# A Computational Model of Dynamic Group Formation on Social Media

Nick V. Flor(✉) and Erin D. Maestas

University of New Mexico, Albuquerque, NM 87131, USA
{nickflor, emaestas00}@unm.edu

**Abstract.** We need process theories of how groups form dynamically on social media in response to news events. We perform an analytic ethnography of a case where groups formed on social media in response to incomplete news of an incident. The groups framed the incident in terms of existing narratives and called for action against those actors they perceived as the aggressors. Later, footage showed this framing to be inaccurate. Based on the analytic ethnography, we propose a computational model of how groups form in response to incomplete or inaccurate reports on social media.

**Keywords:** Dynamic group formation · Viral messages · Analytic ethnography

## 1 Introduction

We lack both process theories [1] and computational theories about how groups form dynamically on social media. A better understanding of this dynamic formation will allow developers of policy, information assurance, or security to better leverage the beneficial capabilities of such groups, or to design interventions that constrain the formation and activity of groups with detrimental capabilities.

To help us craft such a theory, we studied viral messages surrounding a contentious event on the social media platform Twitter known as the Covington Catholic controversy [2]. In this event, Twitter users widely distributed and criticized a picture of a high school student, wearing a red cap with President Trump's campaign slogan Make America Great Again ("MAGA"), who appeared to be smirking at an elder Native American man with a drum (refer to Fig. 2). Behind the smirking student was a large group of students many of whom were also wearing MAGA caps. Early tweets (postings) framed the students as disrespecting the Native American elder, and many users called for identifying the students, and also for punishing the students, their parents, and their school in various ways. This criticism continued for several days, until later investigation by the news media revealed that this early framing was incorrect and that much of the criticism was unwarranted [3].

What allowed this incorrect narrative to persist for several days? To help answer this question we performed an analytic ethnography of viral messages surrounding the Covington Catholic event on Twitter, with the goal of understanding the different kinds of groups spreading messages on Twitter and their dynamic formation. Distributed cognition [4] forms the theoretical foundation for this research.


G. Meiselwitz (Ed.): HCII 2019, LNCS 11578, pp. 324–335, 2019.
https://doi.org/10.1007/978-3-030-21902-4_23

## 2 Analytic Ethnography

Our specific analytic ethnography consisted of identifying and characterizing the top five to six shared postings (retweets) over a period of six hours that contained the hashtag #CovingtonCatholic. To triangulate our characterizations, we performed frequency analyses on the words, users, mentions, and hashtags, and then used word clouds to depict these frequencies.

### 2.1 Hour 2PM–3PM (18-Tweets)

Figure 1-left depicts the very first tweet our scraper returned that contained the hashtag #CovingtonCatholic by user @KaySch10. This tweet criticized the school's slogan and insulted a student, while hash tagging the state of Kentucky, and Kentucky Senator Mitch McConnell who serves as the Senate Majority Leader.


**Fig. 1.** The first tweet containing the hashtag #CovingtonCatholic (left), and an expanded view of the embedded tweet (right).

While this tweet did not have many shares, it contained an embedded quote ("quote-tweet"), requesting the identity of the student and berating the student as "a POS disrespectful MAGA loser that is gleefully bothering a Native American Student" (see Fig. 1-right), where MAGA is an acronym for President Trump's 2016 campaign slogan Make American Great Again, and where people who wear these hats are viewed as supporters of the President.

The quote-tweet was shared times 6,055 times, which represented 12.08% of @IndivisibleNet's followers. In turn, this quote-tweet contained itself a quote-tweet by user @2020fight (see Fig. 2-left), who identified the boy as a "MAGA loser gleefully bothering a Native American protestor".

This quote-tweet was retweeted 14,490 times, which represented 36% of @2020Fight's total followers. Finally, @2020Fight's tweet contained a reply thread that was shared 3,931 times, which represented 605% of @lulu_says2's followers (see Fig. 2-right). The reply thread reframes the situation from an individual bothering a Native American protester, to an entire mob displaying "ignorance, racism & disrespect" towards the Native American protestor (see Fig. 3).

