

# Users, User Roles, and Topics in School Shooting Collections of Tweets

Liuqing Li, Rishabh Anand, and Edward A. Fox  
[\[liuqing, risha97, fox\]@vt.edu](mailto:[liuqing, risha97, fox]@vt.edu)

Presentation for WADL 2019

<http://fox.cs.vt.edu/talks/2019/20190606LiAnandFoxWADL2019.pdf>

Department of Computer Science, Virginia Tech  
Virginia Polytechnic Institute and State University  
Blacksburg, VA 24061 USA

June 6, 2019

# Background

- School shootings in U.S.
  - More than 100 people were killed or injured in school shootings in the United States in 2018\*.
  - These mass shootings have had considerable effect on survivors, families, schools, and the public.

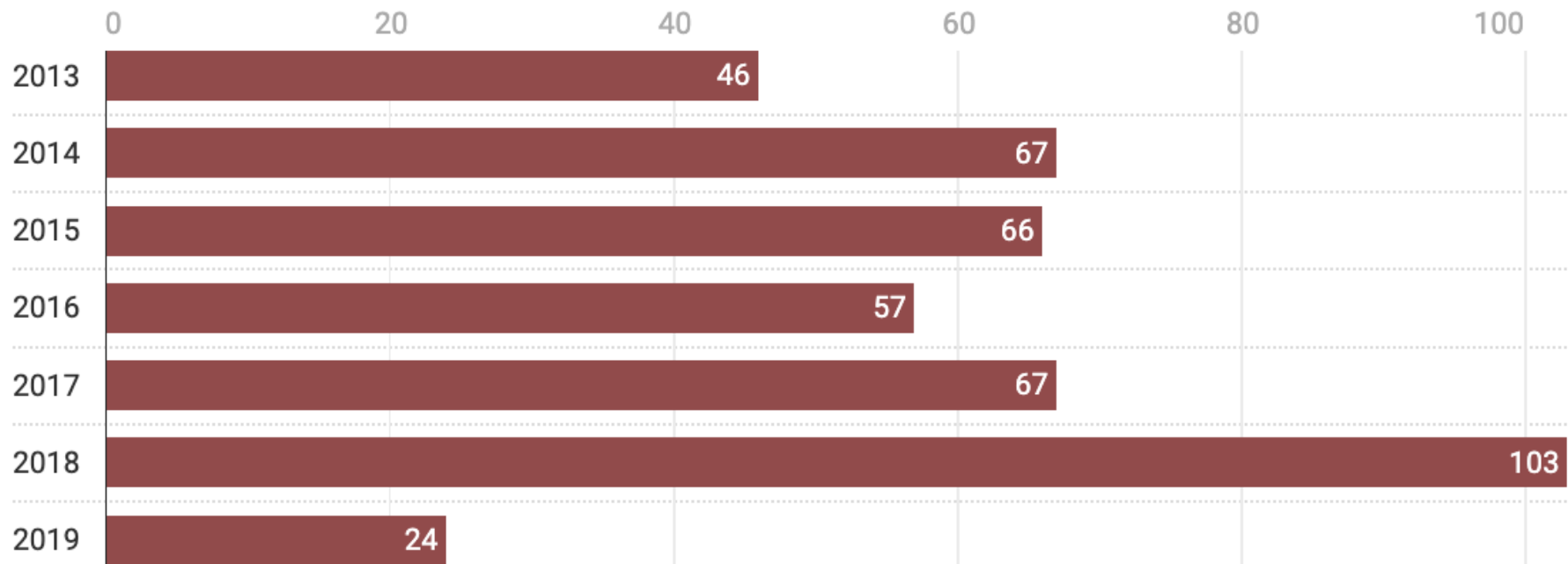
\* BBC. 2018 “Worst Year for US School Shootings”. <https://www.bbc.com/news/business-46507514>. Accessed: 2019-06-

# Background

- Count of school shootings in U.S.

## Gunfire on School Grounds\*

Every time a firearm discharges a live round inside or into a school building or onto a school campus or grounds (including college or university), as documented by the press



\* Chart: Kelly Zegers

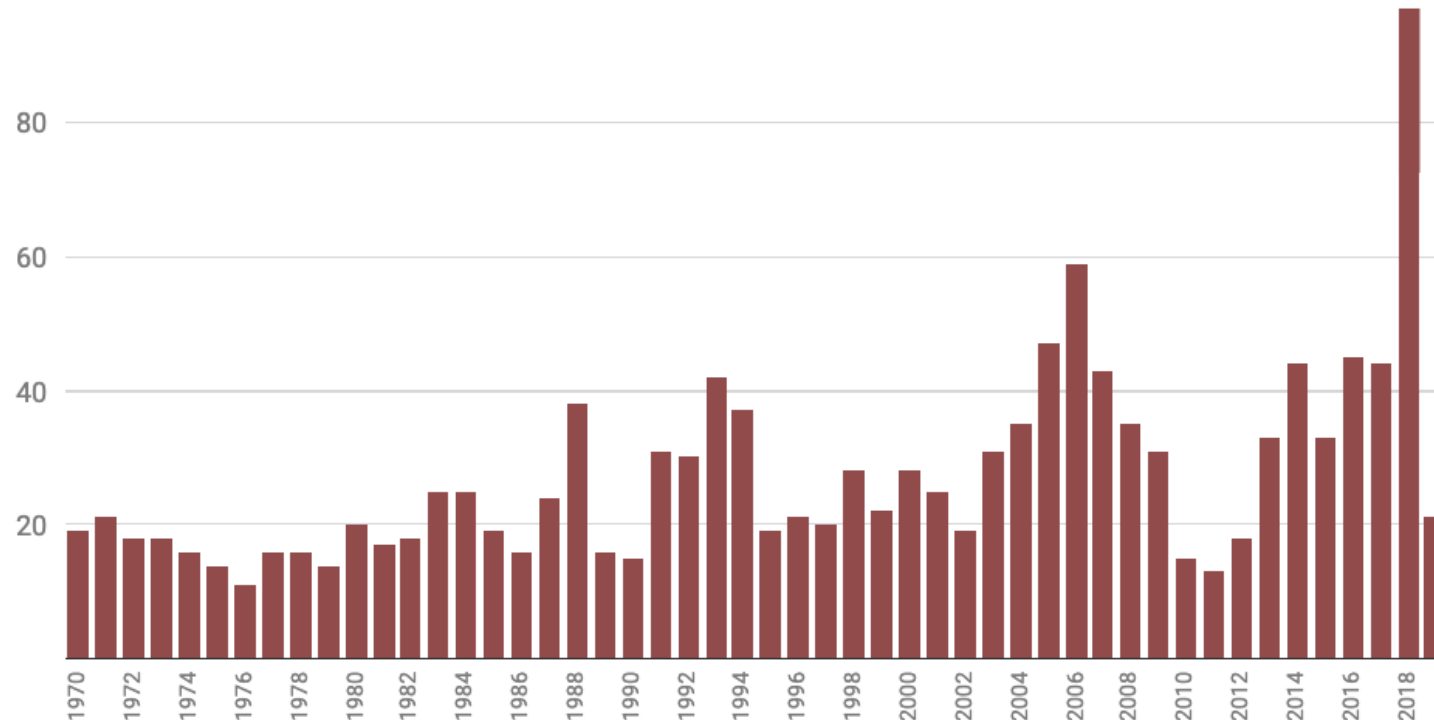
Source: Everytown for Gun Safety Support Fund

# Background

- School shootings in U.S.

