The Influence of Feedback with Different Opinions on Continued User Participation in Online Newsgroups

Teng Wang, Keith C. Wang, Fredrik Erlandsson,
S. Felix Wu, Robert Faris

University of California, Davis
tewang@ucdavis.edu

Davis Social Links

Introduction

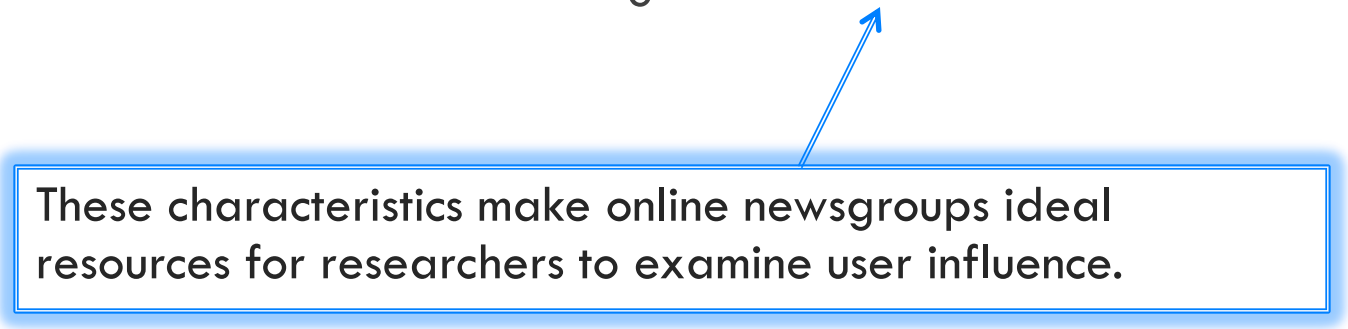
- Online Groups
  - Popularity
    - Over 50% of online users stay in contact with an online group regularly.
    - Up to Feb. 2010, 620 million groups on Facebook. (Nick, 2010)
  - Impact on people lifestyle
    - Information retrieval
    - Communication

Introduction

- Online Newsgroups – a special case

Introduction

- Online Newsgroups – a special case
  - Threaded discussion on different topics
  - Group membership is fairly loose.
    - Less of social component
    - Members are relative strangers with few offline communication



Introduction

- One Important issue – Member retention and continued participation
  - The success of online newsgroups depends on user contribution to provide value to their services.

> In a sample of 578 Usenet newsgroups, only 11.5% of people who posted in one month returned to post in a second month. (Jones et al. 2004)

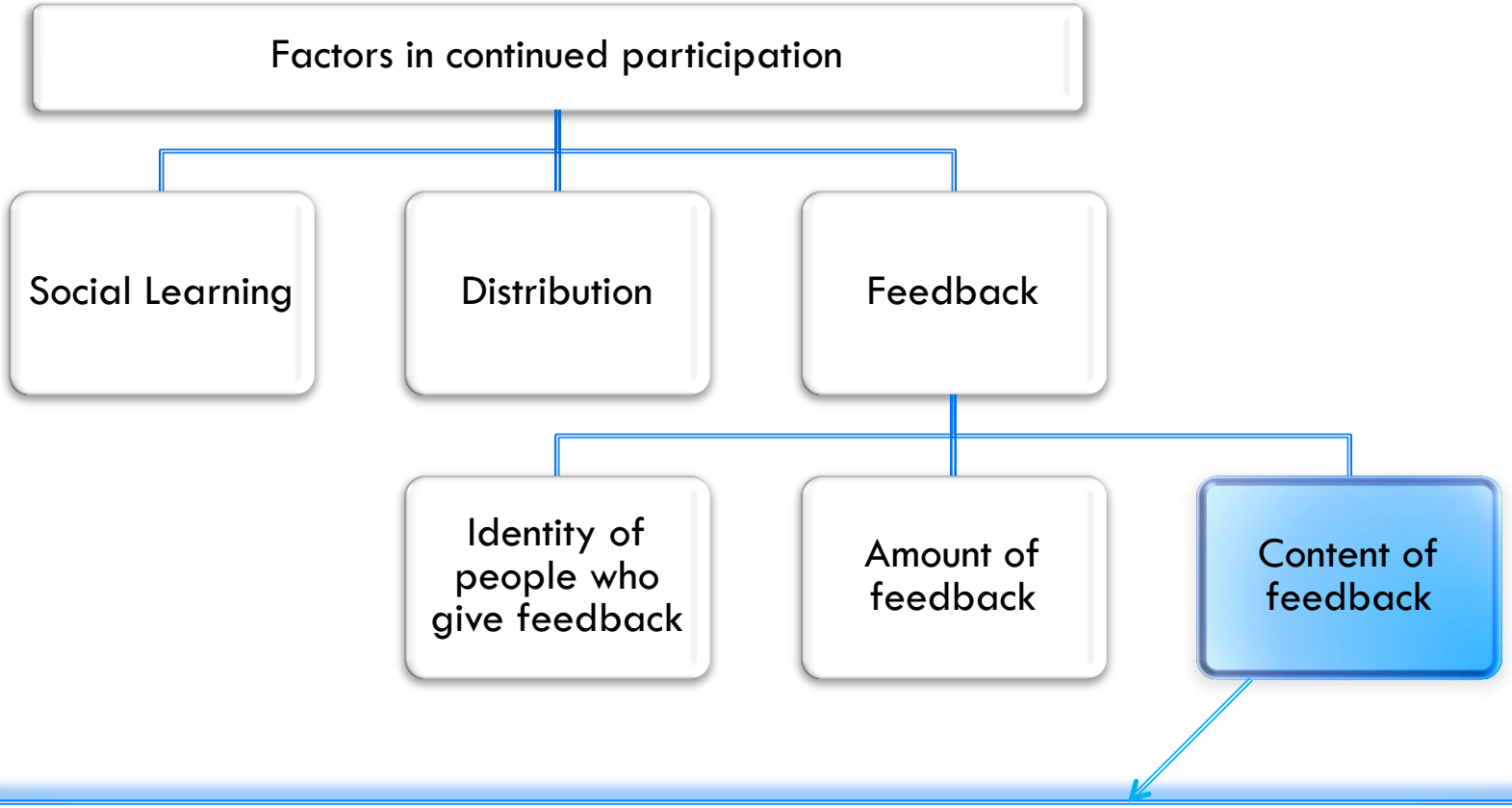
- Research Topic

Examine factors that influence user continued participation in online newsgroups.

