Hot issue publics on Twitter:
Analyzing Charlotte protests with computer-assisted text analysis

Ryan Wesslen, Tiffany Gallicano, Jean-Claude Thill & Sagar Nandu
UNC Charlotte
IC²S², July 12, 2017
The Incident
Tuesday, Sept. 20, 2016

The Scott shooting scene, reconstructed
This map shows approximate locations of people and vehicles when Keith Lamont Scott was shot Sept. 20, reconstructed from videos, maps and aerial photos. CMPD identified Brentley Vinson as the officer who shot Scott. Dotted lines show the path of Scott from his car, and of the uniformed officer wearing the body-cam.

Research: Helen Schwab  Mike Homan - mhoman@charlotteobserver.com
Lyric Scott, Scott’s daughter, live streamed as she confronted police about the shooting.
Charlotte shooting: Disabled black man shot dead by police 'was reading book while waiting for son's school bus'

Keith Lamont Scott, 43, was shot by police who say he was armed while getting out of a vehicle
Keith Lamont Scott's wife releases video of deadly Charlotte encounter
Charlotte Police Chief Says Dead Man Had a Gun, Not a Book

Keith Lamont Scott was shot and killed by police in North Carolina on 20 September 2016.

By Brooke Binkowski  Sep 23rd, 2016
Protest Tweet Counts: 48 Hours

- **Tuesday Sept. 20th**
  - Crowd gathers

- **Wednesday Sept. 21st**
  - 8:30pm: Protester Shot
  - Police Deploy Tear Gas

- **Thursday Sept. 22nd**
  - Looting in Uptown Charlotte
  - 11pm: State of Emergency Declared

PR Theory: Hot Issue Publics
What is a public?
People who “begin as disconnected systems of individuals experiencing common problems” with the potential to “evolve into organized and powerful activist groups” (Grunig, 1997, p. 9)
A group of people that engages in activism on...

...a single issue ➔ Single Issue Public (e.g., #BLM)

...many issues ➔ All Issues Public

...a single issue that:

• involves nearly everyone in a population
• has received extensive media coverage

➔ Hot Issues Public
• **Hot issue publics** only become active on an issue when it affects their lives, **often due to locality.**

• Involvement with the issue is more **ego-related** than related to the actual problem itself.
Research Questions

RQ 1. How do hot issue publics’ Twitter topics differ from single/all issue publics?

RQ 2. How do local ties (e.g., city or state level) affect Twitter topics?

RQ 3. What effect does time have on Twitter topics?
Structural Topic Modeling (STM)

• **Topic models** are unsupervised text algorithms to detect *word co-occurrence clusters*.

• **Structural Topic Modeling (STM)** incorporates a regression framework to test the *effect of metadata* (e.g., covariates) on topics.

# What and How Effects on Topics

## Prevalence Covariates: Effect on Topic Proportions (“what”)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Levels</th>
<th>Definition</th>
</tr>
</thead>
</table>
| Publics Type     | Hot Issue, All/Single       | All/Single = tweeted #BLM 14 days prior to incident  
Hot Issue = did not tweet #BLM 14 days prior... |
| Locality         | Charlotte, Carolinas, Non-Local or Missing | • Inferred from profile information.  
• Supervised classifier to predict missing (if possible). |
| Time             | Hour                        | Spline to smooth                                                              |

## Content Covariate: Effect on Word Proportions (“how”)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Levels</th>
<th>Definition</th>
</tr>
</thead>
</table>
| Locality  | Charlotte, Carolinas, Non-Local or Missing | • Inferred from profile information.  
• Supervised classifier to predict missing (if possible). |
Dataset & Pre-Processing

1.3MM tweets over seven days: Sept. 20-26, 2016.

- Seven hashtags: e.g., #KeithLamontScott, #CharlotteProtest
- Gnip Historical PowerTrack: 100% Twitter firehose

Additional firehose data for feature extraction.

- **Hot Issue**: Did not tweet #BLM 14 days prior to protests.
- Predict **Missing Locality**: Classifier based on 14 days prior tweets.

Exclusions:

- **Remove retweets** (83% of data set) due to “echo effect,” problem of attribution, and limited tweet firehose allocation.
- Excluded short, non-English tweets & trending bots (with supervised classifier)
Hypothesis

Hot issue public topics will be short-lived.

- Expect they will be high in peaks (esp. 2nd night) but should quickly subside in days that follow.

- Interaction with locality => Local & Hot Issue

- Alternatively, Non-local & Single/All Issue Topics should be the most persistent.
Hot Issue & Local Topics

- Cluster topics into four categories (k-means).
- Expect Non-Local & Single/All Issue (red) will be the most persistent.
- Charlotte/Hot Issue (purple) expected to be the least persistent.

This example only considers the “Charlotte” effect. See Appendix for “Carolinas”.