## School Gun Incidents by Year (K-12)

Every instance in which a gun is brandished, fired, or a bullet hits K-12 school property for any reason, regardless of the number of victims, time, day of the week, or reason



\* Chart: Kelly Zegers Source: Riedman, David, and Desmond O’Neill. “CHDS – K-12 School Shooting Database.” Center for Homeland Defense and Security

# Background

- School shootings in U.S.
  - More than 100 people were killed or injured in school shootings in the United States in 2018\*
  - These mass shootings have had considerable effect on survivors, families, schools, and the public.
- User-based Topic Analysis
  - Voluminous discussions among users on social media
  - We treat Twitter as a barometer of users and topics, to investigate the patterns regarding school shootings.
  - We compare the behavior of different types of tweeters.

\* BBC. 2018 “Worst Year for US School Shootings”. <https://www.bbc.com/news/business-46507514>. Accessed: 2019-06-

# Related Work

- Single Event Analysis
  - Sentiments/emotions (e.g., sadness, anxiety, surprise)
  - Sentiment/emotion patterns (e.g., shift, desensitization)
  - Tweeting patterns (e.g., keywords, mentions, and hashtags)
- Multiple Events Analysis
  - Keyword filtering and topic modeling on a random sample of tweets
  - A domain network graph for mass shooting events by extracting embedded URLs from tweets
- User-based Analysis
  - Users' emotion patterns across three school shooting collections

# Research Questions

1

Is there any **pattern regarding user engagement** in school shootings?

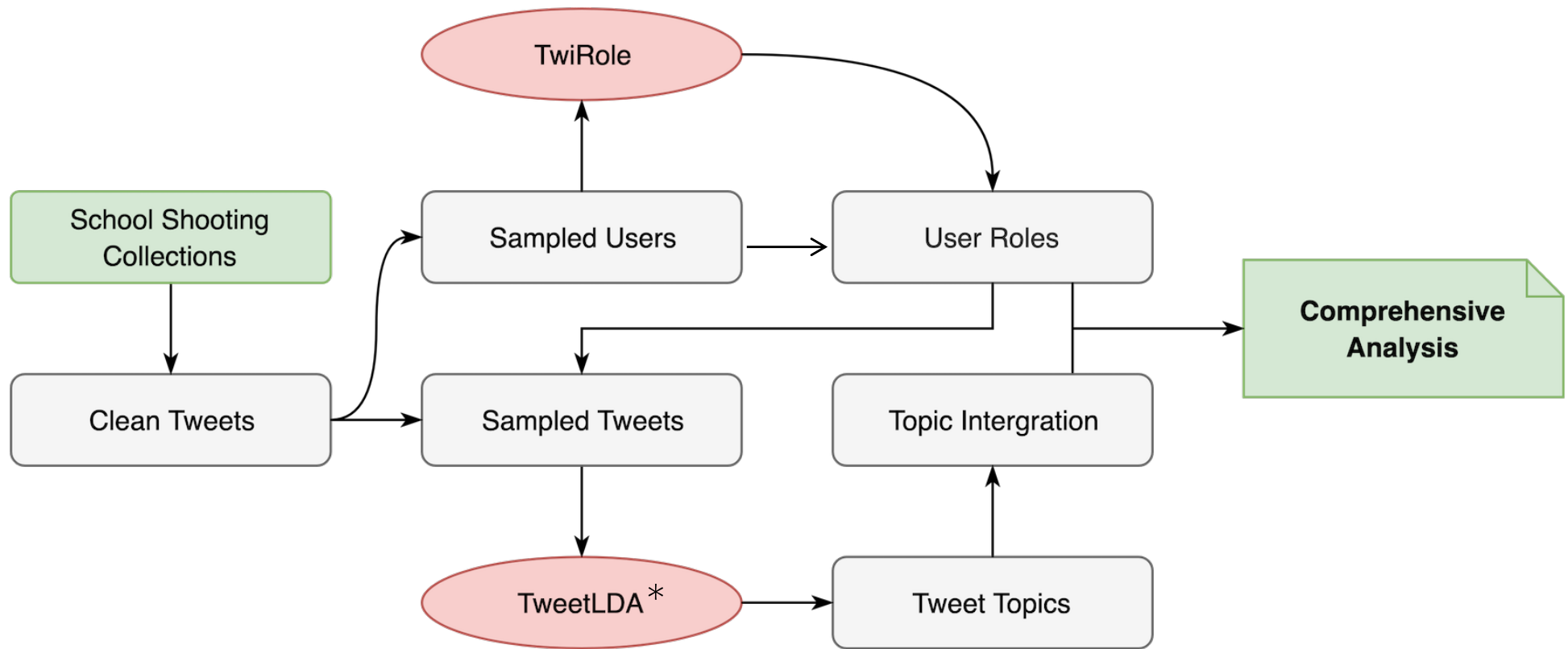
2

Is there any **pattern regarding topics** across different collections?

3

Do different users have **different concerns** regarding specific topics?

## Data Flow Diagram



\* Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., and Li, X. Comparing Twitter and Traditional Media using

Topic Models. In European conference on information retrieval (2011), Springer, pp. 338–349.

