Quantifying Controversy in Social Media

Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, Michael Mathioudakis. Aalto University, Finland

WSDM 2016

Presenter: Robert Baykov

Background

Social media is a hot area of examination in data mining.

Offers vast and rich medium for analysis .

Example: 500M tweets are sent daily, 37M blogs hosts on wordpress and more.

Background

In this work we will focus on interaction between users on social media.

Interaction in the context of discussion amongst users, namely agreement and disagreement.

(i.e. **controversy**)

Motivation

Why controversial discussion?

Offers many useful insights into *real* social phenomena:

- Peer influence and bias.
- Topic framing.
- Public opinion evolution.

Challenge

Domain Specific Knowledge: in order to identify controversy we often require domain specific information in order to discern user-affiliation/features. (e.g. topic:politics sides: republican/democrat)

- How can we identify controversy lacking DSK?

Quantification: Not all discussions are alike, which metrics or features are useful for controversy quantification? How can we apply metric effectively across different topic domains?

- Given a topic, how can we discern the *degree* of controversy?

Problem

Many existing works identify controversial discussions in social media.

However, they have several limitations:

- Address mainly a single topic, e.g. politics.
- Handle specific major events, e.g. elections.
- Rely heavily on domain specific knowledge for analysis (e.g. hand picked hashed tags)

These limitations make it difficult to analyze new topics.

Goal

Overcome previous limitations:

1. Identify controversial topics of discussion in ***any domain*** (e.g. politics, economics, etc).

2. ***Quantify***, (measure) the degree of controversy.

Intuition

1. We can represent topic discussions between users as ***conversation graphs***, where:
    - Vertices are users
    - Edges are interactions (posts, endorsements etc.)

2. Hypothesise, that such graphs will exhibit a ***clustered structure***.

Why clustered?

Our conversation graph is amongst users.

Individuals will often endorse and amplify each other's arguments*, resulting in cluster formulation.

Analysis of clustering identifies controversy. (two sides?)

*Adamic et al. 2004, Akoglu et al 2014, Conover et al. 2011

Conversation Graph Features

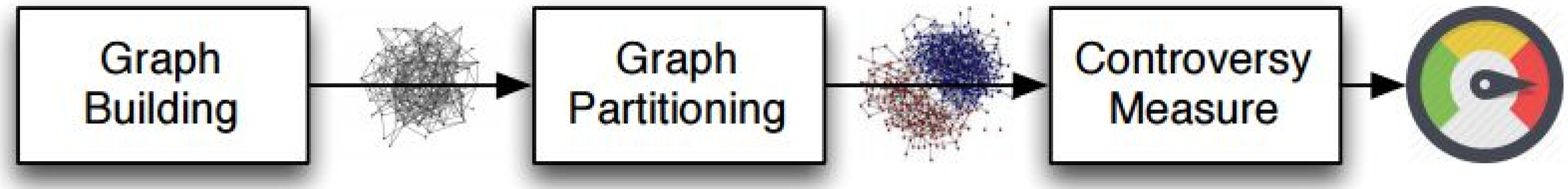
Analyze candidate features of controversy:

1. Conversation graph structure, ***who agrees*** with whom.
2. Structure of social network, ***who is connected*/participating** with whom in conversation.
3. ***Content*** or keywords in topic.
4. ***Sentiment* or tone**, of discussion.

Extract features and compute conversy score of topic.

High Level Approach

1. Build *conversation graph* of topic, each edge represents two users *in agreement.*
2. Identify potential sides based on graph structure.
3. Quantify amount of controversy.



Graph Data Set

This work focuses graphs build on top of twitter's dataset.

Graph represents conversations related to ***one topic***.

A ***topic***, represents conversation results to a *query*.

Example, results for the query “#gunsense” are conversations on the topic of gun violence in the United States.

Table 1: Datasets statistics: hashtag, sizes of the follow and retweet graphs, and description of the event. The top group represent controversial topics, while the bottom one represent non-controversial ones.