Outline

- ~~Introduction~~
- Research Background
- Classify Different Opinions
- Influence of Feedback with Different Opinions on User Continued Participation

Research Background



Analyze the influence of feedback with different opinions on user continued participation in online newsgroups

Challenges

Analyze the influence of feedback with different opinions on user continued participation in online newsgroups

- Classify the content of feedback into different opinions.
  - Graph partition
  - Linguistic Analysis
- Analyze the influence of different opinions on user continued participation.
  - Multiple regression analysis

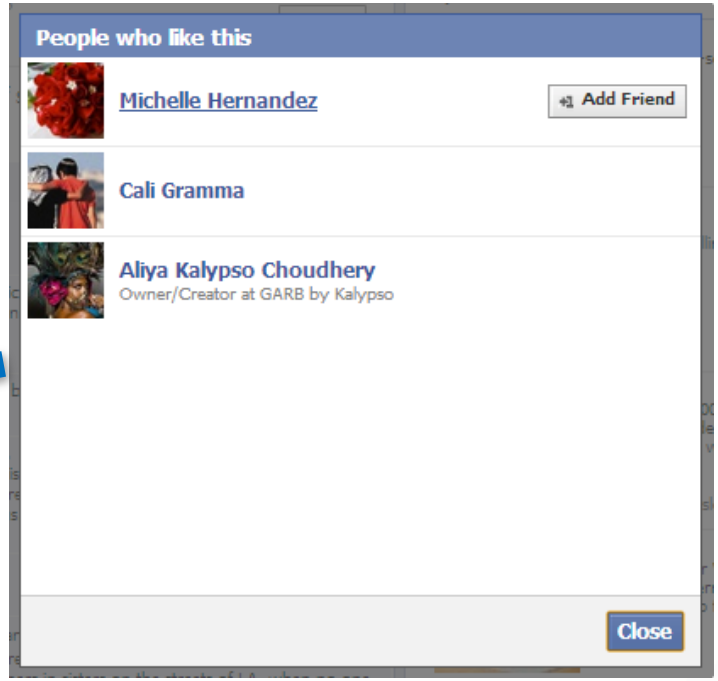
Outline

- ~~Introduction~~
- ~~Research Background~~
- Classify Different Opinions
- Influence of Feedback with Different Opinions on User Continued Participation

Classify Different Opinions

- Opinion Classification Method Based on User-Like graph
- Evaluation from Linguistic Features

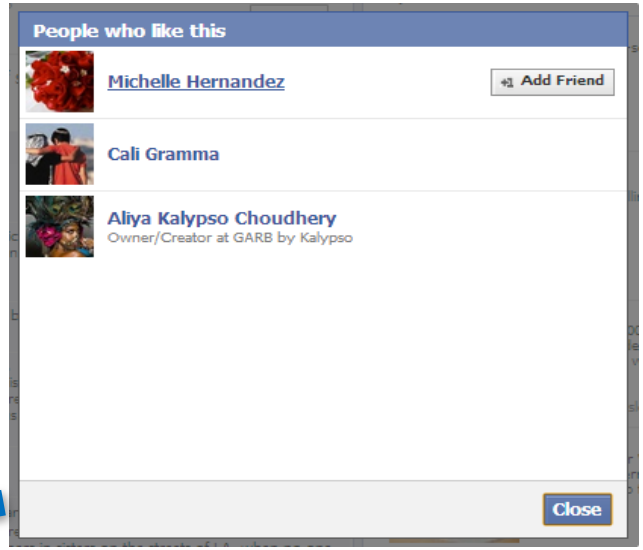
Posts in Facebook Public newsgroups

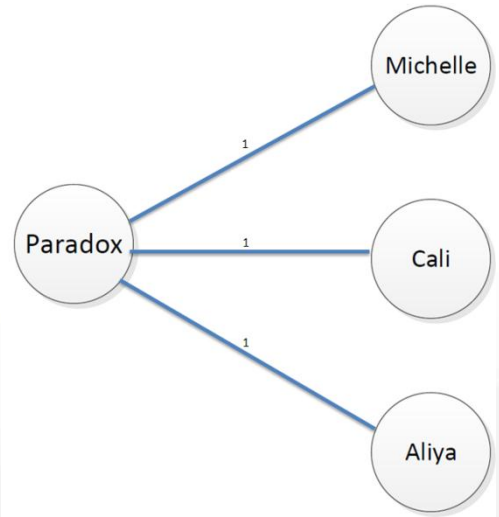


User-Like Graph

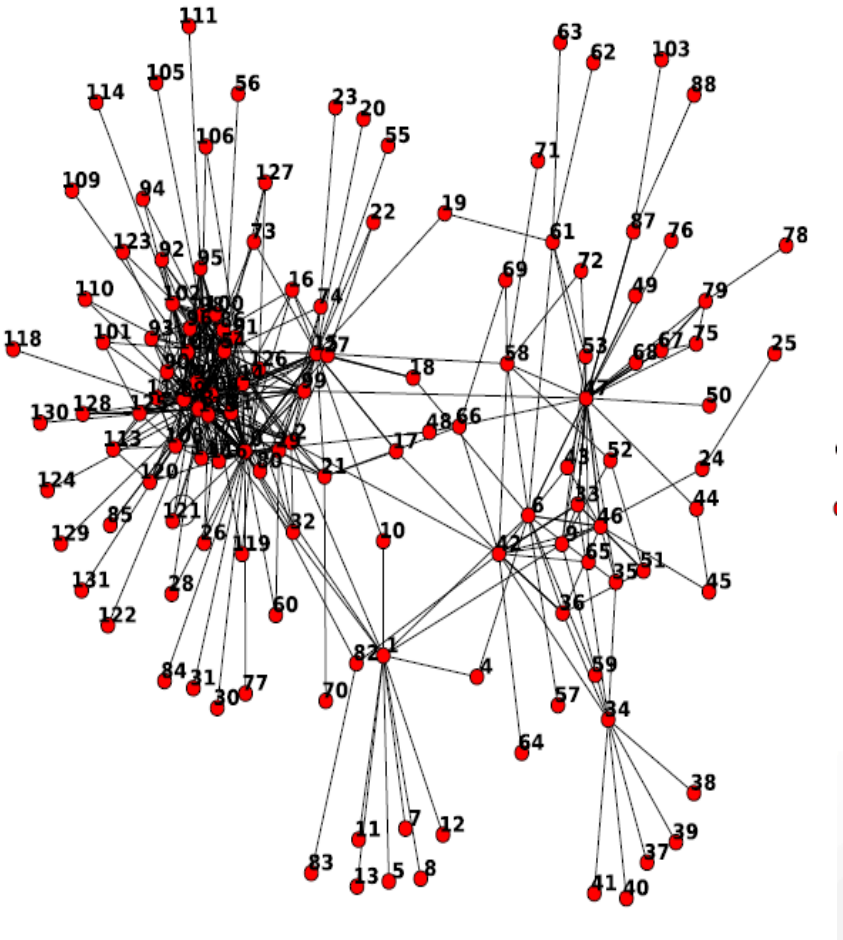
- Undirected Graph $G = (V, E)$
  - V is a set of nodes which stand for users who have clicked like button on any comments or whose comments were liked by others
  - E is a set of edges among V. An edge e stands for a connection between the author of a comment and the people who clicked *like* on it.