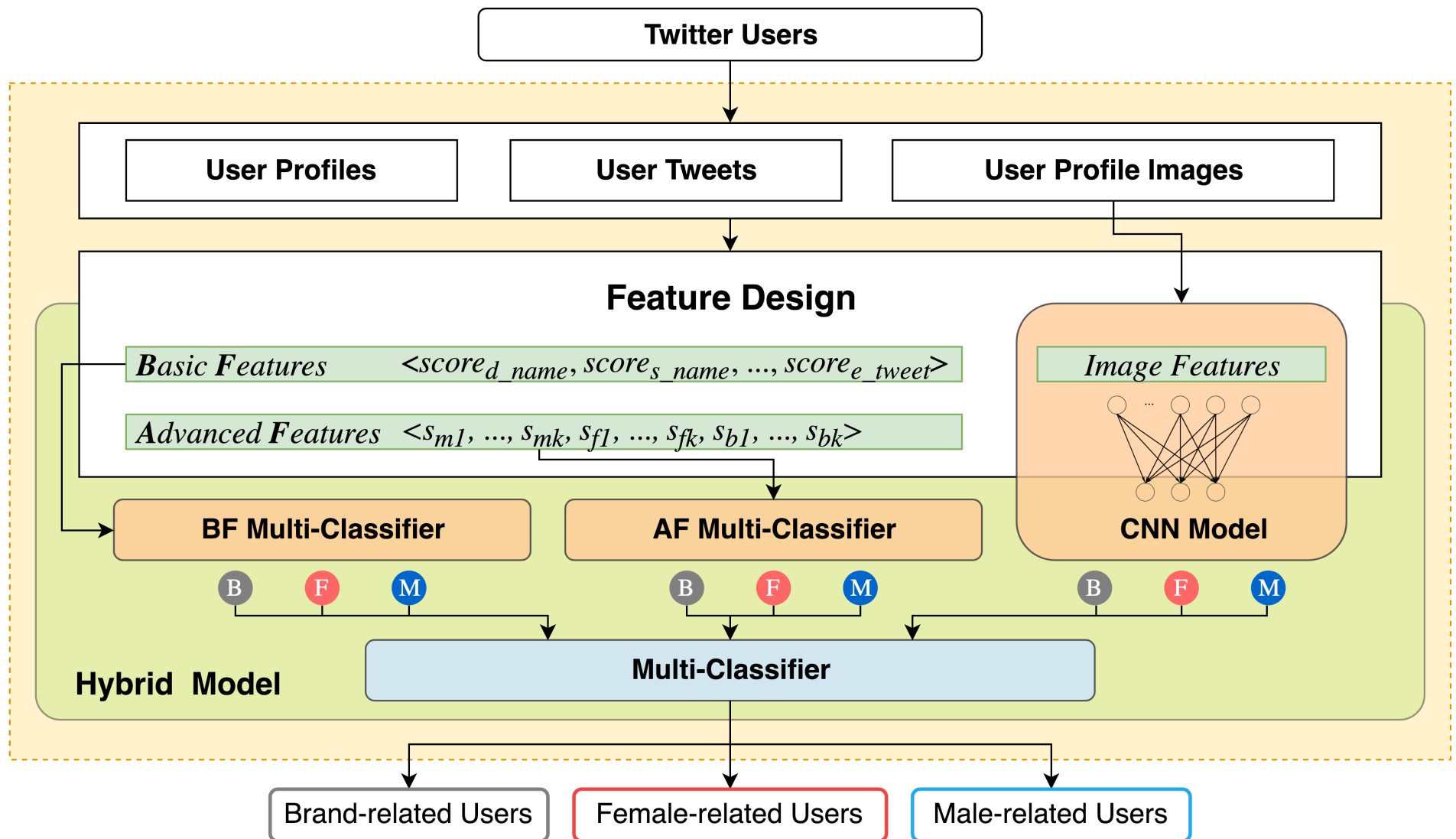


## Data Description

- **Tool:** GetOldTweets3
- **Collection:** Five tweet collections on school shootings
- **Time range:** Covers up to about one month
- **Preprocessing:** Retweets (RTs) and non-English tweets removal

Collection	Start Date	Deaths	Injuries	# of Tweets
Sandy Hook Elementary School shooting	12/14/2012	28	2	97,283
Santa Monica College shooting	06/07/2013	6	4	23,093
Umpqua Community College shooting	10/01/2015	10	8	17,821
Stoneman Douglas High School shooting	02/14/2018	17	17	22,321
Santa Fe High School shooting	05/18/2018	10	14	47,708

## TwiRole Architecture



## TwiRole Online Classification

Webpage Link: <http://vis.dlib.vt.edu:3001>

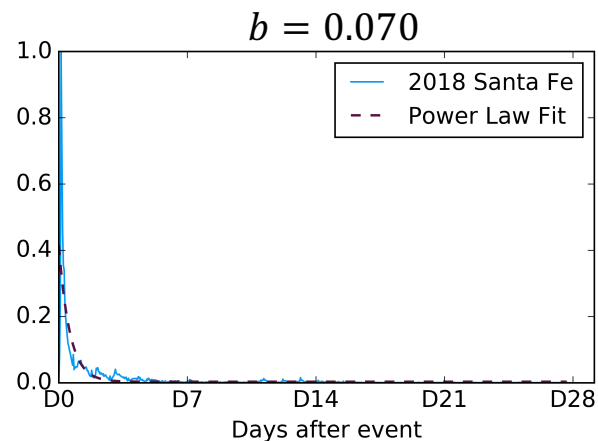
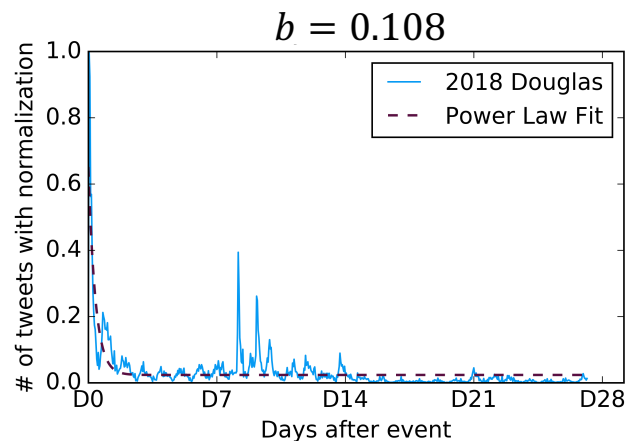
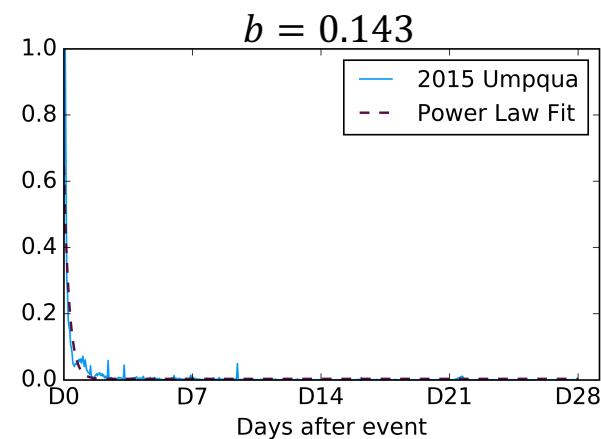
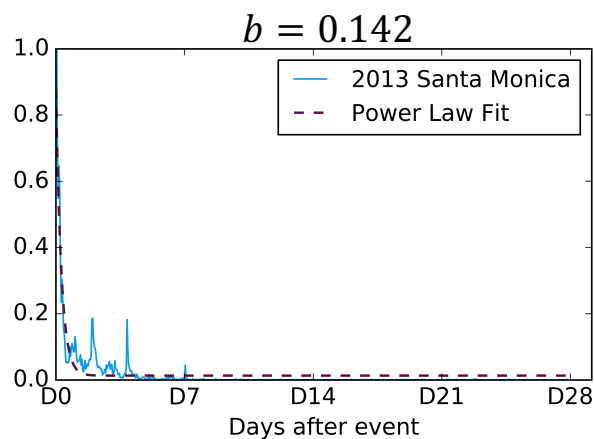
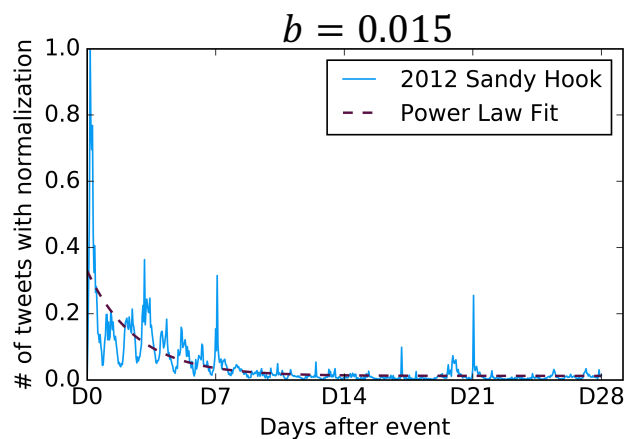
The screenshot displays the TwiRole web application interface. At the top, there is a Twitter logo and the text 'Repository Publication'. The main heading is 'TwiRole'. Below the heading, there is a brief description: 'TwiRole is a hybrid model for role-related user classification on Twitter. It makes use of multiple classifiers and various types of features to predict the roles of users (i.e., female, male, and brand). TwiRole is a graduate research of the Global Event and Trend Archive Research (GETAR) project supported by NSF (IIS-1619028 and 1619371) in Digital Library Research Laboratory (DLRL) at Virginia Tech.'

The interface features a search bar with the text '@LindaEMc' and a blue 'Classify user' button. Below the search bar, three classification results are shown, each with a profile picture, a circular progress indicator, and three horizontal bars representing the confidence scores of the 'Basic Classifier', 'Advanced Classifier', and 'Image Classifier'.