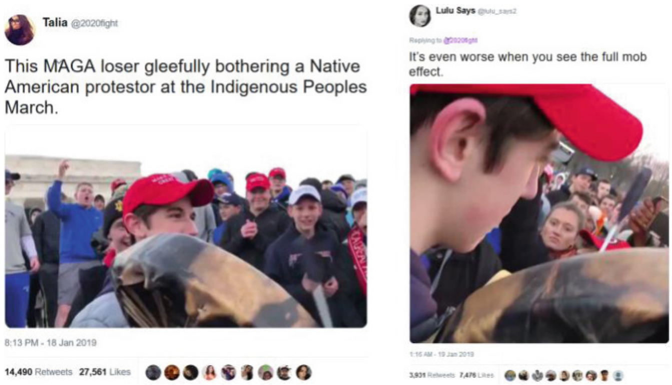


**Fig. 2.** User @2020Fight's Quote Tweet (left), and the first reply in the response thread (right). (Color figure online)

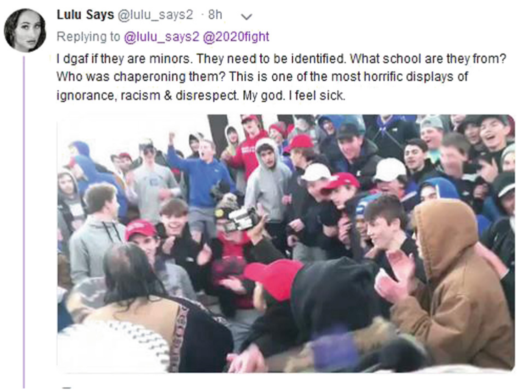


**Fig. 3.** A portion of @lulu_says2's response thread

The reply thread also called for identifying the individuals in the picture, the school they were from and, lastly, who their chaperones were.

Thus, the very first tweet containing the hashtag #CovingtonCatholic, if you follow the quote-tweets and responses contained a complex set of emotions, information, intent, attribution, and requests for action, the latter calling for the identification of the minors, their chaperones, and the school they attended.

Figure 4 depicts a word cloud of the tweets, the participants, and the most viral tweet for the hours of 2PM–3PM UTC (Coordinated Universal Time). Our Twitter scraper collected 18 tweets during this time. The tweet word cloud suggested that users discussed topics including the participants going to the March for Life, the kids, and their chaperones. The most active user was @PiattPatti, and the most shared tweet was a threat to reevaluate a user's grandson going to the school, and a request for apologies and comments from @CovCathColonels—the school's Twitter account.

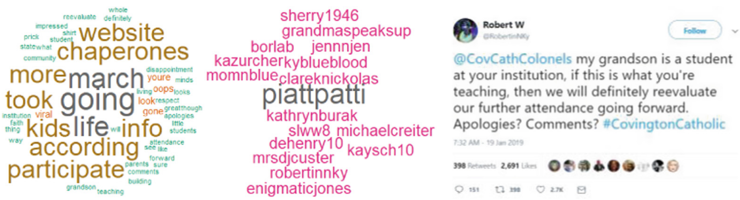


**Fig. 4.** Wordclouds of the tweets, the users, and the most viral tweet from 2–3 PM

## 2.2 Hour 3PM–4PM (182-Tweets)

By analyzing the top-5 tweets shared (retweets) during this time period, we can infer not only the topics discussed but how the topics were evolving relative to the previous hour (refer to Fig. 5). The top retweet by @KAZurcher provided a reason for the students being in DC, and as subtext blamed the chaperones for not supervising the students during the incident. The #2 retweet portrayed the school as anti-gay, and unlikely to discipline the students. The #3 retweet contained three calls to action: release the names of the students, make parent information public and, for colleges, to reject admission to the students. Retweet #4 noted that the topic had gone viral, embedding a tweet from an influencer (blue checkmark near their name) that linked to @2020Fight's tweet (see Fig. 2-left). Retweet #5 is the top tweet from the previous hour.