User-Like Graph





Examples of User-Like Graph



http://www.facebook.com/occupyLA/posts/349187595101578

Examples of User-Like Graph

Occupy Los Angeles · 55,553 like this
November 9, 2011 at 7:37pm ·

Like

MUST BE SEEN: In what world is this an acceptable response to peaceful protest? Berkeley Police clearly on the offensive against unarmed students protesting at Sproul Plaza. At least one instructor in urgent care. ~HSL

Occupy Cal 11/9/11
www.youtube.com
Berkeley police beating Cal students before retreating;)

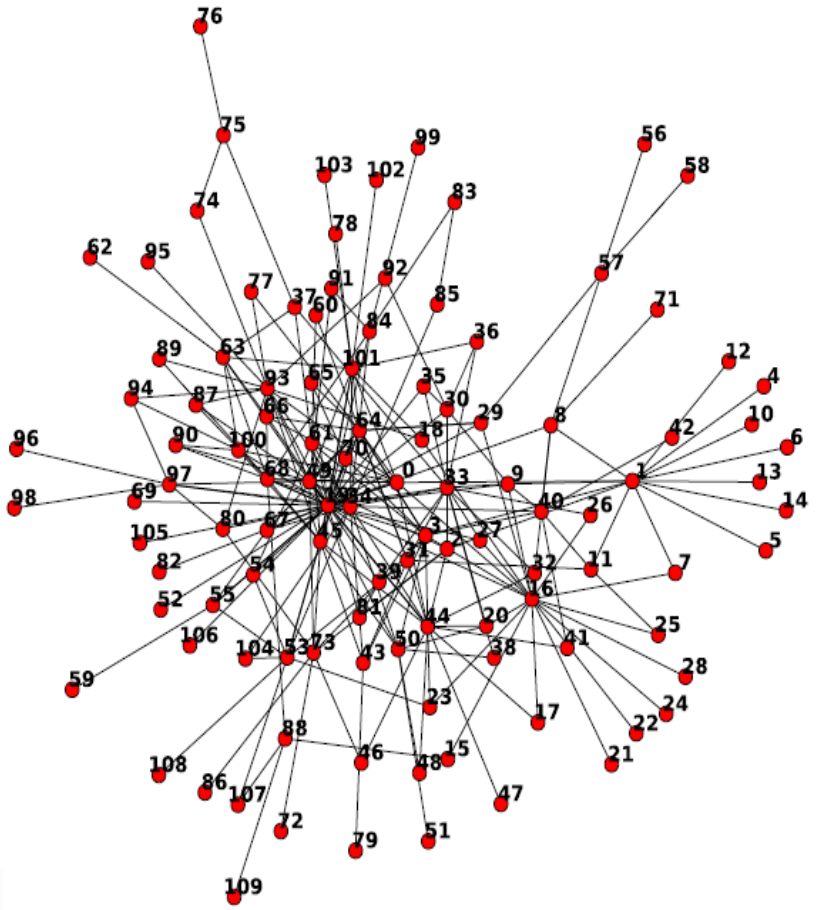
Like · Comment · Share 234

167 people like this.

View previous comments 38 of 126

Protester Alpha This is disgusting. I am posting an article on this at TheWallStreetProtest.com. This site gets thousands of visitors, so I'll help this video go viral. This is bullshit.
November 9, 2011 at 9:46pm · Like · 1

Chris Arroyo Honestly someone needs to die in order for there to be change. In History every revolution has had bloodshed occur. Peaceful protesting although it is taking the higher road and showing how calm and collect you are compared to them obviously does not g... See More
November 9, 2011 at 9:50pm · Like · 1



http://www.facebook.com/occupyLA/posts/137019263068274

Opinion Classification Method

- Objective:

  Find a partition of the graph $G = (V, E)$ such that edges between groups have a very low weight and edges within one group have high weight.

- Spectral Clustering Algorithm (Andrew Y. Ng et al., 2002)

  For a given number k of subset, we use spectral clustering algorithm to solve this RatioCut problem.

Choosing Cluster Number

- For simplicity, we only choose the cluster number from 1 to 2.
- General steps
  - Cluster the user-like graph with $k = 2$ by spectral clustering algorithm.
  - Use Newman Modularity Value Q (M.E.J. Newman et al., 2004) to check the quality of our partitioning result.
  - Decide the number of clusters

Classification Results

Occupy Los Angeles · 55,596 like this
December 31, 2011 at 6:20pm ·

Like

Obama signs NDAA 2012, but says,

"I have signed this bill despite having serious reservations with certain provisions that regulate the detention, interrogation, and prosecution of suspected terrorists," Obama said. "I want to clarify that my Administration will not authorize the indefinite military detention without trial of American citizens. Indeed, I believe that doing so would break with our most important traditions and values as a Nation."

Also notable:

"The Obama administration had expressed concerns about the Iran sanctions, which the White House feared could backfire and limit its ability to persuade other countries to join the United States in multilateral sanctions by forcing Iran to drive up oil prices.

Congress revised the bill to give the administration six months to apply the sanctions if the White House determines they could disrupt the oil markets."