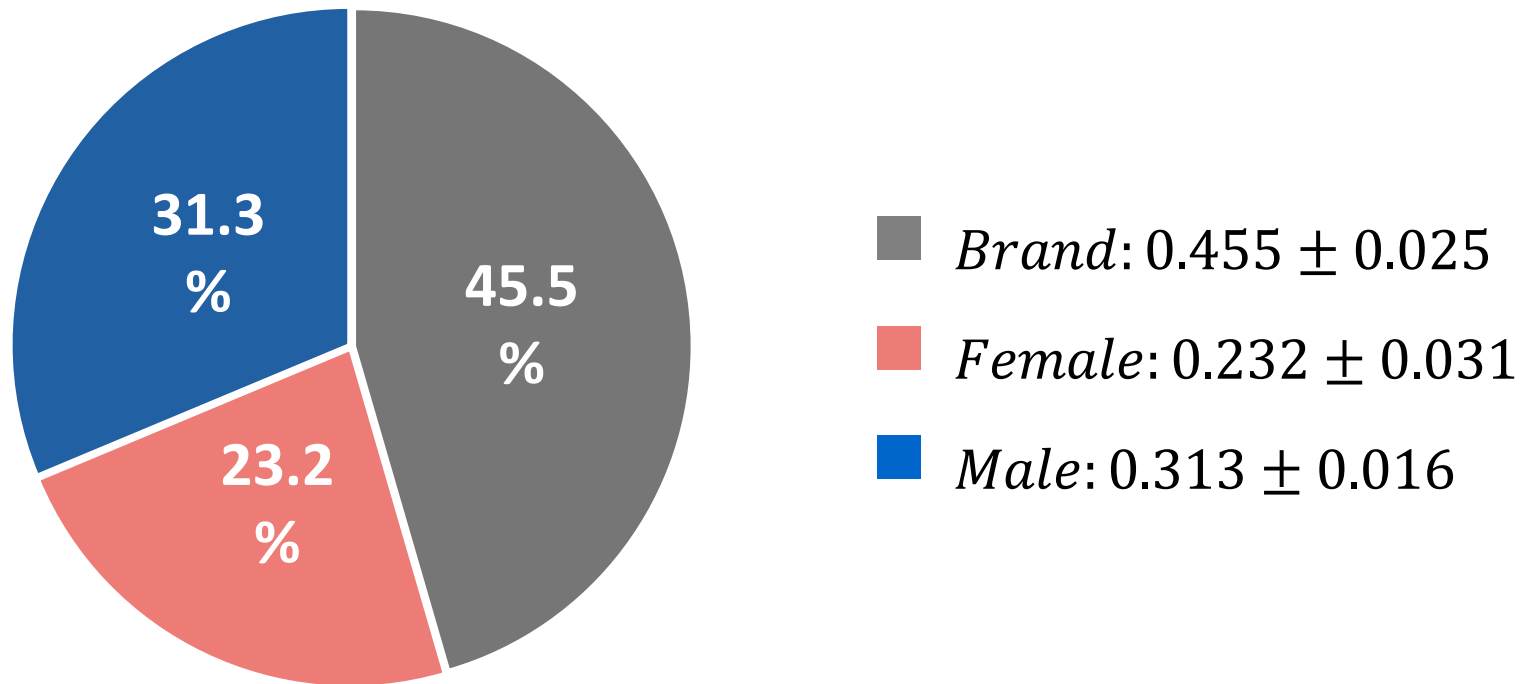
- Female:** The 'Image Classifier' bar is the most prominent, followed by the 'Advanced Classifier' bar. The 'Basic Classifier' bar is the shortest.
- Brand:** The 'Advanced Classifier' bar is the most prominent, followed by the 'Image Classifier' bar. The 'Basic Classifier' bar is the shortest.
- Male:** The 'Image Classifier' bar is the most prominent, followed by the 'Basic Classifier' bar. The 'Advanced Classifier' bar is the shortest.