Hashtag	# Tweets	Retweet graph		Follow graph		Description and collection period (2015)
		$\lvert V \rvert$	$\lvert E \rvert$	$\lvert V \rvert$	$\lvert E \rvert$	
#beefban	84 543	1610	1978	799	6026	Government of India bans beef, Mar 2–5
#nemtsov	183 477	6546	10 172	2156	46 529	Death of Boris Nemtsov, Feb 28–Mar 2
#netanyahuspeech	254 623	9434	14 476	4292	297 136	Netanyahu's speech at U.S. Congress, Mar 3–5
#марш	118 629	2134	2951	1189	16 471	Protests after death of Boris Nemtsov ("march"), Mar 1–2
#indiasdaughter	167 704	3659	4323	1542	9480	Controversial Indian documentary, Mar 1–5
#baltimoreriots	218 157	3902	4505	1441	28 291	Riots in Baltimore after police kills a black man, Apr 28–30
#indiana	116 379	2467	3143	946	24 328	Indiana pizzeria refuses to cater gay wedding, Apr 2–5
#ukraine	287 438	5495	9452	3383	84 035	Ukraine conflict, Feb 27–Mar 2
#gunsense	318 409	7106	11 483	1821	103 840	Gun violence in U.S., Jun 1–30
#leadersdebate	1 139 344	25 983	44 174	9566	344 088	Debate during the U.K. national elections, May 3
#sxsw	343 652	9304	11 003	4558	91 356	SXSW conference, Mar 13–22
#1dfamheretostay	501 960	15 292	26 819	3151	20 275	Last OneDirection concert, Mar 27–29
#germanwings	907 510	29 763	39 075	2111	7329	Germanwings flight crash, Mar 24–26
#mothersday	1 798 018	155 599	176 915	2225	14 160	Mother's day, May 8
#nepal	1 297 995	40 579	57 544	4242	42 833	Nepal earthquake, Apr 26–29
#ultralive	364 236	9261	15 544	2113	16 070	Ultra Music Festival, Mar 18–20
#FF	408 326	5401	7646	3899	63 672	Follow Friday, Jun 19
#jurassicworld	724 782	26 407	32 515	4395	31 802	Jurassic World movie, Jun 12-15
#wcw	156 243	10 674	11 809	3264	23 414	Women crush Wednesdays, Jun 17
#nationalkissingday	165 172	4638	4816	790	5927	National kissing day, Jun 19

Graph Construction

For each topic hashtag, one vertex to user who employs hashtag.

Construct agreement edges between pairs of users, *u* and *v* if:

1. Two or more common *retweets* between *u* and *v*.
2. One user follows another or vice versa.
3. Similar Content:
    - *a.* use of identical non-topic hashtags.
    - b. sharing of identical URL or URL domains.
4. Hybrid Metric: Ruan et al. 2013, high cosine similarity between frequency vectors of hashtags employed by *u* and *v.*

Graph Partitioning

Use a standard algorithm, METIS for partition creation.

Visualize graph with Gelphi's ForceAtlas2 several render methods.

Manual observation of visualizations.

Counterintuitive observation, *content based feature selection is very ineffective!*