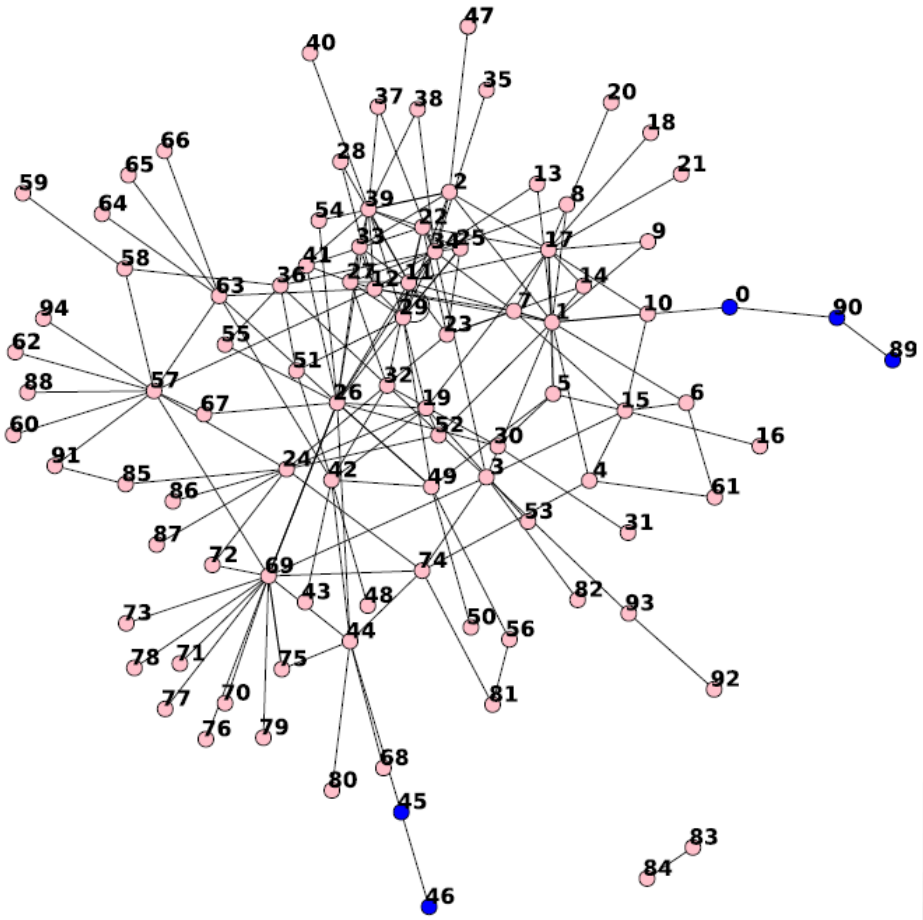
What road will this take us down?

~nb

Obama signs defense bill, pledges to maintain legal rights of U.S. citizens
www.washingtonpost.com
President signs legislation but expresses misgivings about several provisions.

Like · Comment · Share 25

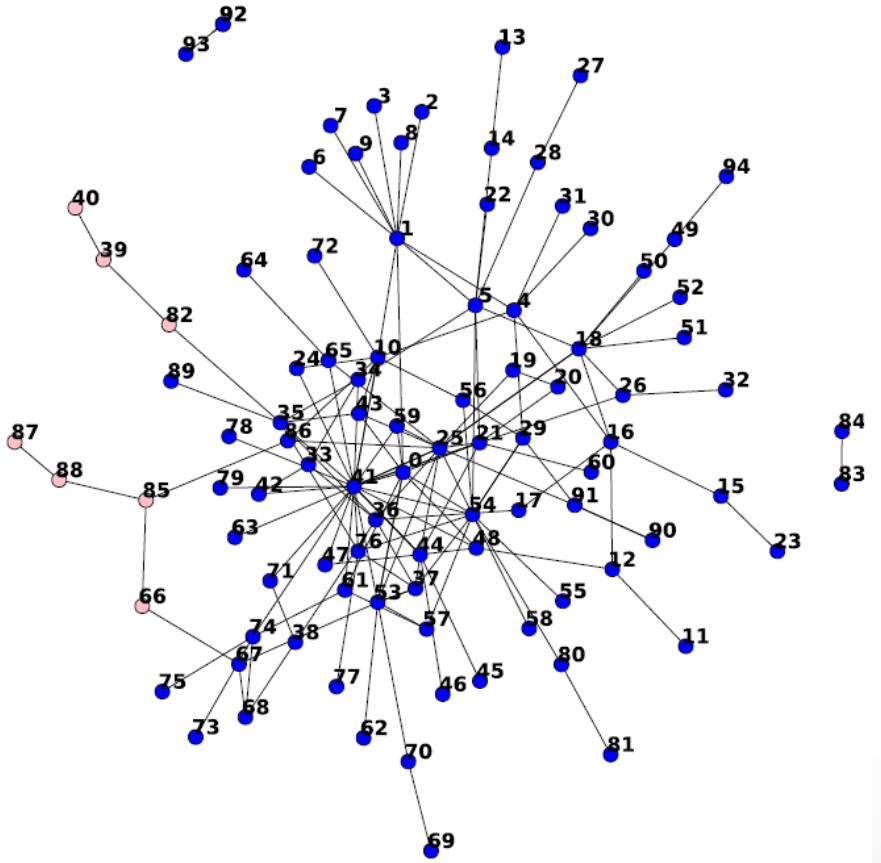
37 people like this.



(a) Q = 0.03. Consider it as one cluster

http://www.facebook.com/occupyLA/posts/336031626407405

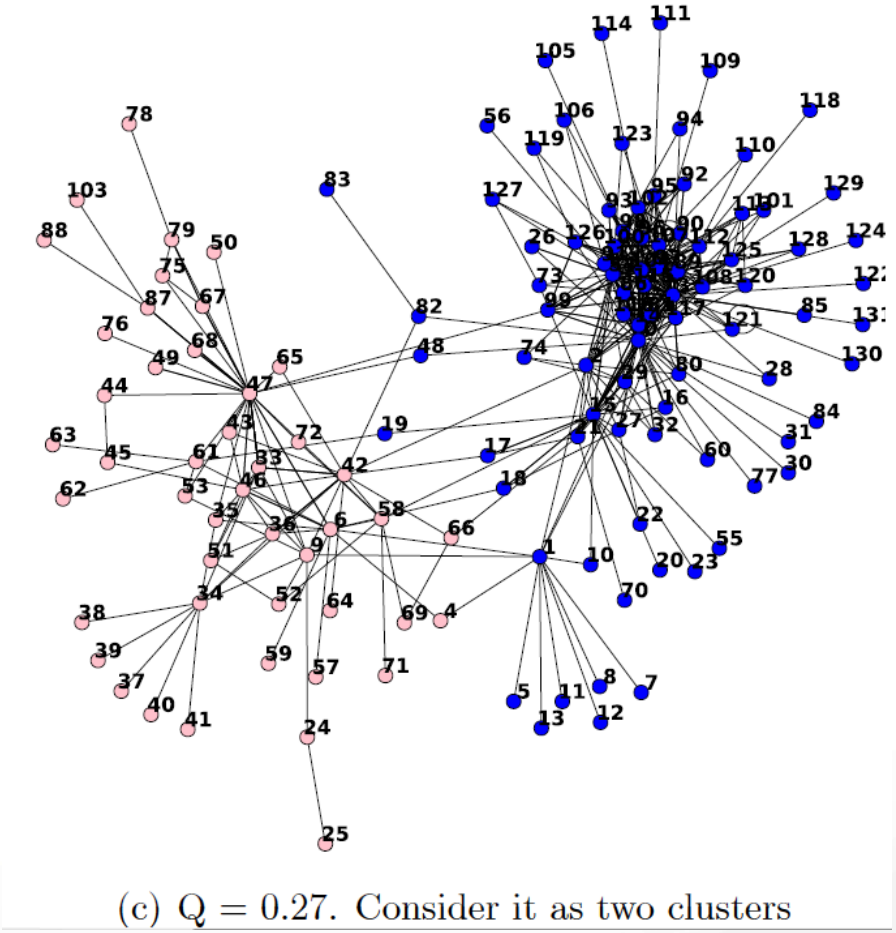
Classification Results



(b) Q = 0.06. Consider it as one cluster

http://www.facebook.com/occupyLA/posts/137019263068274

Classification Results



(c) Q = 0.27. Consider it as two clusters

http://www.facebook.com/occupyLA/posts/349187595101578

Classification Results

Occupy Los Angeles · 55,596 like this
October 10, 2011 at 1:42pm ·

Like

Heard some of the LAPD dropped off 3 crates of supplies: hygienic products, snacks, sunscreen,etc. for Occupiers! -M.K.