## Tweeting Patterns

Power Law Fit:  $f(x) = a \cdot e^{-bx} + c$



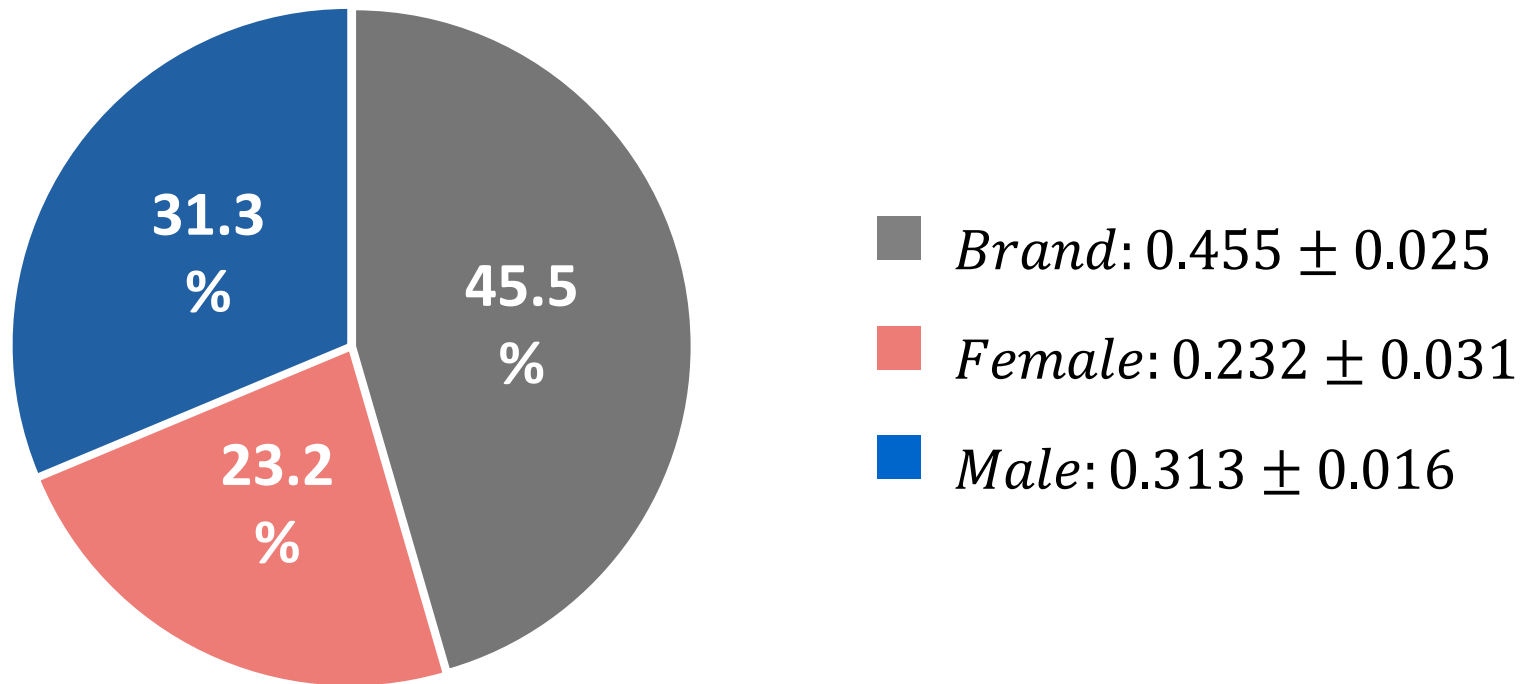
## User Engagement



## Evaluation on Transfer Learning

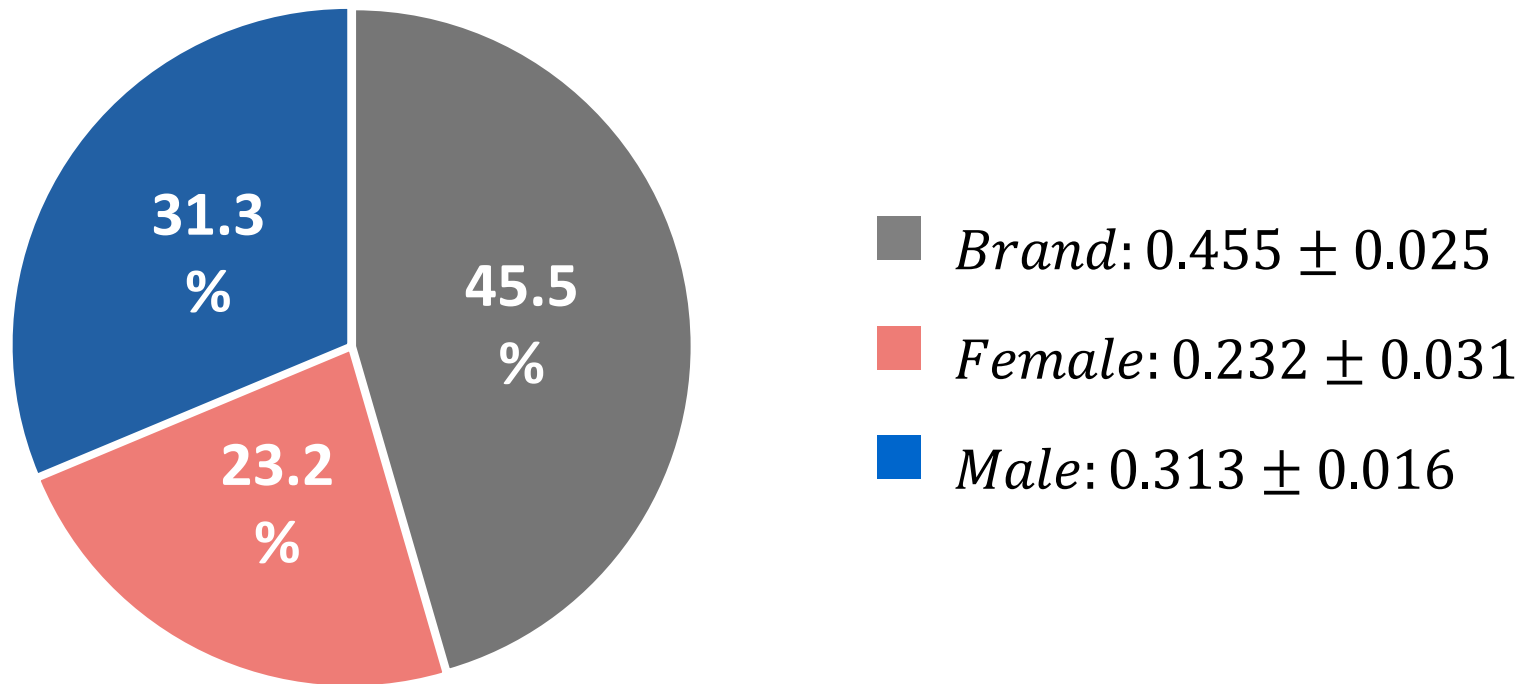
Manually check 100 users per role:  $Acc_{overall} = 82.7\%$

## User Engagement



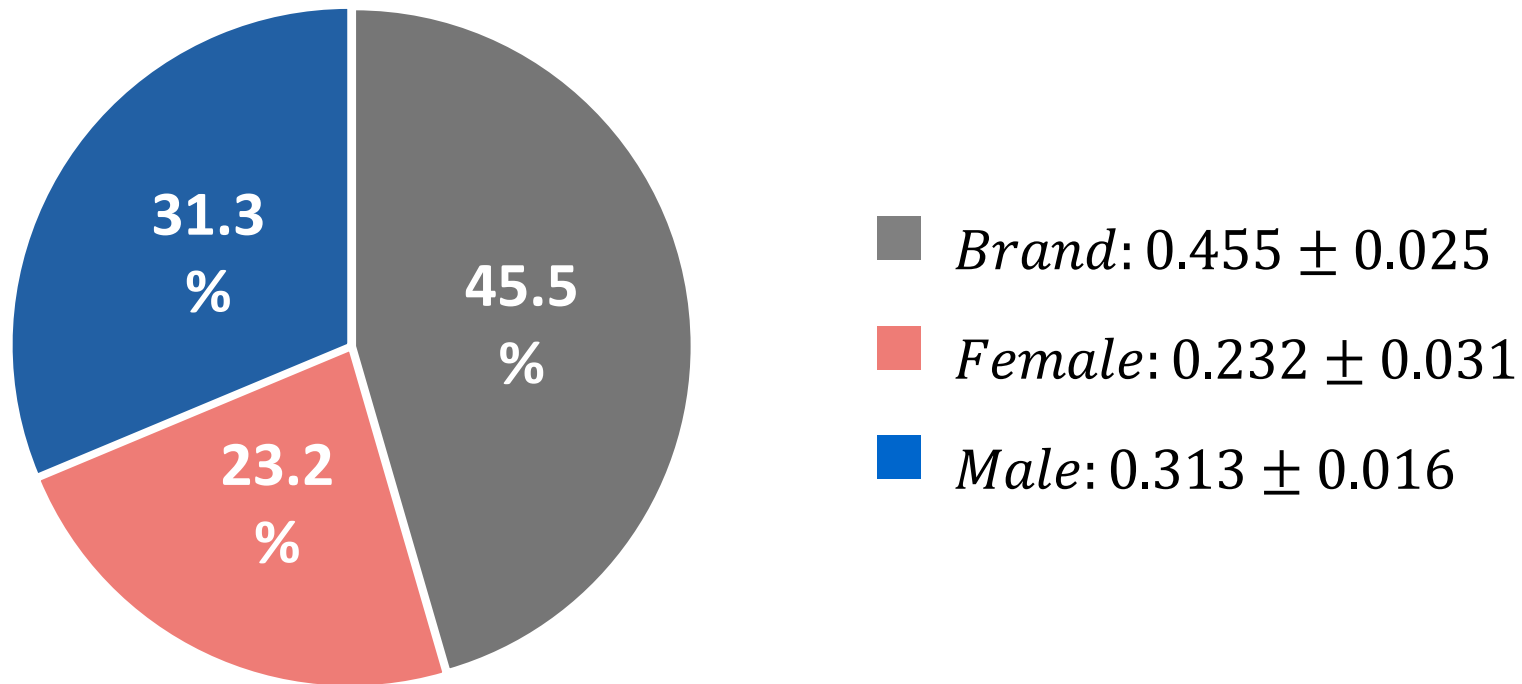
The engagement of different roles of users is relatively consistent among all school shootings.

## User Engagement



**Brand users are the primary group who might publish information or deliver messages.**

## User Engagement



Male users participated more actively than female users in discussions of the selected school shootings.