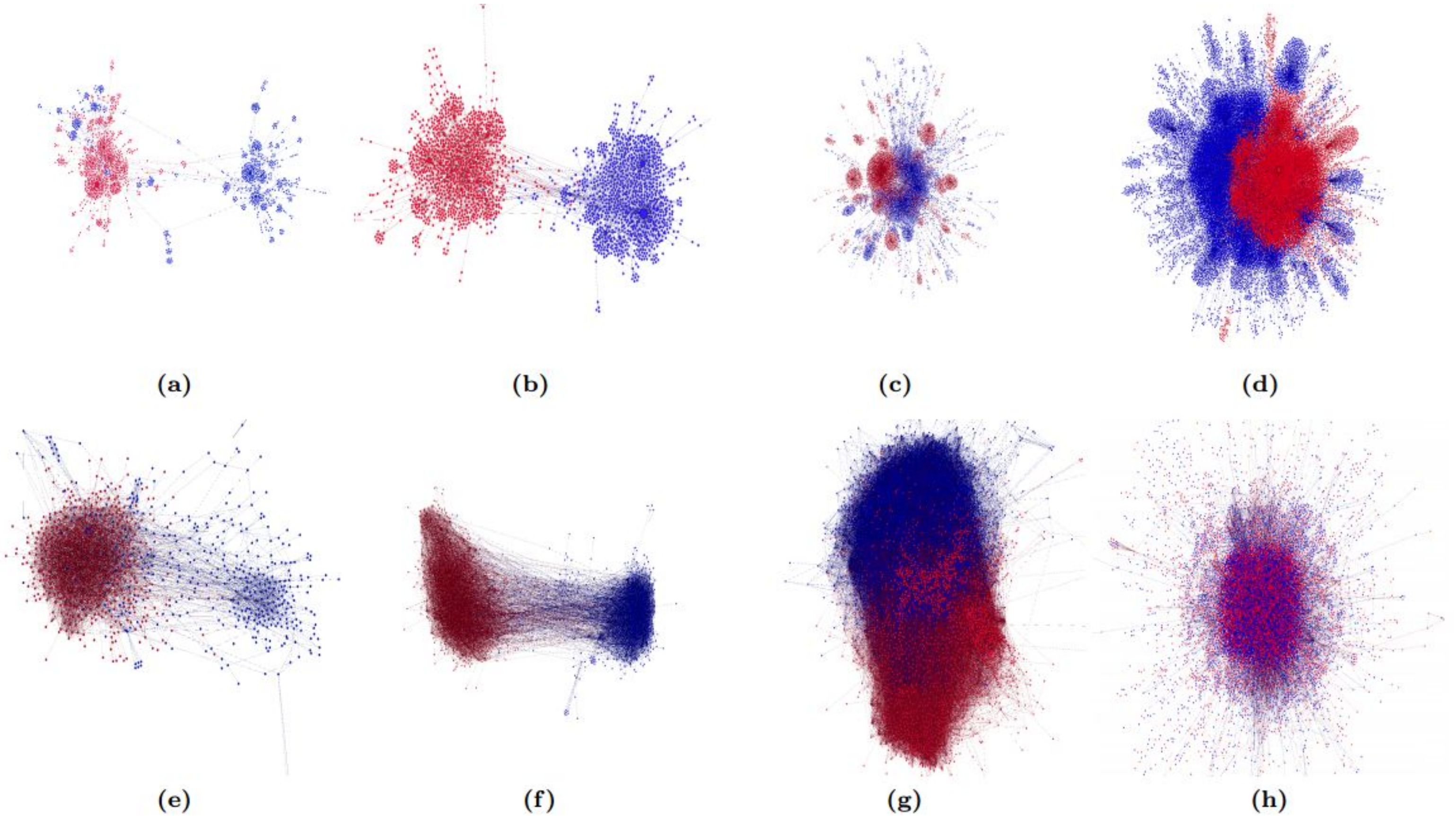


Figure 2: Sample conversation graphs with retweet (top) and follow (bottom) features (visualized using the force directed layout algorithm in Gephi). The left side is controversial, (a,e) #beefban, (b,f) #russia_march, while the right side is non-controversial, (c,g) #sxsw, (d,h) #germanwings.

Controversy Measure

Five metric methods to *quantify* topic controversy:

1. Random walk
2. Betweenness
3. Embedding
4. Boundary Connectivity
5. Dipole Moment

*Random Walk

Controversial topics have authors on both sides:

- What is the likelihood of a random user on one side to be exposed to *authoritative* (high value) content from the *other* side?

Find k highest-degree vertices, the *authoritative* vertices (large number endorsements flow towards this user).

Choose random vertex in either partition, randomly traverse graph ending on authoritative vertex ***from either side.***

Random Walk Controversy (RWC)

Given two graph partitions (sides of argument) and two random walks.

RWC measures difference in probability of two events:

1. Walks end in origin partition.
2. Walks end in non-origin partition.

Result metric of 0.0 - 1.0, probability of walk ending in origin partition.

In other words: highly controversial topics polarise opinion, users are less likely to associate with opposite side of argument.

Evaluation

Retweet and follow graph evaluation, content and hybrid tossed out because of poor performance.

From Table 1, several million tweets analyzed by hand picked, *different* topics.

To avoid overfitting a 40%:60% train:test split is used from gathered data.

Beanplots (right) show the probability density distribution computed on all topics.

A good beanplot will show high separation between controversial and non-controversial topics.

RWC performs the best. Boundary and Dipole worst.