Like · Comment · Share 99

676 people like this.

View previous comments 43 of 103

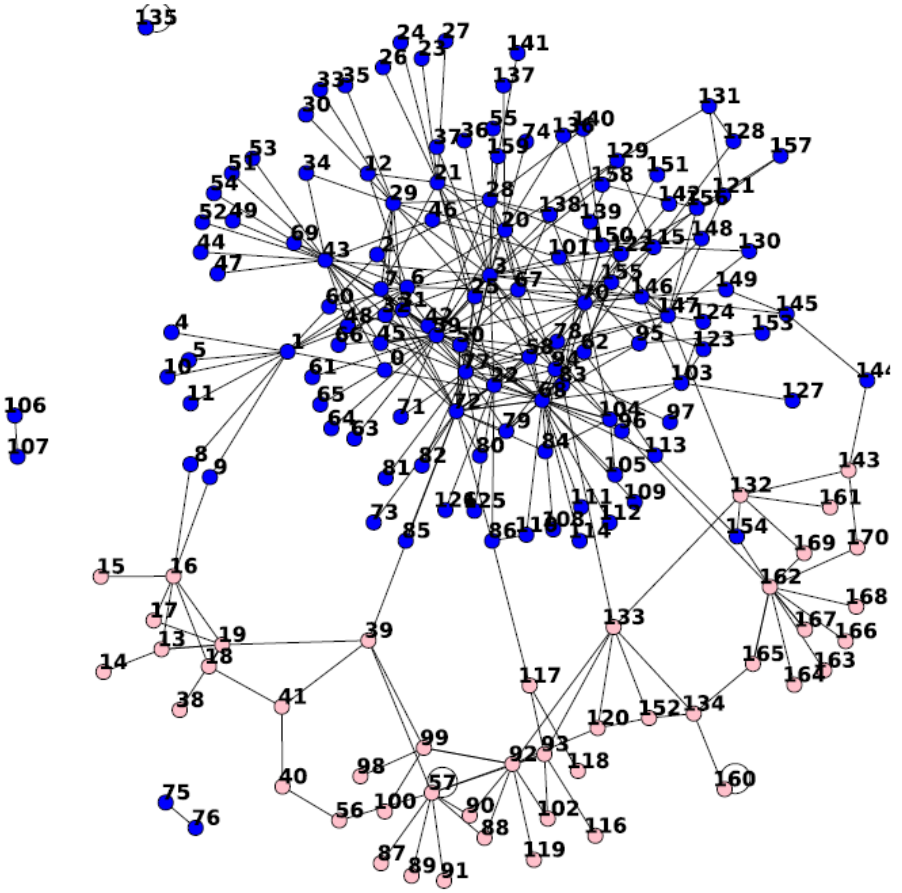
Simone Maritza Missirian Very cautiously optimistic...in the end, we need to win their hearts and minds...'cause they have guns and our movement will grow stronger faster if they don't use them against us.
October 10, 2011 at 2:14pm via mobile · Like · 3

David Sweet awesome - send that tweet out ask to be re-tweeted..!!
October 10, 2011 at 2:16pm via mobile · Like

Paradox Pollack They are looking for photo opps.
Unlike the NYPD who have a vested interest in not pissing off Wall Street, the LAPD and Mayor Villaraigosa want to support this because it is good press and our #'s are fewer. They want to sty on our goo side and not have to pay cops to wrangle us. It's simply good politics.
October 10, 2011 at 2:18pm via mobile · Like · 3

Daniel Rivas Thank you LAPD, we needed that
October 10, 2011 at 2:19pm via mobile · Like · 4

Paulina Selene Jones love ♥ let us hope they change their attitudes towards the people in all of LA. it is awesome and a beautiful thing that they are giving so much support to OLA. But we can't forget the cruel and unusual crimes against our brothers in sisters on the streets of LA, when no one is looking and their word stands higher than the oppressed.
October 10, 2011 at 2:20pm via mobile · Like · 3



(d) Q = 0.31. Consider it as two clusters

http://www.facebook.com/occupyLA/posts/291368864209238

Evaluation

- Dataset
  - Official Occupy LA Facebook page.
  - All the posts, comments and like information in this group from Sep. 2011 to April 2012.
  - 20569 unique users, 88442 comments, 32216 posts, 66758 likes

Evaluation

- Dataset
  - The comments number of posts we examine should not be too small.
  - We select all the posts which contain more than 35 comments. Finally, we get 23372 comments belonging to 401 posts in this Occupy LA group.

Evaluation from Linguistic Features

- Our Theoretical Basis:

  Comments on different opinions have different characteristics on their linguistic features. (R. Abbott et al., 2011)

- General Idea:

  As to each opinion classification result, if there are significant different $(p < 0.05)$ between their linguistic features, we regard it as an acceptable result.

Linguistic Features

- Linguistic Inquiry and Word Count tool (Y.R. Tausczik et al., 2010)
  - It offers 63 linguistic categories based on its dictionary which consists of 3000+ words.
  - For every word in a text, it calculates the frequency with which it matches each of these 63 categories.
  - We select 39 categories from all the features. 32 of them belong to psychological processes and 7 of them belong to personal concern area.