## Topic Distribution

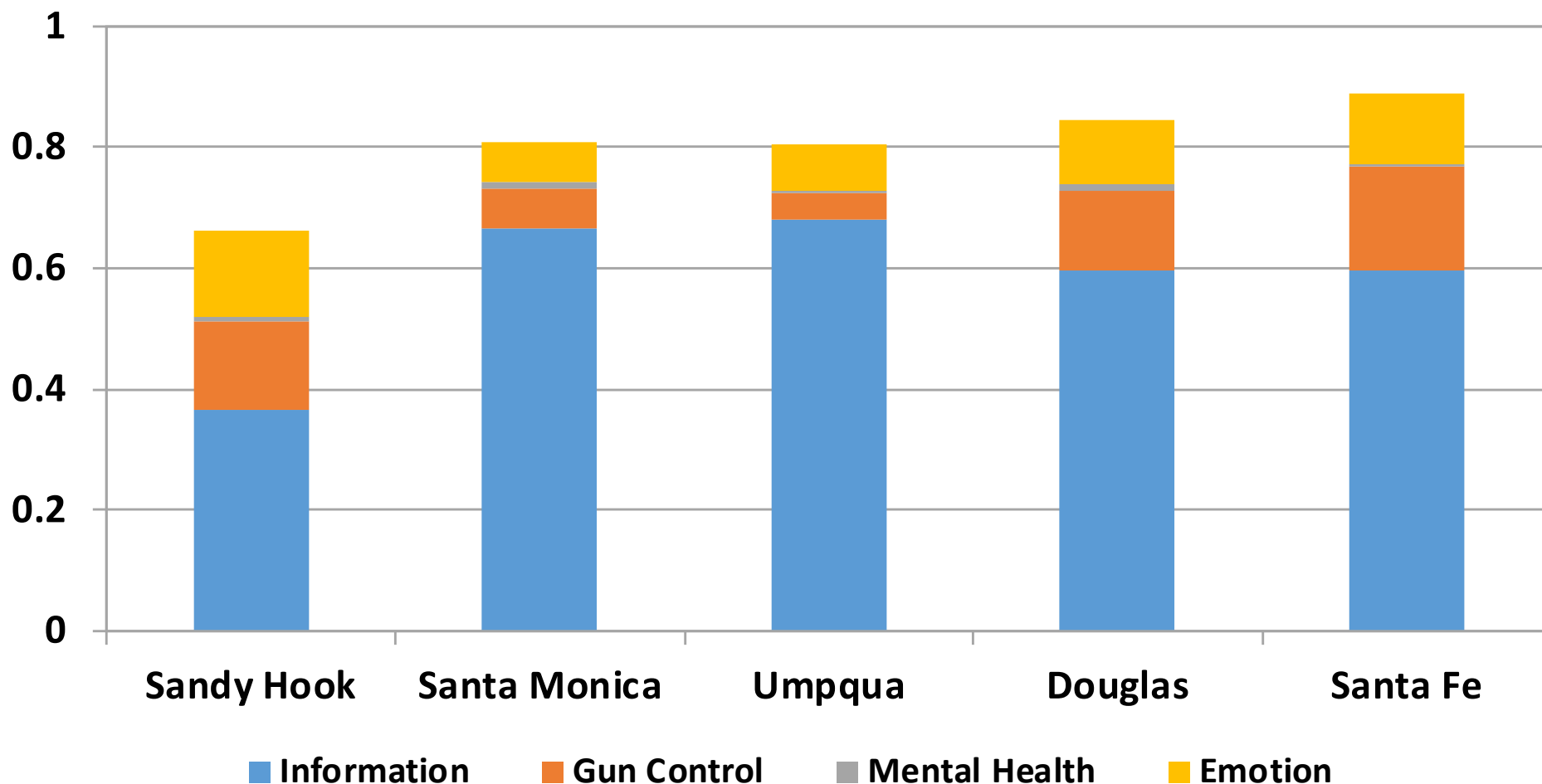
- **Document:** All tweets posted by a labeled user
- **Topic Modeling:** 15 topics generated by TweetLDA
- **Topic Integration:** Combine Zhang et al.'s\* dictionaries and our topics

Category	Typical Words
Information	news, shooter, accidental, gunman, arrest, victim, dead, report, kill, suspect, breaking, injured, custody, wounded, police, fatality, update, fire, campus, student, confirm
Gun Control	gun, control, nra, congress, senate, legislation, safety, violence, illegal, antigun, debate, governor, laws
Mental Health	mental, health, illness
Emotion	absurd, anger, angry, awful, crazy, disgust, frighten, heart, pray, rip, sad, tragic, dedicated, deep, silence

\* Zhang, Y., Wang, Y., Foley, J., Suk, J., and Conathan, D. Tweeting Mass Shootings: The Dynamics of Issue Attention

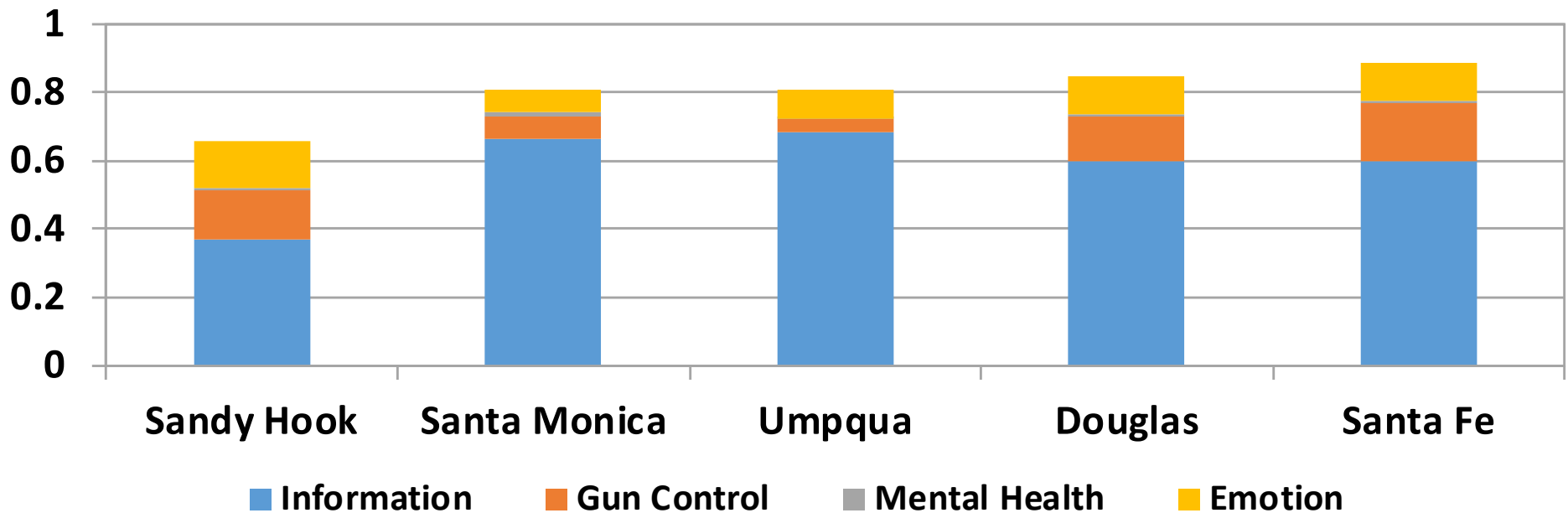
## Topic Distribution

Percentages of different categories among collections



## Topic Distribution

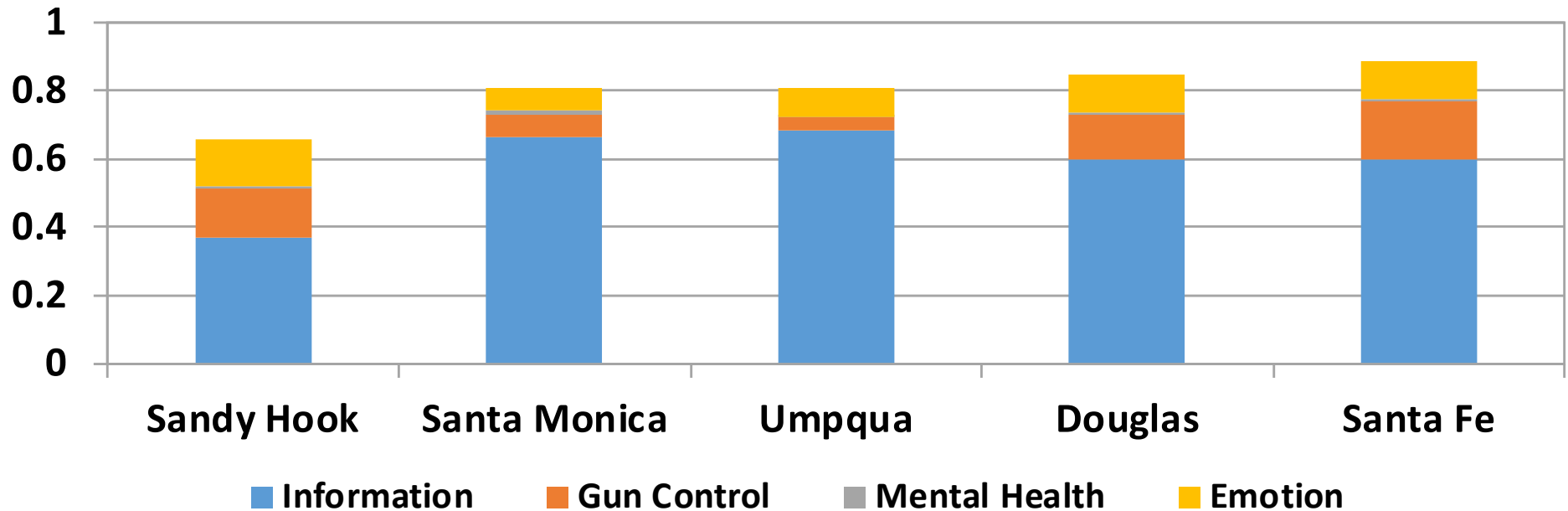
Percentages of different categories among collections



“Information” is the dominant topic in all school shootings.

