology', 'World & Business', 'Science', 'Gaming', 'Lifestyle', 'Entertainment', 'Sports', and 'Offbeat'. The main content area is titled 'News, Videos, Images' and features a list of items. Each item includes a digg count, a title, a source, a brief description, and a 'digg it' button with options for comments, share, and bury. The items listed are:

- 130 diggs**: [\\$500 to \\$100,000: The Biggest Dollar Bills in History \[PICS\]](#) from purpleink.com. Description: This is a collection of all the unknown big dollar bills that America has had in the past. **3 min ago**.
- 109 diggs**: [Obama leads pledged delegate count by 136 delegates](#) from my.barackobama.com. Description: Obama currently leads the pledged delegate count by 136 delegates - 1139 for Obama - 1003 for Clinton. After sweeping the past eight contests by large margins, it now appears that Senator Obama has an insurmountable lead in pledged delegates. **3 min ago**.
- 94 diggs**: [Kung Fu Panda : New Trailer](#) from kungfupanda.com. Description: VIDEO - kungfupanda.com (Comics & Animation) made popular **3 min ago**.
- 380 diggs**: [Drupal 6.0 has been RELEASED - download now](#) from drupal.org. Description: Version 6.0 of Drupal, popular social publishing platform for countless web sites, has been released. This version of Drupal is easier to use for administrators, themers and developers alike. Check out the many great new features such as an improved installer, OpenID support, drag-n-drop, actions and triggers, and better internationalization. **13 min ago**.
- 725 diggs**: ["Imagine No Religion" Billboards Spark National Controversy](#) from imagine-no-religion.com. Description: The national campaign is an effort to let Americans know that there is room for reason and clarity of thought, free from the dogma that organized religion uses to keep its flock in line, as well as donating. **43 min ago**.
- 409 diggs**: [Farthest Galaxy Found, Perhaps](#) from livescience.com. Description: Images taken with the Hubble Space Telescope have revealed the galaxy at almost 13 billion light-years away, making it the strongest candidate for the most distant galaxy ever seen. **53 min ago**.

On the right side, there's a 'Top in All Topics' sidebar with a list of trending items:

- 4466 diggs**: You want the blocks? Best CraigsList Ad Of All Time Ever.
- 3313 diggs**: I've Never Been More Ashamed to be an American.
- 2409 diggs**: Sometimes You Get the Trees, Sometimes They Get You [PICS]
- 1796 diggs**: Mozilla Firefox 3 beta 3 released today!
- 1712 diggs**: Hillary Lost This White Lady's Vote Today.
- 1598 diggs**: Zero Punctuation: Sim City Societies
- 1408 diggs**: jesus gets struck down
- 1310 diggs**: How Google Got Its Colorful Logo
- 1092 diggs**: Men arrested in murder plot of Muhammad drawing cartoonist

- Submit stories.
- Digg (give a thumbs-up/positive vote to) a story you want other people to see.
- Follow users you consider interesting to get informed about their diggs in your news feed.

The **digg** dataset is appropriate for our study because:

- **digg** was very popular at the time the dataset was collected [LGS12]
- **digging** a story has a sense of opinion expression and an urge to influence
- both **diggs** and **follower links** are timestamped

## We use:

a variation of the **DeGroot** model due to Friedkin and Johnsen [FJ90] and the corresponding *game* of [BKO11].

### Each user $i$ maintains:



An *intrinsic* belief  $s_i$   
Remains constant



An *expressed* opinion  $z_i$   
Updated iteratively through averaging

### The cost a user suffers emanates from:



*Suppressing* her intrinsic belief



*Disagreeing* with her friends

## Repeated Averaging

At each time step user  $i$  updates  $z_i$  to *minimize* her cost:

$$z_i = \frac{s_i + \sum_{j \in N(i)} w_{ij} z_j}{1 + \sum_{j \in N(i)} w_{ij}}$$

$N(i)$ : the set of nodes that  $i$  follows

$w_{ij}$ : the *strength* of the influence of  $j$  on  $i$

The averaging process terminates when  $z$  converges to the unique Nash equilibrium, where the social cost is minimized.

## Our intuition:

The influence of  $j$  on  $i$  regarding a specific matter depends on:

The impact  $a_{ij}$  of  $j$  on  $i$       The expertise  $b_j$  of  $j$   
Does  $i$  respect  $j$ 's opinion in general?      Is  $j$  authoritative on this matter?