Linguistic Features

Areas	Categories	Examples
Social processes[1]		Mate, talk, they, child
	Family	Daughter, husband, aunt
	Friends	Buddy, friend, neighbor
	Humans	Adult, baby, boy
Affective processes[1]		Happy, cried, abandon
	Positive emotion	Love, nice, sweet
	Negative emotion	Hurt, ugly, nasty
	Anxiety	Worried, fearful, nervous
	Anger	Hate, kill, annoyed
	Sadness	Crying, grief, sad
Cognitive processes[1]		cause, know, ought
	Insight	think, know, consider
	Causation	because, effect, hence
	Discrepancy	should, would, could
	Tentative	maybe, perhaps, guess
	Certainty	always, never
	Inhibition	block, constrain, stop
	Inclusive	And, with, include
	Exclusive	But, without, exclude

Table 3.1: The 39 textual Categories in 13 areas used in our linguistic evaluation. Areas marked with 1 are psychological processes, and areas marked with 2 are personal concerns

Linguistic Features

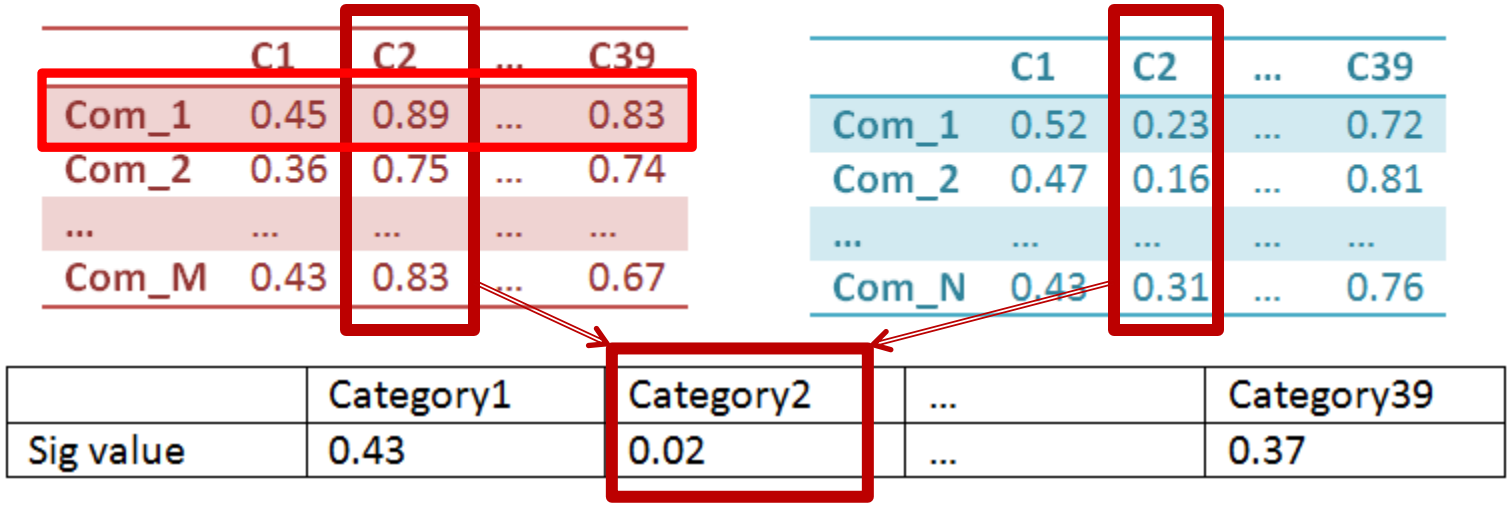
Perceptual processes[1]		Observing, heard, feeling
	See	View, saw, seen
	Hear	Listen, hearing
	Feel	Feels, touch
Biological processes[1]		Eat, blood, pain
	Body	Cheek, hands, spit
	Health	Clinic, flu, pill
	Sexual	Horny, love, incest
	Ingestion	Dish, eat, pizza
Relativity[1]		Area, bend, exit, stop
	Motion	Arrive, car, go
	Space	Down, in, thin
	Time	End, until, season
Work[2]		Job, majors, xerox
Achievement[2]		Earn, hero, win
Leisure[2]		Cook, chat, movie
Home[2]		Apartment, kitchen, family
Money[2]		Audit, cash, owe
Religion[2]		Altar, church, mosque
Death[2]		Bury, coffin, kill

Table 3.1: The 39 textual Categories in 13 areas used in our linguistic evaluation. Areas marked with 1 are psychological processes, and areas marked with 2 are personal concerns

Evaluation from Linguistic Features

- Methods

  For each of the posts which are considered including two different opinions inside, we use T-Test to check whether the LIWC scores of its two groups of comments are significant different $(p < 0.05)$ from each other.



	C1	C2	...	C39
Com_1	0.45	0.89	...	0.83
Com_2	0.36	0.75	...	0.74
...	...	...	...	...
Com_M	0.43	0.83	...	0.67

	C1	C2	...	C39
Com_1	0.52	0.23	...	0.72
Com_2	0.47	0.16	...	0.81
...	...	...	...	...
Com_N	0.43	0.31	...	0.76

	Category1	Category2	...	Category39
Sig value	0.43	0.02	...	0.37

Evaluation from Linguistic Features

- Results
  - Among all the 401 posts in Occupy LA group, 336 of them are selected as posts with two different opinions inside.
  - 240 of these 336 posts reveal different characteristics on their linguistic features, which achieves an accuracy of 71.4%

System Development

- Social Interactive Networking and Conversation Entropy Ranking Engine (SINCERE)
  - SINCERE is a diversified search engine based on user social informatics. Up to Aug. 20, 2013, its database stores all the information of 1542 Facebook public groups.
  - We implemented this opinion classification method into SINCERE as a real-time service.

System Development

Occupy Los Angeles · 55,596 like this
October 10, 2011 at 1:42pm ·

Like

Heard some of the LAPD dropped off 3 crates of supplies: hygienic products, snacks, sunscreen,etc. for Occupiers! -M.K.

Like · Comment · Share 99

676 people like this.

