Abstract

In extreme events such as the Egyptian 2011 uprising, the online social media technology enables many people from heterogeneous backgrounds to join the event in response to the crisis. This form of collectivity (online crowd) is usually formed spontaneously with minimum constraints in the relationships among the members. The theories of collective behavior suggest that the pattern of behavior in the crowd is not just a set of random acts. Instead they evolve toward a normative stage. Because of the uncertainty of the situations people are more likely to search for norms. Understanding the process of norm formation in online social media is beneficial for any organization that seeks to establish a norm. In this study, we propose a longitudinal data-driven approach to investigate the process of emergent norm formation in online crowds in the context of the online crowd formed around the Egyptian Revolution in 2011.

Keywords

Online Social Media, Online Crowd, Emergent Norm, Collective Behavior, Twitter.

Introduction

In January 2011, the Egyptian revolution started, when 29 year-old Egyptian-born Wael Ghonim started a Facebook page in solidarity with a fellow Egyptian (Khaled Said) who was beaten to death by the Egyptian police force (Vargas, 2012). Soon after, many people from heterogeneous backgrounds joined the campaign. The movement went through several phases in its lifetime and eventually ended on February 11, 2011 after resignation of Hosni Mubarak, the president of Egypt at the time. The role of online social media such as Facebook and Twitter in facilitating the Egyptian 2011 movement is undeniable. The online social media technology enabled people from heterogeneous backgrounds to spontaneously form an online collectivity in support of the movement. This form of collectivity is called “crowd”.

Turner and Killian (1964) define Collective behavior as a spontaneous social process or action of crowds. In Turner and Killian’s perspective, collective behavior takes place in unusual situations with the aim of redefining the situation and making sense of confusion. Despite the spontaneous nature of the crowd and the lack of constraints in the relationships among the crowd members formed around the Egyptian movement, the pattern of collective behavior in the crowd is not believed to be just a set of random acts. Instead the collective behavior evolves toward a normative stage. The uncertainty of the situations is the triggering factor which drives people to search for norms.

Understanding the nature of human behavior in online social media is beneficial for any organization that seeks to provide appropriate technology for its social media platform and gives insights on how to improve and adjust the technology to social needs. The process of norm formation and rule establishment in online crowds is a phenomenon of interest which has not been given extensive attention in the field of Information Systems. Thus, our research question is: How does the norm emerge in online crowds?

Understanding the human behavior in social media and the user influence has applications in viral marketing (Cha et al., 2010) and creating public awareness (e.g. emergency situation awareness, public awareness regarding a disease or a group of people suffering), etc. Especially if an agency seeks to establish a norm, they need to know where to start and which group to target. Agencies need to target
users who are influential enough to establish a particular norm of interest. The outcome from this study will also provide development suggestions for social media and contribute to enhance its productivity. The advent of online social media technologies brings us to the new era of online societies and provides us with an abundance of data about the interrelations of online actors in online social environments. This abundance of data provides us a unique opportunity to study and track the underlying process of norm formation in online societies. The norm in a social environment is a set of practiced behaviors which are known by the participants. When a norm is neither pre-defined nor regulated by the agency, it will emerge from the crowd and thus is called the emergent norm. Examples of norms in online social media could include reposting verified information from a specific set of trustworthy sources, inclusion of a reference to the source of content, sharing specific types of content (such as image, text, or video) more frequently, etc.

In this paper, we propose a longitudinal data-driven approach to investigate the process of emergent norm formation in online crowds formed around the Egyptian Revolution in 2011. The data set, the research methodology, and the preliminary results of our study are presented in the paper.

Theoretical Background

In his theory of Structuration, Giddens (1991) suggests that a social system is not merely a mass of individual random acts; instead, human agency and social structure are in constant relationship with each other which gives such systems a dual structure (the micro and macro structure). He believes that structure is not the shape but the process which gives form and shape to social life (Giddens, 1991; Giddens, 2013). Throughout the use of social media technology, crowds gradually gain expectations of the technology and define and redefine the normative behavior of technology use, and the norm gradually emerges.

Turner and Killian (1964) propose Emergent Norm Theory (ENT) in the context of a collectivity of people with heterogeneous backgrounds and perceptions. In extreme events such as a crisis, the emergency situation forces people to give up established and legitimized actions and take the appropriate action in response to the crisis. ENT proposes that crowds shape normative structures through the processes of milling and keynoting. Crowds start to make sense of the situation by interacting with each other and going through a milling stage. Searching for appropriate behavior in a particular unexpected situation is the main concern in the milling stage. At this stage, particular individuals play the role of keynoters in the crowd. The keynoter is an individual who can help explain the situation and reduce the ambiguity. The keynoting process occurs when the crowd starts to build a shared image of what’s happening and separates facts from rumors (Turner and Killian, 1964).

In collective behavior, either norms or relationships or both could be emergent. Based on the level of the emergence of the norms and relationships, Weller and Quarantelli (1973) categorized the collective behavior into four types. Figure 1 shows the topology of the collective behavior introduced by Weller and Quarantelli (1973).