## We define:

$$w_{ij} = a_{ij}b_j$$

Top-20 Cascade Patterns							
1		6		11		16	
2		7		12		17	
3		8		13		18	
4		9		14		19	
5		10		15		20	

## Empirical Analysis of **digg** (cont.)

the total *infections* were  
less than the initial  
*seeders* in more than 92%  
of the stories of digg

## Empirical Analysis of **digg** (cont.)

the total *infections* were less than the initial *seeders* in more than 92% of the stories of digg

a few users caused many cascades while most were unable to cause any

## Empirical Analysis of **digg** (cont.)

the total *infections* were less than the initial *seeders* in more than 92% of the stories of digg

a few users caused many cascades while most were unable to cause any

even the most *authoritative* users were not effective in all stories

# Our Experiments and Assumptions

- We performed repeated averaging in our model for *stories* of **digg** until the opinions, expressed by votes, converged to the unique Nash equilibrium.
- We compared against predictions obtained using a *Neural Network* classifier.

## Model initialization assumptions

**Intrinsic belief**  $s_i$

$$s_i = \begin{cases} 1 & \text{if } i \text{ voted a story before } \textit{any} \text{ user she follows} \\ 0 & \text{otherwise} \end{cases}$$

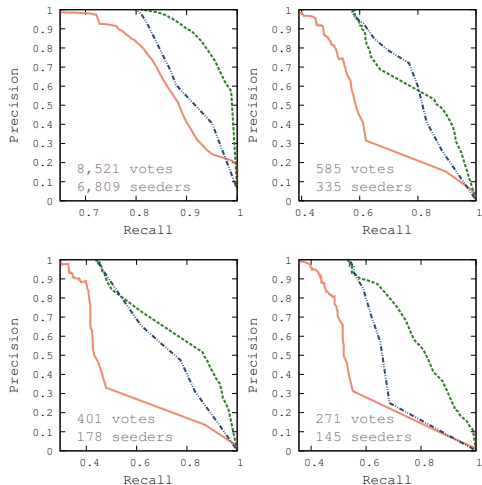
**Influential strength**  $w_{ij}$

Two variants:

$$\begin{aligned} \text{i) } & a_{ij} = b_j = 1 \Rightarrow w_{ij} = 1 \\ \text{ii) } & a_{ij} = \frac{\# \text{ times } i \text{ is influenced by } j}{\# \text{ votes of } j} \\ & b_j = \frac{\# \text{ users influenced by } j \text{ in this story}}{\# \text{ followers of } j} \end{aligned}$$

# Results

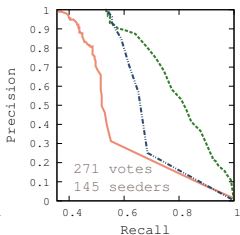
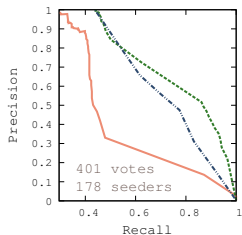
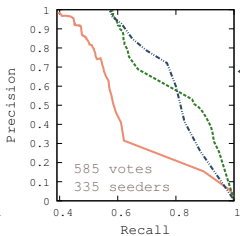
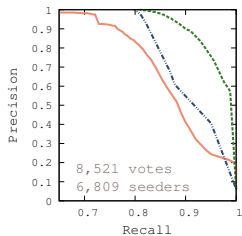
The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:



Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) .....  
Neural Network -.-.-.-

# Results

The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:

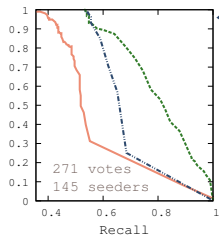
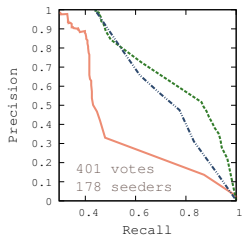
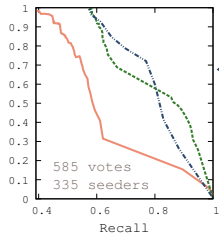
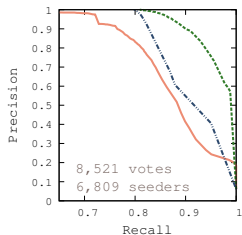


our 2nd variant closely mimics the original social activity

Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) .....  
Neural Network -.-.-.-

# Results

The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:



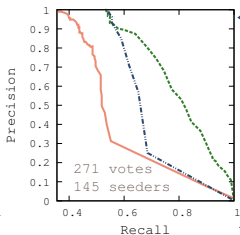
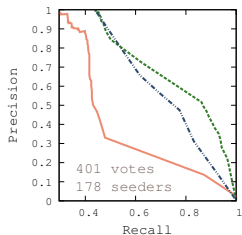
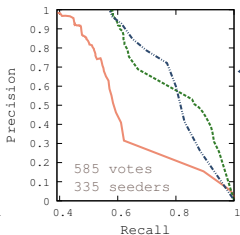
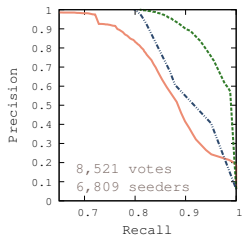
Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) - - -  
Neural Network - · - · -

our 2nd variant  
closely mimics the  
original social activity

the simplistic  
variant behaves  
poorly as expected

# Results

The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:



Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) .....  
Neural Network -.-.-.-

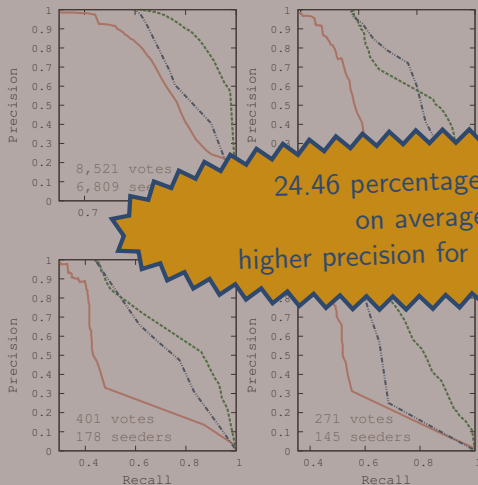
our 2nd variant  
closely mimics the  
original social activity

the simplistic  
variant behaves  
poorly as expected

we consistently out-  
perform the Neural  
Network classifier

# Results

The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:



our 2nd variant  
closely mimics the  
original social activity

24.46 percentage points,  
on average,  
higher precision for 90% recall!

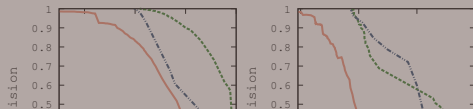
stochastic  
behaves  
poorly as expected

we consistently out-  
perform the Neural  
Network classifier

Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) - - - -  
Neural Network ·····

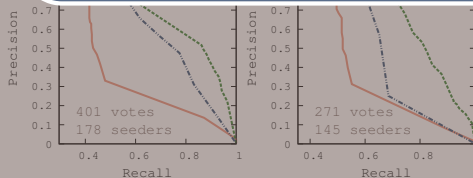
# Conclusion

The fraction of *predicted* votes that were actually *casted* against the fraction of *casted* votes that are *predicted*:



our 2nd variant closely mimics the original social activity

The repeated averaging process, combined with proper influence weights, results to Nash equilibria which are illustrative of how users really behave.



we consistently outperform the Neural Network classifier

Repeated Averaging ( $w_{ij}=1$ ) ———  
Repeated Averaging ( $w_{ij}=a_{ij}b_j$ ) - - - -  
Neural Network - · - · -

# References

- [AKM08] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian, *Influence and correlation in social networks*, SIGKDD (Las Vegas, Nevada, USA), 2008, pp. 7–15.
- [BKO11] David Bindel, Jon M. Kleinberg, and Sigal Oren, *How bad is forming your own opinion?*, FOCS, 2011, pp. 57–66.
- [FJ90] N.E. Friedkin and E.C. Johnsen, *Social influence and opinions*, Journal of Mathematical Sociology **15** (1990), no. 3-4, 193–206.
- [LGS12] Kristina Lerman, Rumi Ghosh, and Tawan Surachawala, *Social contagion: An empirical study of information spread on digg and twitter follower graphs*, arXiv preprint arXiv:1202.3162 (2012).

- We can estimate the price of anarchy in such a network, by comparing the social cost of the Nash equilibrium we computed here with the cost of the optimal setting of this network.
- We can investigate approaches that may reduce the cost of such networks.

# thank you!

for further details visit:

<http://hive.di.uoa.gr/network-analysis/>

or email me at: [katia@di.uoa.gr](mailto:katia@di.uoa.gr)