View previous comments 43 of 103

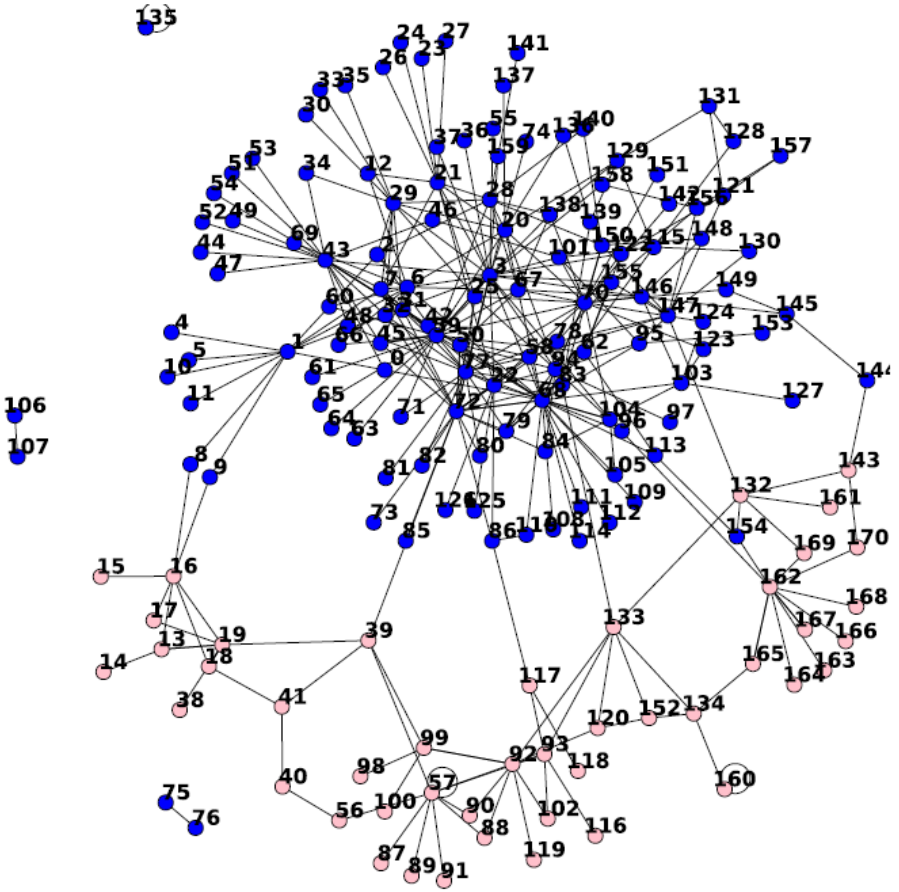
Simone Maritza Missirian Very cautiously optimistic...in the end, we need to win their hearts and minds...'cause they have guns and our movement will grow stronger faster if they don't use them against us.
October 10, 2011 at 2:14pm via mobile · Like · 3

David Sweet awesome - send that tweet out ask to be re-tweeted..!!
October 10, 2011 at 2:16pm via mobile · Like

Paradox Pollack They are looking for photo opps.
Unlike the NYPD who have a vested interest in not pissing off Wall Street, the LAPD and Mayor Villaraigosa want to support this because it is good press and our #'s are fewer. They want to sty on our goo side and not have to pay cops to wrangle us. It's simply good politics.
October 10, 2011 at 2:18pm via mobile · Like · 3

Daniel Rivas Thank you LAPD, we needed that
October 10, 2011 at 2:19pm via mobile · Like · 4

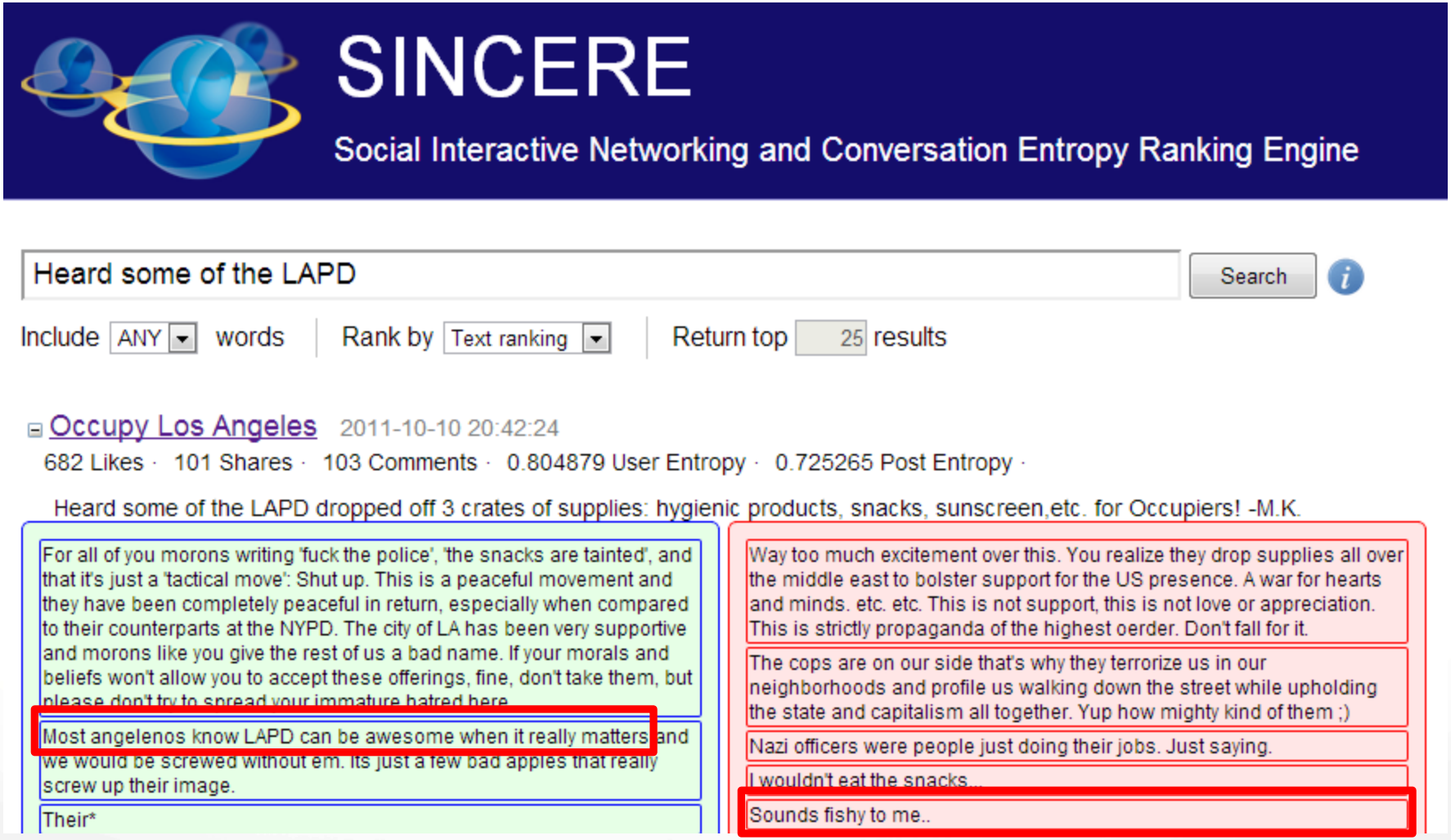
Paulina Selene Jones love ♥ let us hope they change their attitudes towards the people in all of LA. it is awesome and a beautiful thing that they are giving so much support to OLA. But we can't forget the cruel and unusual crimes against our brothers in sisters on the streets of LA, when no one is looking and their word stands higher than the oppressed.
October 10, 2011 at 2:20pm via mobile · Like · 3



