Extracting Diverse Sentiment Expressions With Target-dependent Polarity from Twitter

Lu Chen, Wenbo Wang, Meenakshi Nagarajan, Shaojun Wang, and Amit P. Sheth

Target-specific Corpus

Just saw **The Avengers**! It lived up to the hype.

Just got out of **the Avengers** movie. It was all storyboard, no story. What a clumsy, clunky mess. Got kind of bored 90 min in. B-

Went saw **The Avengers** tonight and man it was epic....a definite must see!!!

President Obama watched **the Avengers**.

The Avengers movie was bloody amazing! Mmm looking good Robert Downey Jr and Captain America ;)

If I can be half as cool as Scarlett Johansson in **The Avengers** when I grow up, my life will be accomplished.

Saw **the avengers** last night. Mad overrated. Cheesy lines and horrible writing. Very predictable.

My best friend wants to see **The Avengers** today, so of course I'll go. Now to decide which Batman shirt to wear...

Okay, **the Avengers** was worth the hype. Best superhero movie of all time. Blown away, here.

Target-dependent Sentiment Expressions

Just saw **The Avengers**! It lived up to the hype.

Just got out of **the Avengers** movie. It was all storyboard, no story. What a clumsy, clunky mess. Got kind of bored 90 min in. B-

Went saw **The Avengers** tonight and man it was epic....a definite must see!!!

President Obama watched **the Avengers**.

The Avengers movie was bloody amazing! Mmm looking good Robert Downey Jr and Captain America ;)

If I can be half as cool as Scarlett Johansson in **The Avengers** when I grow up, my life will be accomplished.

Saw **the avengers** last night. Mad overrated. Cheesy lines and horrible writing. Very predictable.

My best friend wants to see **The Avengers** today, so of course I'll go. Now to decide which Batman shirt to wear...

Okay, **the Avengers** was worth the hype. Best superhero movie of all time. Blown away, here.

Extracting a **diverse** and **richer** set of sentiment-bearing expressions, including formal and slang words/phrases

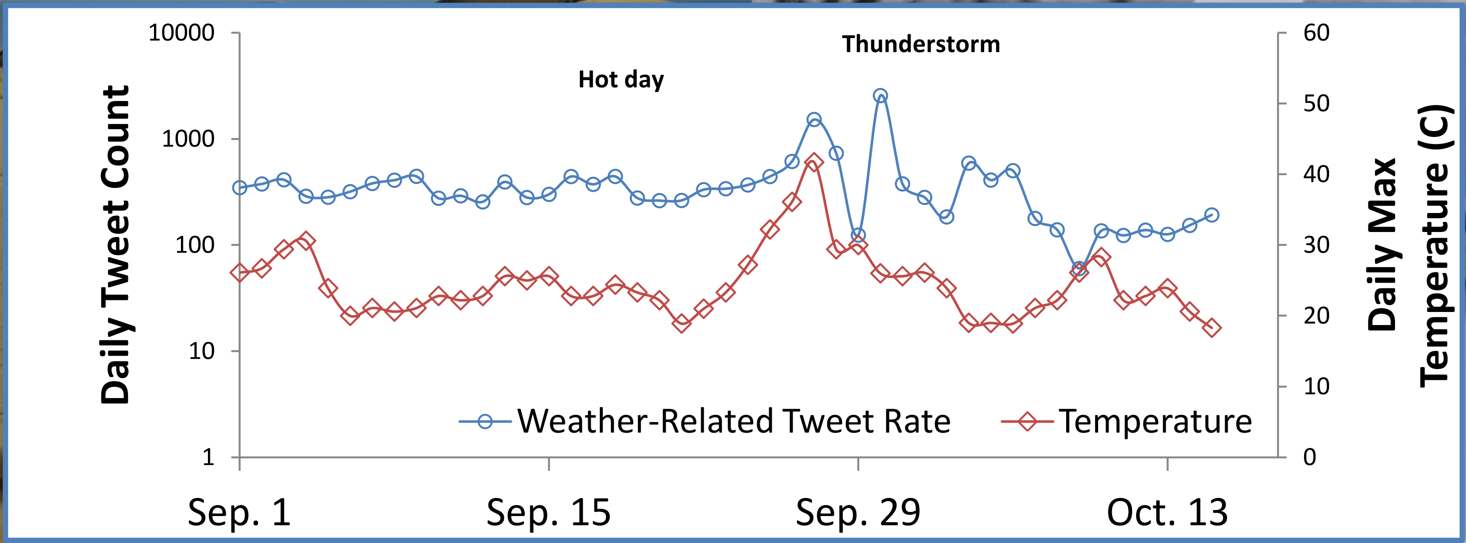
Assessing the **target-dependent** polarity of each sentiment expression

A novel formulation of assigning polarity to a sentiment expression as a **constrained optimization problem** over the tweet corpus

OMG I have to Tweet that!