## Topic Distribution

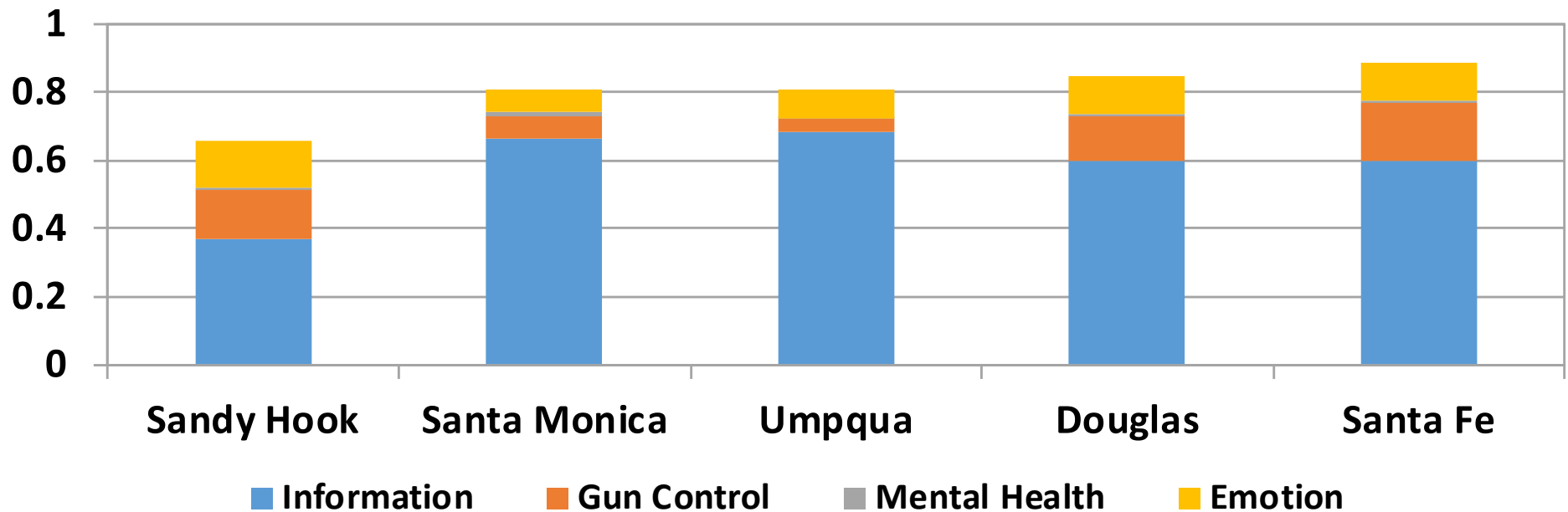
Percentages of different categories among collections



Tweeters posted few tweets about mental health, but focused on the topics related to gun control and emotion.

## Topic Distribution

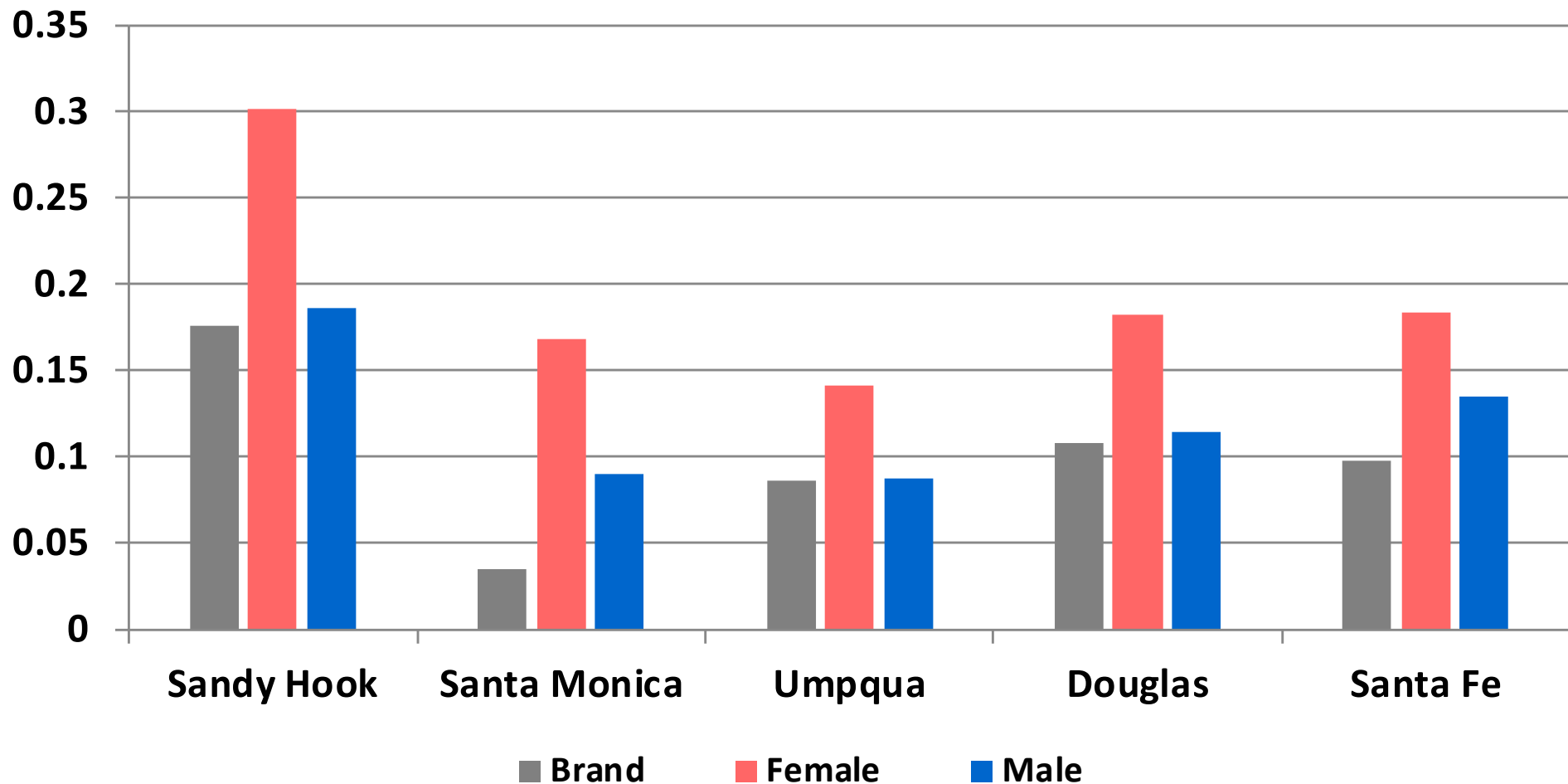
Percentages of different categories among collections



People were more concerned about gun control and expressed their feelings regarding the massacres that had more people dead or injured.