(d) Q = 0.31. Consider it as two clusters

http://www.facebook.com/occupyLA/posts/291368864209238

System Development

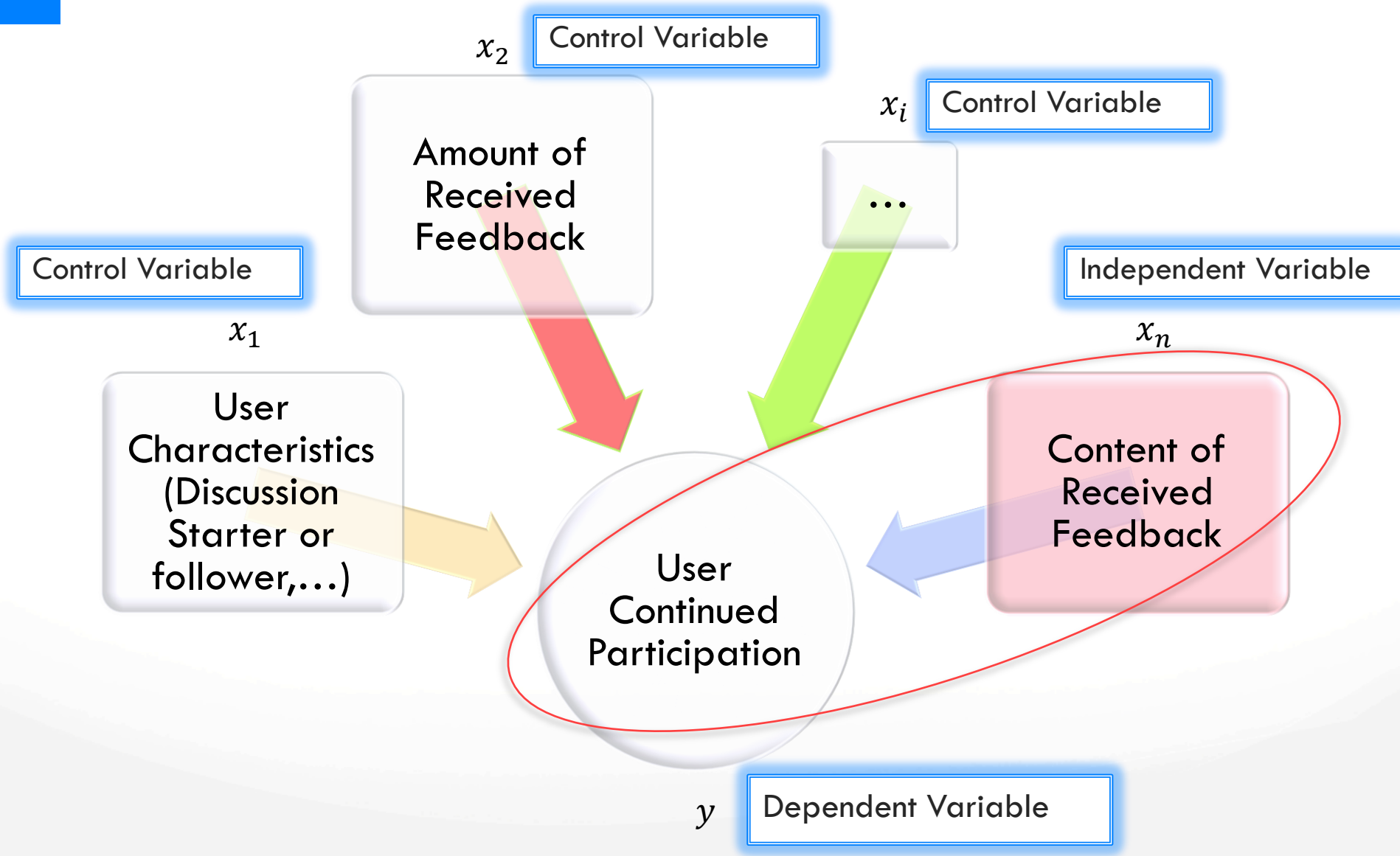


http://sincere.se

Outline

- ~~Introduction~~
- ~~Research Background~~
- ~~Classify Different Opinions~~
- Influence of Feedback with Different Opinions on User Continued Participation

Problem Description



Cox Regression Model

- Mathematical Formulation

$$h_i(t) = h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik})$$

  - x_i denotes the covariates and β_i denotes the estimated coefficients.
  - $h_i(t)$ denotes the instantaneous probability for user to leave this online group at time t.

Results Analysis

Control/Indepedent Variable	coef	exp(coef)	$Pr(<\|z\|)$	Std. Err.
OriginalPostWriter	-0.444	0.641	$< 0.001^{***}$	0.119
NonDiscussionPostsInvolved	-0.325	0.723	0.002^{**}	0.104
ComPerActivity	-0.274	0.760	0.020^{*}	0.118
SamePercent	-0.194	0.823	0.007^{**}	0.072
DifferentPercent	-0.154	0.857	0.005^{**}	0.055
SamePercent * DifferentPercent	-0.267	0.766	0.002^{**}	0.085
ComPerActivity * SamePercent	0.284	1.328	$< 0.001^{***}$	0.070
ComPerActivity * DifferentPercent	0.130	1.139	0.043^{*}	0.064

Table 4.2: Results of Cox Regression Model

Discussion and Conclusion

- Members who started more original posts revealed a longer lifetime in this social newsgroup

  Website designer can motivate users longtime participation by offering more opportunities for normal users to post their own news and become a discussion starter.

- The more feedback one individual receives after his/her comment, the more likely he/she will remain in this newsgroup.
  - This result is also proved in E. Joyce's work about newcomers in newsgroup (E. Joyce et al., 2006) and Yi-Chia's work about all members in online health support group (Yi-Chia et al, 2011).
  - Our result broads this conclusion to the behavior of all the members in online newsgroup.

Discussion and Conclusion

- The content of replies has a significant influence on user's commitment to online newsgroup.
  - Both the feedback comments from the same and different opinions can motivate user longtime participation in online newsgroup.
  - Furthermore, an interaction of feedback from both the same and different opinions during discussion can boost user future participation.

Control/Indepedent Variable	coef	exp(coef)	$Pr(<\|z\|)$	Std. Err.
SamePercent	-0.194	0.823	0.007**	0.072
DifferentPercent	-0.154	0.857	0.005**	0.055
SamePercent * DifferentPercent	-0.267	0.766	0.002**	0.085

Discussion and Conclusion

- Our result proves and broadens the conclusion of Jennifer's research in political communications. (J. Stromer-Galley et al., 2009)
  - Statistical analysis on a much larger dataset.
  - We offer a new opinion classification algorithm to partition feedback comments into different opinions automatically. (SINCERE system)
  - Instead of only focusing on political topics where people's stances are always sharp and in opposition to each other, our experiment broadens user interaction data to general discussion topics.

Thank you!

Q&A

Davis Social Links