Difference Topic Proportions with 99% Confidence Intervals
Color by K-Means Clustering (k = 4)

Non-Local & Hot Issue
- Charlotte/Hot Issue
- Non-Local & Single/All Issue
- Charlotte & Single/All Issue

Local Ties Differences
- Non-Local
- Charlotte
Top 3 topics by hour

Top Topics by Top Five Words

reading + book + disabled + son + sick
black + cop + officer + man + unarmed
drop + gun + planted + weapon + hand
#keithlamontscott + #video + #policeshootings
riots + telling + erupt + calm + also
violence + stop + act + congress + must
looting + rioting + go + country + home
#charlotteprotest + #charlit + marshall
guard + gas + tear + factsbymax + nomypti
obama + thugs + rioters + obamas + blm
release + wife + video + released + releases
#charlotteriots + #msnbc + organization + group
curfew + midnight + #hillary + #charlotteprotests
last + hide + died + bullets + rubber

Timeline created by Microsoft’s Timelines Storyteller
https://timelinestoryteller.com/
Results and Discussion

The most persistent topics tended to be by Non-Locals and Single/All Issue Publics (i.e., previously engaged in #BlackLivesMatter).

However, we find an unexpected persistence of “release the video” topic, which was Local & Hot Issue.

Explore the role local and national organizations (e.g., Color of Change, ACLU) to influence topics with “cluster tweets” (Gallicano, Wesslen and Thill, 2017).
Future Work

Need for additional studies
- “Natural Experiment”: Compare 2014 Charlotte incident vs 2016.
- Other #BlackLivesMatters and non-protests events.

Problems with topic modeling/STM:
- Zero-sum problem of topic models (Fong & Grimmer, 2016)
- Effect of retweets: weighted STM?

Social Network & Information Diffusion
- Can hot issue relate to periphery vs. core? (Barberá et al., 2015)
Questions?

Thanks to Tiffany Gallicano, Jean-Claude Thill, Sagar Nandu, Wenwen Dou and Ali Karduni.


Fong, C., & Grimmer, J. (2016). Discovery of Treatments from Text Corpora. In ACL (1).


Feature Extraction

Local ties extraction on Twitter profile location

1. Use a list of North and South Carolina terms.
2. Location entity recognition & geolocation.
3. Build supervised classifier (CV Ridge Regression) on out-of-sample tweets (10% two weeks prior) to predict labels for missing; remaining are “Missing.”

Hot Issue vs. Single/All Issue Publics

◦ Single/All Issue = Tweeted #BLM two weeks before protest
◦ Hot Issue = Did not tweet #BLM two weeks before protest

<table>
<thead>
<tr>
<th>Count of Tweets</th>
<th>Charlotte</th>
<th>Carolinas</th>
<th>Non-Local</th>
<th>Missing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot Issue</td>
<td>7,860</td>
<td>6,024</td>
<td>81,423</td>
<td>13,404</td>
<td>108,711</td>
</tr>
<tr>
<td>Single/All Issue</td>
<td>739</td>
<td>1,914</td>
<td>37,787</td>
<td>1,888</td>
<td>42,328</td>
</tr>
<tr>
<td>Total</td>
<td>8,599</td>
<td>7,938</td>
<td>118,210</td>
<td>15,292</td>
<td>151,039</td>
</tr>
</tbody>
</table>
## Local Ties Effect on Content

### Topic 7: drop + gun + planted + weapon + hand + marijuana + @thejusticedept

- Carolinas
  - cmpd putney #wral chief investigation #cmpd sbi evidence #releasethetape kerr marijuana gun case drop

- Charlotte
  - putney investigation chief sbi says kerr marijuana case possession handgun evidence information original gun

- Missing
  - hear drop ankle holster didn't planted day empty weapon knew wearing cmpd commands hand

- Not Local
  - gun drop planted evidence holster transparency weapon @thejusticedept putney chief hand scene marijuana planting

### Topic 15: looting + rioting + country + home + protesting + way

- Carolinas
  - home rioting looting work protesting people close you're way really bring country destroying

- Charlotte
  - city home together work people come way looting go close protesting rioting isn't leave

- Missing
  - rioting looting bunch start city country isn't movement cause everyone go way respect acting

- Not Local
  - looting rioting go country protesting home way change destroying like property you're work animals
Non-Local vs Carolinas/Missing
Predicting Missing Locality

- Used 5-fold Ridge Regression to predict users with missing profile location.

- Built two models: Charlotte and Carolinas. Used the “known” labels and held out 20% for testing.

- Achieved 90%+ accuracy but low precision/recall.

<table>
<thead>
<tr>
<th>Carolinas</th>
<th>FALSE</th>
<th>TRUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE</td>
<td>7,325</td>
<td>167</td>
</tr>
<tr>
<td>TRUE</td>
<td>295</td>
<td>81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Charlotte</th>
<th>FALSE</th>
<th>TRUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE</td>
<td>6,948</td>
<td>411</td>
</tr>
<tr>
<td>TRUE</td>
<td>348</td>
<td>161</td>
</tr>
</tbody>
</table>

