Visualizing the Spread of Rumors on Twitter during Disaster Events

CSE 512 Final Project

Cynthia Andrews
caandrew@uw.edu

Graeme Britz
grbritz@uw.edu

Rama Gokhale
ramag@uw.edu

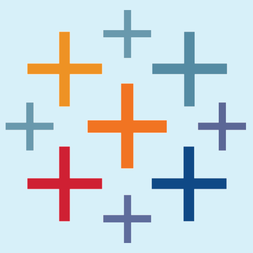
Zeno Koller
kollerz@uw.edu

Introduction

The **Emerging Capacities of Mass Participation** (emCOMP) Lab at the University of Washington studies the collective sense-making process that takes place on Twitter during crisis events. During an event, they collect tweets related to specific topics, or rumors, and analyze them with a variety of existing methods.

They are currently investigating how specific Twitter accounts can have a large effect on the direction that a rumor spreads. They want to be able to identify which accounts are the "major players", i.e. those who strongly influence the spread of a rumor in a certain direction, while still getting an overall sense of the rumor. Our work provides a new tool they can use to further their research.

Existing Methods



Members of the Data Science sub-group write database queries and create plots of affirmation and denial volume over time in Tableau.

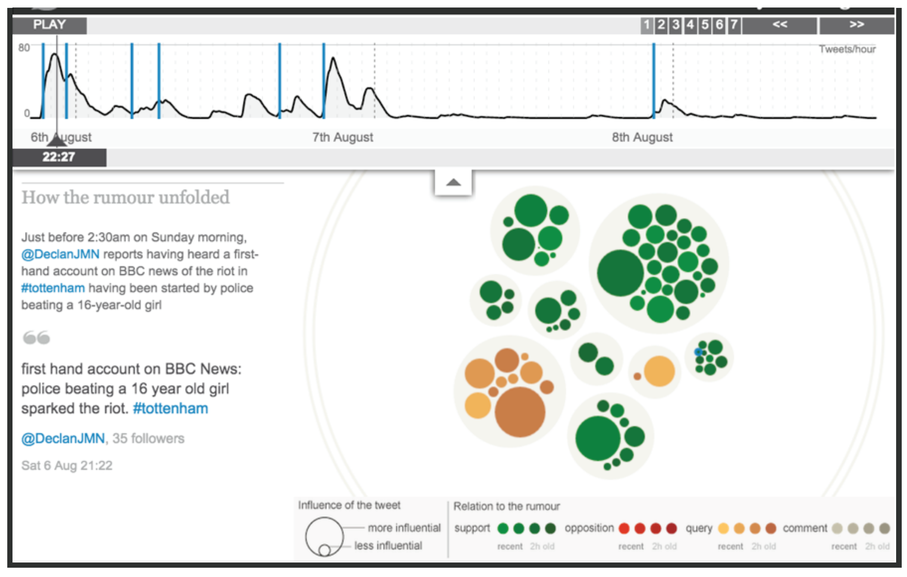
They investigate periods with high tweet volume and manually try to identify accounts that are being heavily re-tweeted at that time.

IDENTIFICATION

Each account's individual tweet volumes are graphed on a time-series plot with volume represented by area, and rumor encodings represented by color..

VISUALIZATION

Existing Work



http://bit.ly/1gV4Ncs

This is one of the few visualizations that focuses on the spread of rumors and how the sentiment of participants changes over time. Its prime directive is to "show the birth and death of rumours on Twitter".

Our Design

We had two main goals when designing our tool. The first was to build visualizations that brought more emphasis to their primary task, finding the accounts that were "major players", than existing methods. The second was to automate many of the database queries and data processing needed to analyze the data.

The main panel displays the retweet lifecycle of individual tweets over time. Each line is a tweet, whose color indicates its code. A tweet travels from left to right over time on the x-axis. Its total exposure (an aggregate of the follower counts of all accounts that have emitted the tweet) is measured on the y-axis.

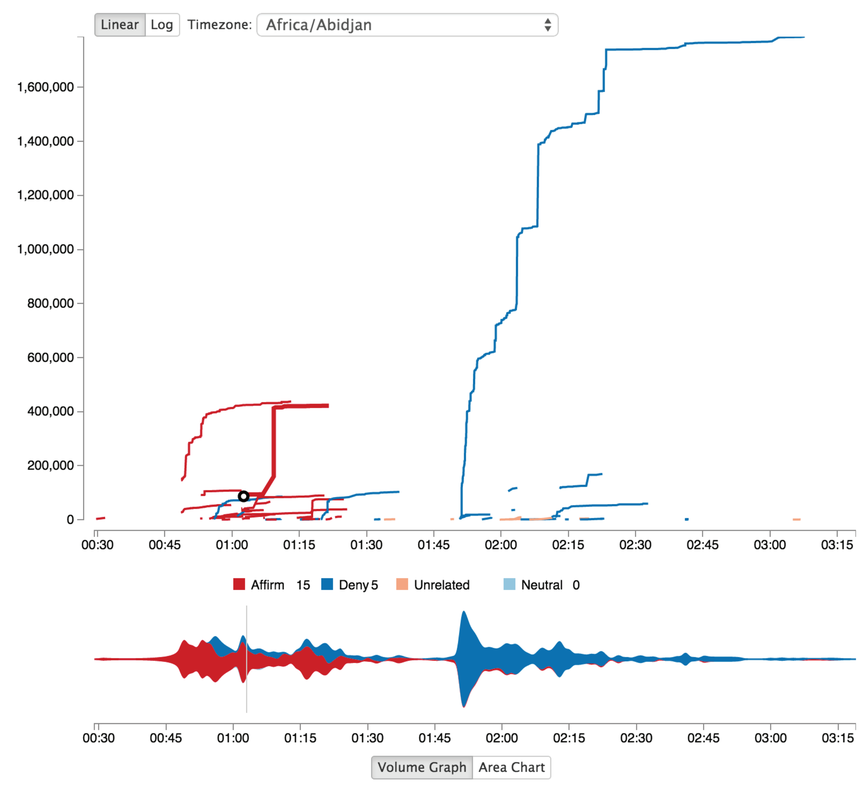
Users can hover or click on a line in the main panel to get details about that tweet. This information includes the original tweet and the names and follower counts of all accounts emitting that tweet.

7News Yahoo!7 @Y7News
01:02 · 84790 followers

Reports of police raids at #Lakemba, south-west of Sydney. Rolling coverage on #siege situation: http://t.co/jfQHQxCKbb

Name	Time ↑	Followers
nfeatherby	01:02	220
MarijaJovan7	01:02	1268
Loupo85	01:02	2835
Mon_Commentaire	01:02	3157

Name	RTs	Exposure ↓
AFPmedia	425	1784926
Y7News	27	336719
SkyNewsAust	128	291765
Mo_Taha1	47	101216
safimichael	52	83290
cerenomri	12	68988
Lisa_Wilkinson	19	55023



The leaderboard allows users to view an overall summary of the most influential accounts across the current time viewport.

This component shows the coded tweet volumes over time. When a selection is made here, the other components update to reflect the new time viewport. Two views, one by area, and one by volume, are provided for a better understanding of different trends.

bit.do/twitter-rumors