**Fig. 5.** Top-5 Retweets for the 3–4 PM time period, from left-to right, top-down.

Figure 6 depicts the word clouds for tweet words, users, mentions, and hashtags. The words and the hashtags suggest that the focus of the discussion for the 3PM–4PM h is on exposing the school, and Christian schools in general, as places where students learn to hate, and that there must be consequences for this behavior. The users who tweeted the most were @blinksup and @chileman55. The most mentioned users were @covcathcolonels, who the users were directing their outrage towards, and @2020fight who users were either retweeting or replying to.

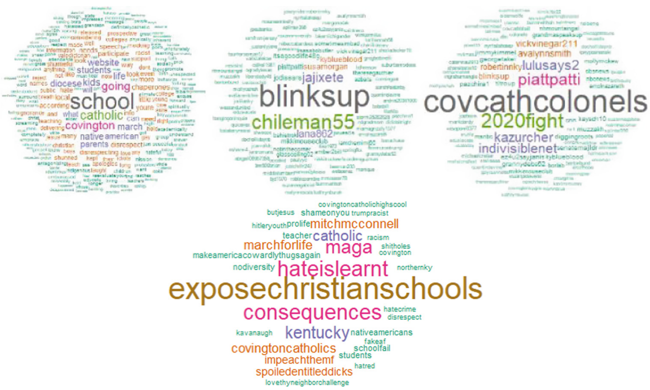


Fig. 6. Word clouds for words, users, mentions, and hashtags (3 pm–4 pm).

## 2.3 Hour 4PM–5PM (605 Tweets)

The top retweets for the 4PM–5PM time slot (see Fig. 7) indicated that users were discussing the mocking behavior of the boys towards the Native America elder (#1, #3), while questioning the school's ability to teach students higher values such as respect (#4–#6). User @PiattPatti contrasted the students' behavior towards the elder, with the same behavior towards a priest, asking users to imagine what would happen to the students if they behaved this way during mass.


Fig. 7. Top-5 Retweets for the 4–5 PM time period, from left-to right, top-down.

The word clouds (see Fig. 8) corroborate the top retweets, with the terms *school*, *boys*, *man,* and *elder* being in the top-tier of most frequent words, and in the second tier, *respected*, *veteran*, *ritual,* and *mass*. The most common hashtag was #ExposeChristianSchools, indicating topics similar to top retweets #4–#6 (Fig. 7) that suggest alleged hypocrisy in Christian schools.

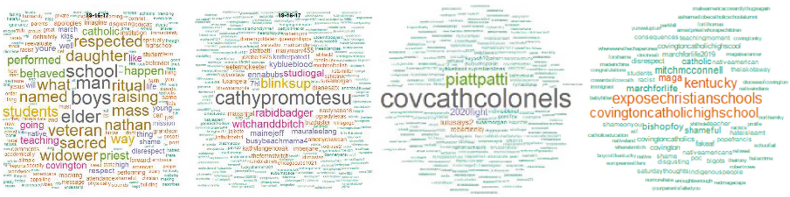


**Fig. 8.** Word clouds for words, users, mentions, and hashtags (4 pm–5 pm).

The most active user was @cathypromotesu, with the previous timeslot's most active user, @blinksup, falling into the second tier of most active users. The main user mentioned was @CovCathColonels who users continued to direct their outrage towards, and @PiattPatti who had two of the top three retweets for this timeslot.