### Carolinas Top Features

<table>
<thead>
<tr>
<th>Rank</th>
<th>Coeff</th>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.600672</td>
<td>clt</td>
</tr>
<tr>
<td>2</td>
<td>0.589524</td>
<td>raleigh</td>
</tr>
<tr>
<td>3</td>
<td>0.438203</td>
<td>mccrory</td>
</tr>
<tr>
<td>4</td>
<td>0.436338</td>
<td>durham</td>
</tr>
<tr>
<td>5</td>
<td>0.410346</td>
<td>@coliegestudent</td>
</tr>
<tr>
<td>6</td>
<td>0.400823</td>
<td>#repealhb2</td>
</tr>
<tr>
<td>7</td>
<td>0.393499</td>
<td>#asheville</td>
</tr>
<tr>
<td>8</td>
<td>0.390386</td>
<td>uptown</td>
</tr>
<tr>
<td>9</td>
<td>0.381046</td>
<td>@wbtv_news</td>
</tr>
<tr>
<td>10</td>
<td>0.379816</td>
<td>ecu</td>
</tr>
<tr>
<td>11</td>
<td>0.372657</td>
<td>shrimp</td>
</tr>
<tr>
<td>12</td>
<td>0.352878</td>
<td>@roycoopernc</td>
</tr>
<tr>
<td>13</td>
<td>0.347727</td>
<td>tackling</td>
</tr>
<tr>
<td>14</td>
<td>0.343472</td>
<td>triad</td>
</tr>
<tr>
<td>15</td>
<td>0.326004</td>
<td>trippin</td>
</tr>
<tr>
<td>16</td>
<td>0.319209</td>
<td>unc</td>
</tr>
<tr>
<td>17</td>
<td>0.312744</td>
<td>brackeen</td>
</tr>
<tr>
<td>18</td>
<td>0.307389</td>
<td>@cameronnewton</td>
</tr>
<tr>
<td>19</td>
<td>0.300823</td>
<td>@snapchat</td>
</tr>
<tr>
<td>20</td>
<td>0.298132</td>
<td>charlotte</td>
</tr>
</tbody>
</table>

### Charlotte Top Features

<table>
<thead>
<tr>
<th>Rank</th>
<th>Coeff</th>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.125938</td>
<td>clt</td>
</tr>
<tr>
<td>2</td>
<td>0.920002</td>
<td>uptown</td>
</tr>
<tr>
<td>3</td>
<td>0.711054</td>
<td>@coliegestudent</td>
</tr>
<tr>
<td>4</td>
<td>0.708484</td>
<td>@wbtv_news</td>
</tr>
<tr>
<td>5</td>
<td>0.694427</td>
<td>mod</td>
</tr>
<tr>
<td>6</td>
<td>0.679505</td>
<td>parkway</td>
</tr>
<tr>
<td>7</td>
<td>0.559287</td>
<td>charlotte</td>
</tr>
<tr>
<td>8</td>
<td>0.548697</td>
<td>gemini</td>
</tr>
<tr>
<td>9</td>
<td>0.53564</td>
<td>sized</td>
</tr>
<tr>
<td>10</td>
<td>0.529665</td>
<td>hannah</td>
</tr>
<tr>
<td>11</td>
<td>0.524252</td>
<td>tt</td>
</tr>
<tr>
<td>12</td>
<td>0.522303</td>
<td>@cameronnewton</td>
</tr>
<tr>
<td>13</td>
<td>0.520000</td>
<td>debuted</td>
</tr>
<tr>
<td>14</td>
<td>0.517259</td>
<td>reservation</td>
</tr>
<tr>
<td>15</td>
<td>0.515416</td>
<td>menace</td>
</tr>
<tr>
<td>16</td>
<td>0.513435</td>
<td>#charobs</td>
</tr>
<tr>
<td>17</td>
<td>0.50302</td>
<td>ashes</td>
</tr>
<tr>
<td>18</td>
<td>0.499604</td>
<td>plugs</td>
</tr>
<tr>
<td>19</td>
<td>0.491436</td>
<td>oak</td>
</tr>
<tr>
<td>20</td>
<td>0.489624</td>
<td>tailgate</td>
</tr>
</tbody>
</table>
Cluster Tweets

- Barbara Fister: 
  #TerenceCrutcher, #KeithLamontScott...the violence must stop. Tell Congress to act. aclu.org/preventtragedi... You can sign this ACLU letter.

- Mindy Crouchley: 
  Demand justice after police shootings now! #KeithLamontScott #TerenceCrutcher action.momsrising.org/sign/scott-cru ...

- Fredric Wertham: 
  North Carolina police shouldn’t be able to hide the footage of #KeithLamontScott’s murder. Take action:

---

Graph showing activity from 20 Sep to 26 Sep with peaks for 
- ACLU
- MomsRising
- Color of Change
Number of Topics Evaluation

Held-Out Likelihood

Residuals

Semantic Coherence

Lower Bound

Number of Topics (K)

Number of Topics (K)

Number of Topics (K)

Number of Topics (K)
Racial Diversity

Green: African-Americans

Blue: Caucasians

Source:
2010 US Census
Racial Dot Plot Map
University of Virginia
Weldon Cooper Center
Charlotte ranks last in upward mobility study

The study ranked Charlotte dead last among the country's 50 largest cities for upward mobility, meaning it is harder to climb out of poverty here than in any other major U.S. metro area. The odds of moving up in Charlotte - just 4.4%.