A Study of Factors that Influence Tweet Rates

Emre Kıcıman, emrek@microsoft.com



Event Diffusion Patterns in Social Media

Australian National University

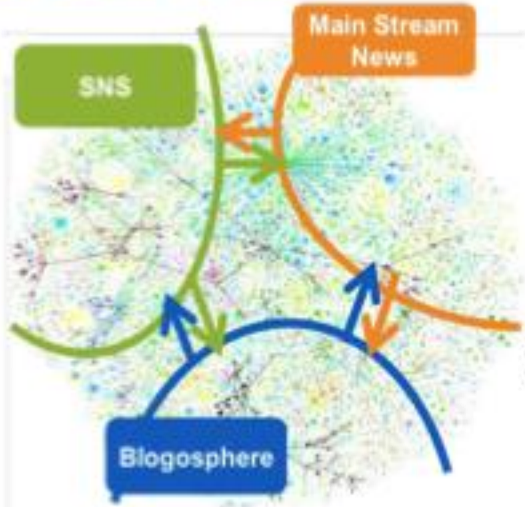
Analyze how real-world events spread and interact with other events across diverse social media (news, blog and SNS) and understand the underlying structure of user network using *ICWSM'11 Spinn3r* dataset.

Cricket Game Cancelation

Queensland Floods 2011

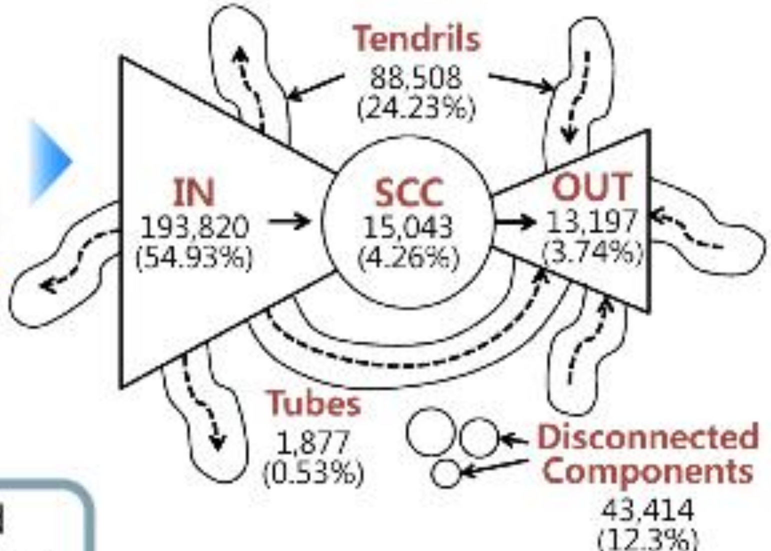
Australian Open 2011

Events are connected across different categories.



Information is propagated across sites through hyperlinks in the text of a post.

90% of SNS posts and **50%** of blog posts make citations to news articles.



Unbalanced interactions between users on events: 55% of users cite less than 5% other users.

Minkyoung Kim, Lexing Xie, Peter Christen

Viraliy and Susceptibility in Information Diffusions

Tuan-Anh HOANG, Ee-Peng LIM
Living Analytic Research Centre
School of Information System, Singapore Management University

Viral (information) items: pieces of information that are easily to be diffused through word-of-mouth

Viral users: people that are good at diffusing information

Susceptible users: people that are easily to be convinced

The image was taken from http://www.wordscanheal.org

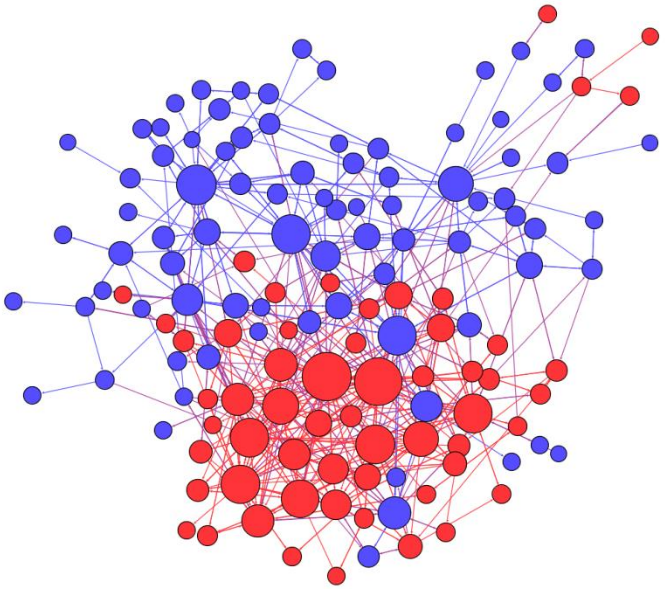
Would you follow a (automated) complete stranger?