## 2.4 Hour 5PM–6PM (798 Tweets)

Four of the top six retweets from the 5PM–6PM timeslot (see Fig. 9) were repeats from previous timeslots, suggesting that some of the discussion was stabilizing around those topics. Retweet #4 labels the students racist and an embarrassment to the religion. Of particular note is retweet #5, which provides phone numbers to Covington Catholic High School, and to the area's diocese, as well as an e-mail and a physical address for the school:

*RT @kyblueblood: Their school\nCovington Catholic High School\nPhone (859) 491-2247\nFax (859) 448-2242\n\nTheir diocese.\nCovington Catholic Diocese \nPhone: (859) 392-1500\nEmail: info@covdio.org\n1125 Madison Ave.\nCovington, KY 41011-3115\n#CovingtonCatholic* — Top Retweet #5, Deleted.


**Fig. 9.** Top-6 Retweets for the 5–6PM time period, from left-to right, top-down.

Both the words cloud and the hashtags cloud, indicate the main topic focus was the school (see Fig. 10). The most mentioned user was once again @CovCathColonels, the school's social media account, with user PiattPatti—who had several top retweets—in the second tier. The user who spread the tweet with the phone and the contact numbers (@KyBlueBlood) was the most active user.

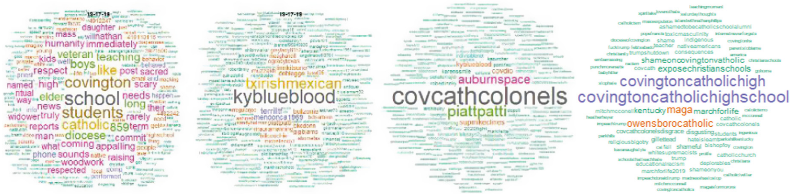


**Fig. 10.** Word clouds for words, users, mentions, and hashtags (5 pm–6 pm).

## 2.5 Hour 6PM–7PM (1315 Tweets)

The 6PM–7PM timeslot indicated a shift in main topic, where the top-retweet contained contact information for Covington Catholic High School and the Covington Diocese, including phone numbers, e-mails, and mail addresses. Three of the top-6 retweets were from the previous timeslot (#2, #5, #6), showing overlap in discussion. Retweet #3 described another activist group @LPJLeague that confronted the students at the rally. Finally, retweet #4 contained an embedded tweet of the Native American elder giving his account of what happened at the rally (Fig. 11).


**Fig. 11.** Top-6 Retweets for the 6 pm–7 pm time period, from left-to right, top-down.

The words cloud (see Fig. 12) aligned with the content of the top retweet. The most prominent word was Covington, with the phone numbers in the lower tiers suggesting the spread of these numbers to other users so that they could contact the school with their complaints. A new most-active user emerged, @AnitaThom57, with the previous most active-user falling into the second tier. Hashtags about Covington Catholic are the

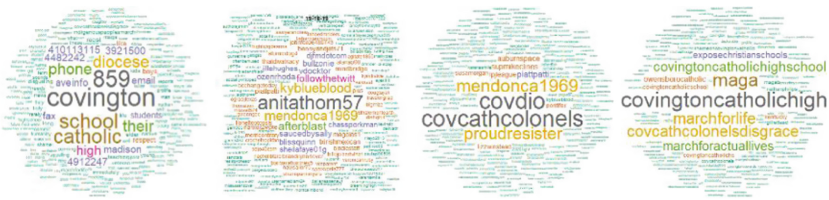


**Fig. 12.** Word clouds for words, users, mentions, and hashtags (6 pm–7 pm)

most prominent in the hashtags cloud, but several “march” hashtags gained prominence, specifically *MarchForLife* and *MarchForActualLives*.

## 2.6 Hour 7PM–8PM (1590 Tweets)

The first three most viral retweets were the same as the previous time slot. Retweet #5 had appeared in several timeslots in the past and re-emerged after not being in the top 5 in the previous time slot. Retweet #4 was a unique call to action from a Twitter influencer, which suggested Covington Catholic teachers should be fired for not watching their students (Fig. 13).