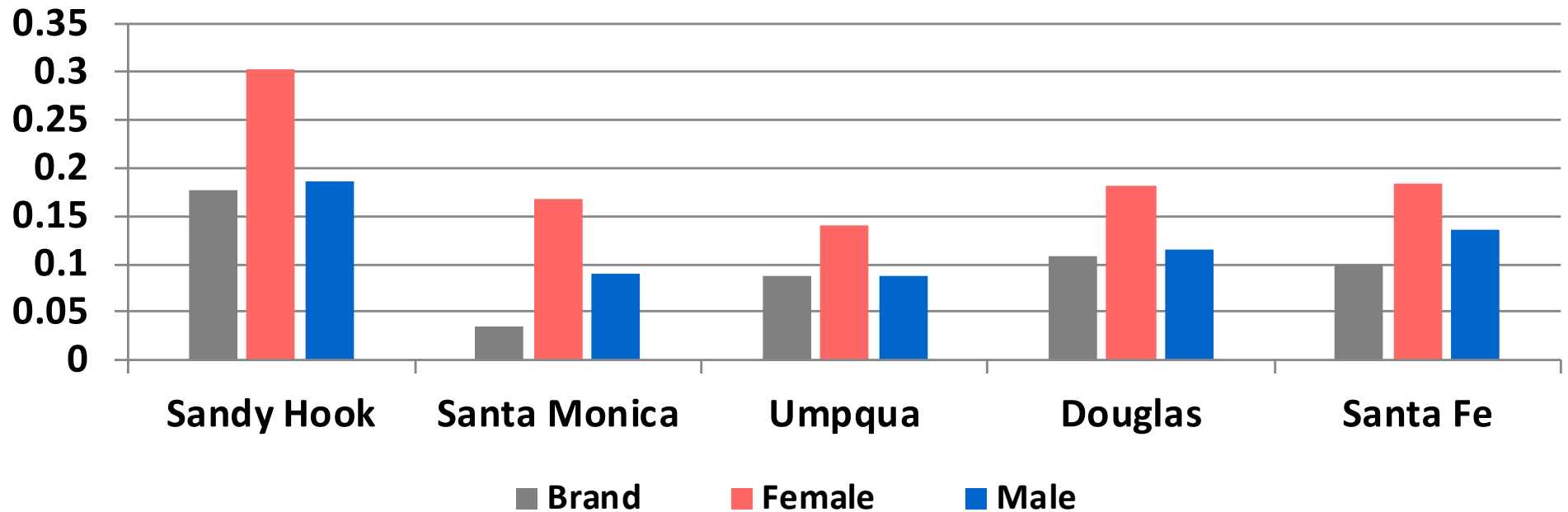
## User Role Differentiation

Role differentiation in the “emotion” topic



## User Role Differentiation

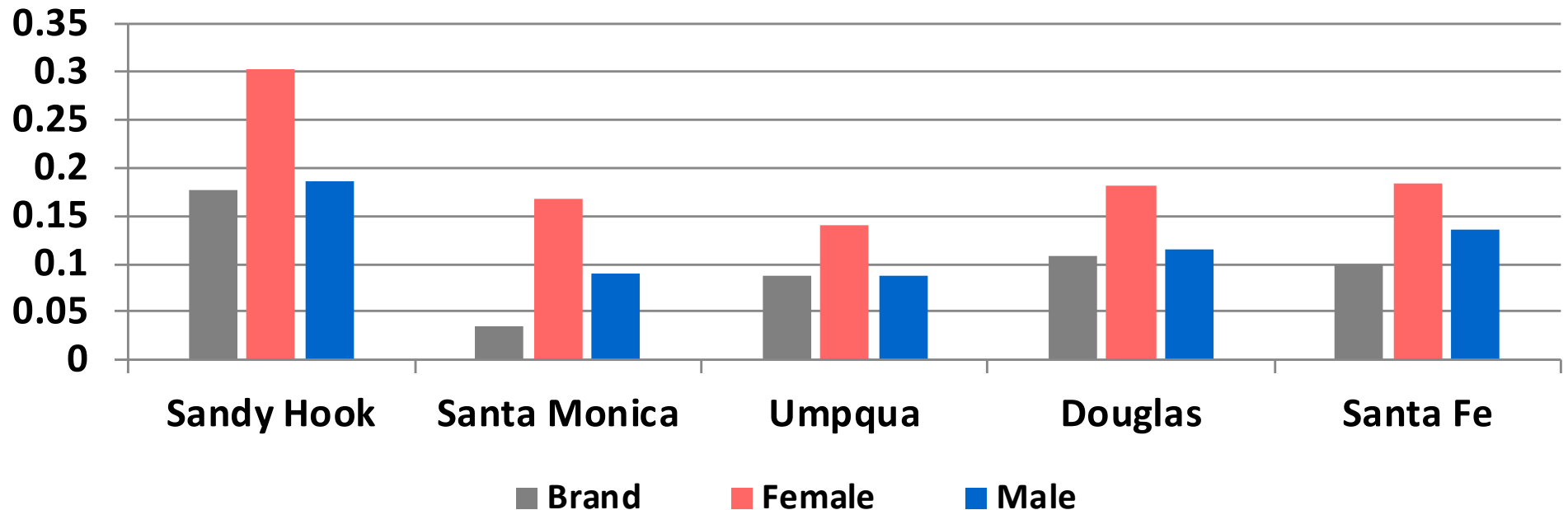
### Role differentiation in the “emotion” topic



Female users were posting more tweets containing emotional words than male and brand users.

## User Role Differentiation

### Role differentiation in the “emotion” topic



A higher percentage of the tweets reflected emotion, for all of the groups of users, in the Sandy Hook Elementary School shooting.



# Conclusion

- A comprehensive analysis of five school shooting collections, considering users, user roles, and topics.
- We observed
  - The relative percentages of user engagement appear stable across collections ( $P_{Brand} > P_{Male} > P_{Female}$ )
  - “Information” is the dominant topic in all school shootings while “mental health” was seldom discussed. Tweeters were more concerned about “gun control”;
  - Tweeters expressed their feelings most when the victim count was highest;
  - Female users posted more emotional tweets than male and brand users.

# Future Work

- **We will improve**
  - the current TwiRole with new features and architecture, leading to more accurate prediction results
- **We will carry out**
  - a temporal analysis to further explore the changes of different emotions across user roles over time
- **We will deploy**
  - our approach on other types of events to help with further discovery of patterns

# Acknowledgements

- Projects
  - U.S. NSF for supporting the Global Event and Trend Archive Research (GETAR) project, through grant IIS-1619028 as well as IIS-1619371 to partner Internet Archive
- Student Team
  - The Bluebirds (TwitterRoleClassification) team in Spring 2019 CS4624

# Thank you !

## Q & A



### Authors

Liuqing Li      Email: [liuqing@vt.edu](mailto:liuqing@vt.edu)      Webpage: <http://liuqing.dlib.vt.edu/main/>

Rishabh Anand      Email: [risha97@vt.edu](mailto:risha97@vt.edu)

Edward A. Fox      Email: [fox@vt.edu](mailto:fox@vt.edu)      Webpage: <http://fox.cs.vt.edu/>



### Resources



TwRole Code      <http://github.com/liuqingli/TwiRole>



TwRole UI      <http://vis.dlib.vt.edu:3001/>



Shooting Datasets      <https://bit.ly/2LdOcej>



GETAR Project      <http://eventsarchive.org/> NSF (IIS-1619028 and 1619371)