Online **social experiment**

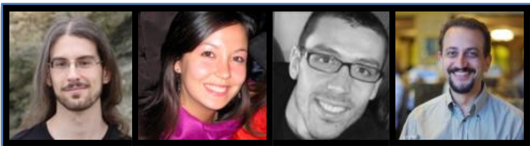
Bot interacting with real users

Gain **popularity** with simple automated activity

Polarization of reactions



People are Strange when you're a Stranger:
Impact and Influence of Bots on Social Networks



L.M. Aiello, M. Deplano, R. Schifanella, G. Ruffo
ARC²S Group - http://arcs.di.unito.it

Jaram Park video presentation

You Too!? Mixed-Initiative LDA Story Matching To Help Teens in Distress

Karthik Dinakar+ Birago Jones+ Henry Lieberman+ Rosalind Picard+ Carolyn Rose* Matthew Thoman$^{\Psi}$ Roi Reichart$^{\theta}$

+ MIT Media Lab $^{\theta}$MIT CSAIL *Carnegie Mellon University $^{\Psi}$Northeastern University

Teenage Distress

- Negative effects of bullying are well known
- Bullying in the context of teenage drama
- Psychiatry: need to foster cognitive empathy

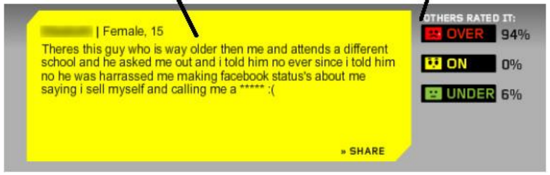
Mitigation: hypothesis

- Detection of distributions of teenage drama
- Using detection to power reflective thinking
- Indexing appropriate help material

Corpora

Distressing personal experience

Severity poll ratings



- MTV's ww.athinline.org – fighting bullying
- 5500 personal stories of distress
- Severity poll ratings for each story
- Third party advice as comments for each story

Extracting high-level themes

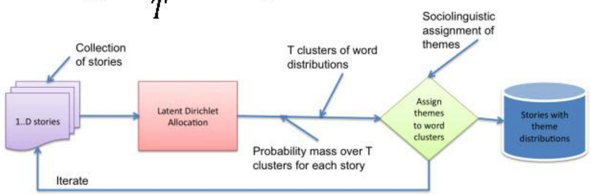
Let T = # of themes in teenage stories
D = # of stories
N = # words in the corpus.

Let P(z) = distribution over themes z in a particular story, & P(w|z) = probability mass over word w given theme z.

$$P(w_a) = \sum_{b=1}^{T} P(w_a|z_a = b)P(z_a = b)$$

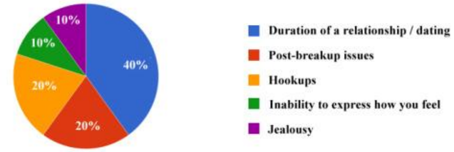
$$\theta^{(d)} = P(z) \qquad \varphi^{(b)} = P(w|z = b)$$

$$\alpha = \frac{50}{T} \qquad \beta = 0.01$$



Story thematic distributions

"okay so i had this boyfriend, and i kissed another guy, but i told him. and we broke up. then months later we got back together and his friend came up to me and kissed me. and then her told my boyfriend, so he broke up with me. but i want him back!"



Thematic story matching

- Apply model to a new story to get a thematic distribution
- Similarity metric: Kullback-Liebler divergence to fetch similar old stories

Evaluation+ Error Analysis

- Q1: Validity of themes extracted
- Q2: Similarity & usefulness of showing matched stories
- Strong results for LDA+ versus control
- New stories with outlier themes didn't match well
- Promising results: current actual deployment on MTV www.athinline.org

Results

n = 12 participants control = tf-idf cosine similarity

	% Strongly Agree		% Agree		% Strongly Disagree		% Disagree	
	LDA+	Control	LDA+	Control	LDA+	Control	LDA+	Control
Q1	45.0%	0%	22.1%	3%	15.9%	46%	17%	51%
Q2	35.3%	0%	23.0%	8.3%	15%	31.0%	13.2%	35.1%

Media Laboratory

Massachusetts Institute of Technology