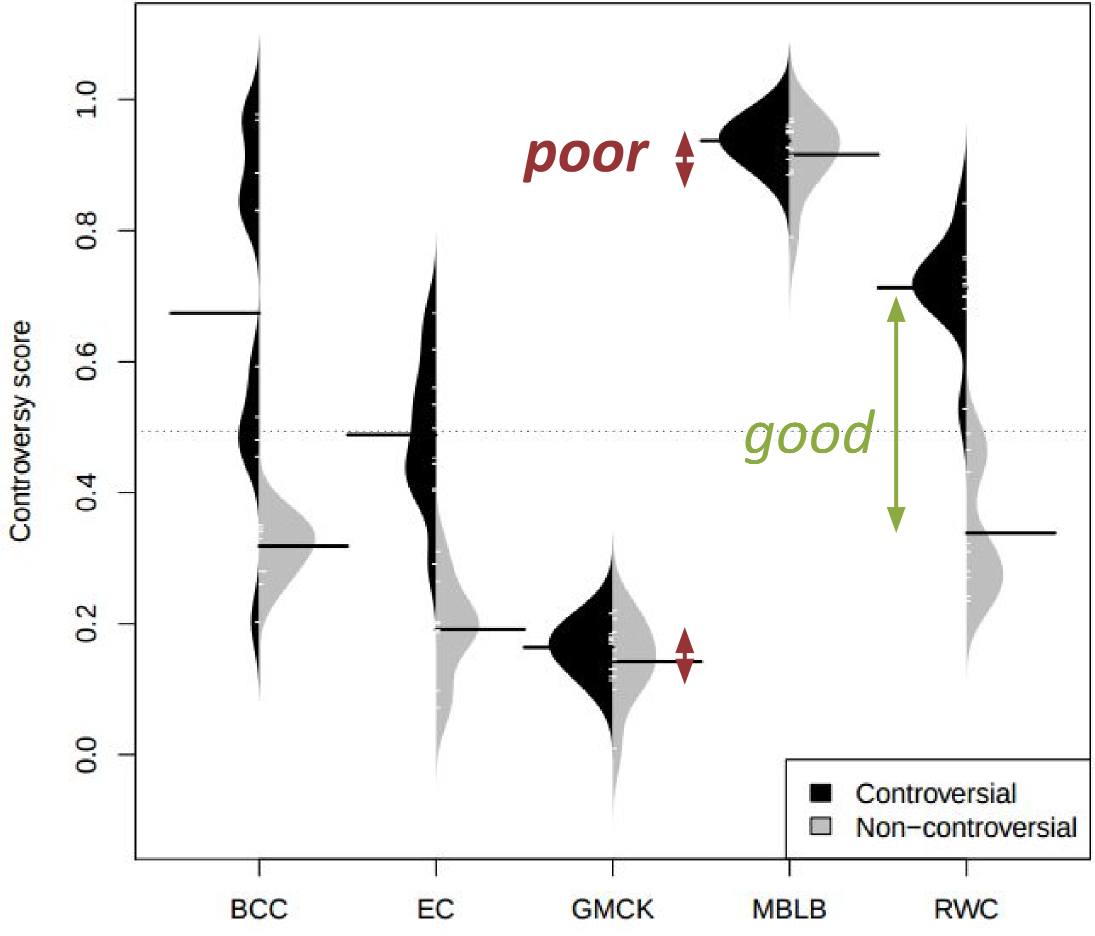


Figure 4: Controversy scores on *retweet* graphs of various controversial and non-controversial datasets

External Datasets

Table 2: Results on external datasets. The 'C?' column indicates whether the previous study considered the dataset controversial (ground truth).

Dataset	$\lvert V \rvert$	$\lvert E \rvert$	C?	*RWC*	*BCC*	*EC*	*GMCK*	*MBLB*
Political blogs	1222	16 714	✓	0.42	0.53	0.49	0.18	0.45
Twitter politics	18 470	48 053	✓	0.77	0.79	0.62	0.28	0.34
Gun control	33 254	349 782	✓	0.70	0.68	0.55	0.24	0.81
Brazil soccer	20 594	82 421	✓	0.67	0.48	0.68	0.17	0.75
Karate club	34	78	✓	0.11	0.64	0.51	0.17	0.11
Facebook university	281	4389	✗	0.35	0.26	0.38	0.01	0.27
NYC teams	95 924	176 249	✗	0.34	0.24	0.17	0.01	0.19

We can see the qualification metrics are a good indicator of controversy when applied to data sets from existing works.

*Karate club, poor outcome on small graph.

Limitations

A key limitation is the use of Twitter, albeit it is a good representation given its current popularity.

Multi-sided arguments are not represented in this work.

Reliant on graph partitioning, computationally hard problem.

Data choice is difficult, controversy can be a matter human of opinion.

Conclusion

This work presents a novel approach to identifying and quantifying controversy in social media on *arbitrary topics.*

First large scale analysis of multiple cross domain topics on Twitter.

Random walk method produces highest quality results.

Results of this work can be practically applied to provide “news diet” to social media consumers.

Thank You

Q&A

Betweenness

Betweenness centrality gives us a metric for vertex partition cross-over facilitation based on shortest paths in the graph.