**Fig. 13.** Top-5 Retweets for the 7 pm–8 pm time period, from left-to right, top-down.

The words cloud again aligned with the top retweet, with *Covington School* in the first tier of words, *859*, the area code of the school and the diocese in the second tier, and the phone numbers in the later tiers. There was a new most active user, and @CovCathColonels remained the most mentioned user (Fig. 14).

Of note in the hashtags cloud is the appearance of MAGA in the top-tier. While this hashtag could signal the arrival of MAGA supporters in the discussion, an examination of the tweets containing #MAGA show that non-MAGA supporters were using the hashtag to criticize MAGA supporters (Fig. 15).

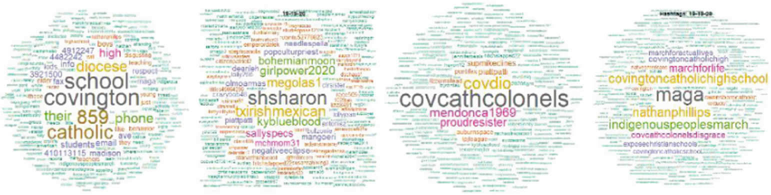


**Fig. 14.** Word clouds for words, users, mentions, and hashtags (7 pm–8 pm).


**Fig. 15.** A sample of critical Tweets containing #MAGA (7 pm–8 pm).

# 3 Towards a Computational Model

We summarize the results above using both a UML use case diagram and a UML communication diagram. First, we defined the different kinds of actors in the system, based mainly on: (a) the type of information posted; (b) the source of this information, which is largely inferred from the posting; (c) the target for the information, determined by mentions, hash tags, or posting content; and (d) any beliefs intended by the actor who posted the information. We refer to this as a *functional-intentional* categorization of actors.

## 3.1 Actors

Based on function and intention, we identified at least 9 different kinds of actors in our analytic ethnography. Given a news event consisting of people or groups in which a user can identify ideological allies or enemies, these actors are:

- Instigators—users who first comment on a news event, frame the event in terms of existing narratives, and include a unique hashtag for organizing further discussion of the event. An instigator can also provide links to more information about the event.
- Investigators—users who collect information on the web about the people or the groups in a news event, then post their findings along with any questions that arise during their investigations so that others can search for the answers. They may also post predictions about the behaviors of the people in the event based on their findings.
- Attributers—users who add specific information about people or groups in an event, as well as their actions at the event, intended to make self-identified allies in the

event appear more sympathetic, and to make enemies appear more immoral. They do so by confirming stereotypes about allies and enemies.

- Doxxers—users who post contact information about people or groups in an event who the user identifies as enemies.
- Punishers—users who post descriptions about the people in the event who should be punished, what their punishments should be, and who should execute the punishments.
- Confronters—users who send messages directly to the people or the groups at an event who they identify as enemies. Such messages include telling the people what they must do to atone for their actions; and demanding apologies or answers while mentioning potential consequences for actions.
- Debaters—users who post about the hypocrisy between a group's stated beliefs and their actions.
- Counterfactualists—users who post imaginary scenarios, typically to make immoralities in behaviors explicit.
- Eyewitnesses—users at the event who post their experiences.

### 3.2 Use-Case Analysis

Figure 16 depicts the nine different actors described previously, and their primary actions.

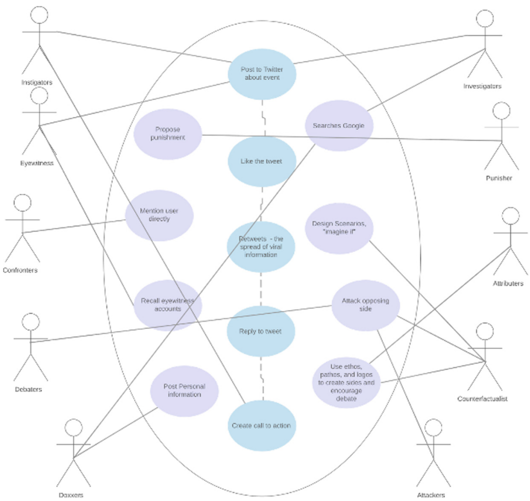


**Fig. 16.** Use case analysis for the nine actors.

### 3.3 Communication Diagram

Figure 17 depicts the communication diagram for the nine actors. Also included is an object for the event, which contains both ally and enemy objects.

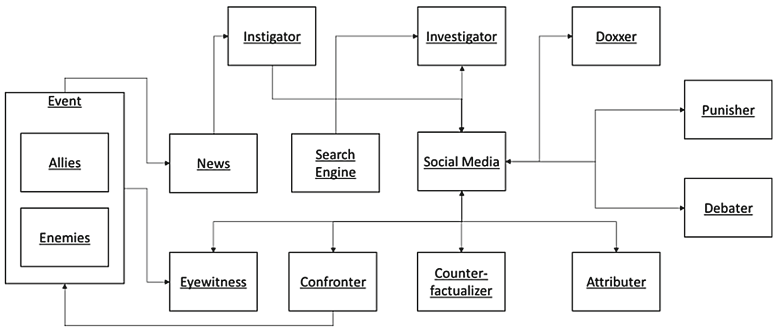


**Fig. 17.** Communication diagram.

## 4 Discussion

Based on an analytic ethnography, we identified nine distinct types of actors responsible for the volume of tweets surrounding the #CovingtonCatholic controversy on the Twitter social media platform. The nine actors are: instigators, investigators, doxxers, punishers, debaters, attributers, counterfactualizers, confronters, and eyewitnesses. We based the actor categories on function—the type of information posted, the source of the information, and the target for the information—as well as intent, viz., beliefs the posting conveyed to other users. Social media users can switch from any of the nine actor roles. For example, a user that acted as an investigator for one posting, may switch roles and act as a punisher for another posting.

The nine different types of actors were responsible for the top retweeted messages in the six-hour period spanning 2PM–8PM on January 19, 2019, which is the day after the March for Life event occurred.

We can view the users that retweeted the same posting as a group that are all trying to spread the same message across their individual networks. Furthermore, each message is associated with a distinct type of actor. Thus, by studying the change in top retweets, we can understand the shift in actor types, and therefore the dynamic formation of groups over time based on the function and intent of the viral messages.

Further research is necessary to determine a more comprehensive set of use cases and set of actors, and to determine the properties and methods of each actor-object in the communication diagram. This will allow us to implement the computation model, perhaps developing social media bots that execute the different actor roles.

Another important research area is to understand the process through which the community of users participating in the event, co-construct a shared understanding of the event through the dynamic formation of groups as indicated by the shifts in actor roles for the top shared messages over time. These shifts, and perhaps specific patterns of shifts, can be viewed as an attempt by the community to construct a consistent and defensible shared understanding of the individuals, the groups, and the behaviors exhibited in the event.

Finally, an understanding of the dynamic formation of groups around specific messages may generalize and help us create sociotechnical systems for spreading scientific or technical information across social media.

**Acknowledgments.** This material is based partly upon work supported by the National Science Foundation (NSF) under CMMI–1635334. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

## References

1. Mohr, L.B.: Explaining Organizational Behavior. Jossey-Bass, New York (1982)
2. Miller, M.E.: Viral standoff between a tribal elder and a high schooler is more complicated than it first seemed. New York Times, 22 January 2019. https://www.washingtonpost.com/local/social-issues/picture-of-the-conflict-on-the-mall-comes-into-clearer-focus/2019/01/20/c078f092-1ceb-11e9-9145-3f74070bbdb9_story.html. Accessed 14 Feb 2019
3. Soave, R.: The media wildly mischaracterized that video of covington catholic students confronting a native american veteran. Reason, 20 January 2019. https://reason.com/blog/2019/01/20/covington-catholic-nathan-phillips-video. Accessed 14 Feb 2019
4. Hollan, J., Hutchins, E., Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans. Comput.-Human Interact. **7**, 174–196 (2000)