Vertices with high betweenness serve as cross-over points between partitions. More shortest path connections which pass through the vertex increases the score.

Betweenness Centrality Controversy **(BCC)** measures graph divergency (closer to one implies higher divergency)

Embedding

Measure of embedded vertex (rendered graph) distance.

Embedding Controversy **(EC)** between two partitions X,Y is given by:

$$EC = 1 - \frac{d_X + d_Y}{2d_{XY}}.$$

A value closer to one shows higher graph separation, higher controversy.

Boundary Connectivity & Dipole Moment

Boundary vertices are located on the edges of partitions, a higher controversy set of partitions will produce boundary vertices well (more) connected to its *internal vertices.* **(GMCK)**

Dipole measures produces users who associate (more) with only one side, polarised users.

Applying measure **(MLBL)**, Morales et al. 2015, gives a polarization score to partitions.

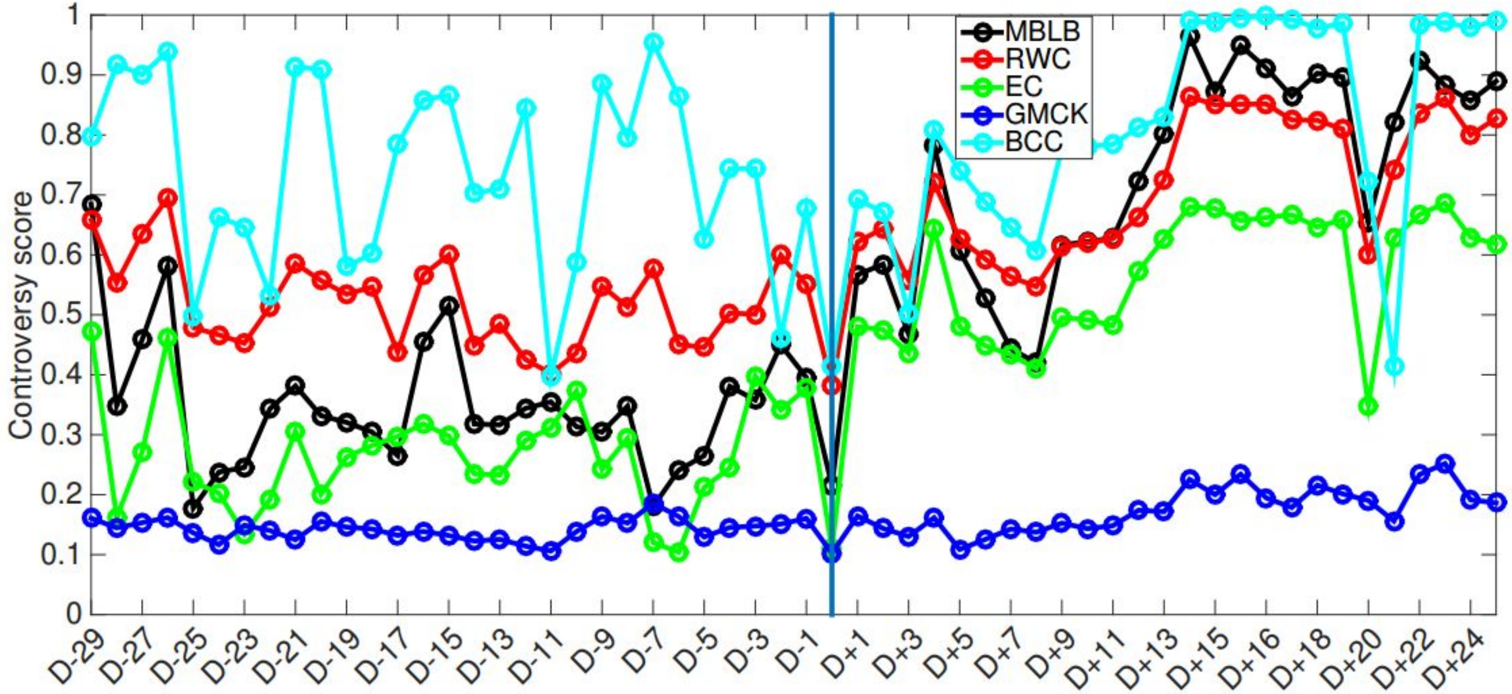


Figure 6: Controversy scores on 56 retweet graphs from Morales et al. Day 'D' (indicated by the blue vertical line) indicates the announcement of the death of president Hugo Chavez.

img placeholder :)