The previous studies on norm formation mostly focus on the enduring end of the two dimensions in figure 1; where either the social relationships or the systems of norms are enduring. There have been few studies on ENT in groups in physical or digital environments (Friedkin, 2001; Postmes et. al., 2000). Postmes et. al. (2000) analyzed the formation of group norms through email communication among users of a computer mediated course (Postmes et. al., 2000). Consistency of communication was defined as an indicator of the emergence of norm. Postmes and colleagues identified various attributes of the communication and measured them through a longitudinal study to see the pattern of changes in the communication style within each group of users. Based on the results of their study, they concluded that the content of communication is normative and the conformity to the emerged norms increases over time. While providing interesting insights, the subjects in this study were members of a group with relatively enduring interpersonal relations, which are fundamentally different from crowds. The context of our study falls on the emergent ends of the typology in figure 1, which represent the collective behavior in new situations.
Research Method

Data
Twitter is one of the most widespread social media technologies used in various contexts. Its unique feature of a 140 character limitation in each tweet post makes it a good fit as a means to spread information fast. Also, because of this limitation, users employ creative ways to make the best use of this technology. The data from the 2011 Egyptian uprising has been collected using Twitter API. The dataset contains more than 343,000 tweets posted during January 12, 2011 to March 10, 2011. More than 20,000 unique Twitter users have been contributing to the body of tweets posted during the lifetime of the event; these users posted about 202,000 original tweets and retweeted around 141,400 posts. This dataset has been pre-processed and the isolated nodes have been cleaned out for increased efficiency.

Data Analysis
Looking through the lens of the Structuration Theory, we take macro and micro approaches together to study the process of norm formation in the Egyptian movement. These two approaches include: (1) defining and identifying the indicators of the normative behavior, and (2) explaining the underlying processes of norm formation. In the first approach, we need to demonstrate that some forms of norms are evident in the pattern of behaviors extracted from the data. This approach is from a macro viewpoint which allows us to uncover the emergent and sustainable patterns of behavior, which convert to norms over time. We used a data-driven approach to address this question. In this methodology, we analyze the tweets to identify the normative patterns of behavior during the lifetime of the movement. The second approach is from the micro viewpoint from which we indicate the influence path and influential groups of users at each time frame and study the scope and reach of the audience of each group. The micro viewpoint allows us to figure out how the norms get established through a social process. The longitudinal study will help us uncover the process of self-organization in the networks and identify the changing patterns of behavior over time.

Previous studies suggest that both interpretive and structural approaches look at the social systems from different perspectives and thus the combination of both approaches is a better option for studying the social systems. The former emphasizes the content of the interaction among different roles while the latter uses social network metrics to differentiate between individual nodes in a network (Gleave et al., 2009). Role identification makes it easier to study the social system and creates the potential to answer the system level questions about the impact of different social roles on norm enforcement (Gleave et al., 2009). In order to account for the impact of different social roles in the crowd, we categorize the crowd members and classify them into different pre-identified roles. Identifying roles will help us study the crowd behavior both from the macro level (by identifying the pattern of behavior of each role in the network), and the micro level (by differentiating among the influential power of each role on their neighbors in the crowd). The next step in our study will be the content analysis of the tweets, then the content characteristic measures will be aggregated for each role from these linguistic perspectives.

Figure 1. Institutional and collective behavior collectivities (Weller and Quarantelli, 1973)
on these characteristics, the tweets will be categorized into different types. We will be able to find the correlations between the roles and the types of tweets each role posts more frequently.

We take the initial step of our analysis to identify the overall patterns of behavior regardless of the user roles as described in the first approach. The results of this part of the analysis are provided in the next section.

**Preliminary Results**

First of all, we analyzed the frequency of original tweets over time. The graph includes peaks in the frequency at different points of time (Figure 2a). These peak times include January 25, January 28, February 2, and February 11, 2011. Looking back at the pre-revolution events on the timeline of the 2011 Egyptian uprising (Table 1), we found the correlation between the critical points of time and the peak times in the tweet frequency graph. This simply tells us that depending on the nature of the movement, the triggering events drive people to disseminate higher number of original tweets to respond to the increasing need of firsthand information. In the first step of our analysis, which is identifying the overall pattern of behavior, regardless of user roles, we extracted the following measures and tracked them over time. We considered two of the main technology features which are commonly adopted by the users of Twitter. The measures include the number of hashtags per tweet and inclusion of a url to the source of the content. The usage of the aforementioned tweet notions follow different patterns in regards to the total number of original tweets over time. Regardless of which role posted the tweets, the results show that, in general, when the number of tweets goes up, the average number of urls per tweet drops down (Figure 2b). This tells us that during an event, when an occurrence of an incident triggers lots of discussions in online environments (especially in critical points of time), inclusion of a source of content is of less importance. That’s when the rumors find their way to spread in the network. However, the average number of hashtags per tweet experiences a relatively stable pattern over time, regardless of number of tweets (Figure 2c). The purpose of using hashtags in tweets posts is to group the messages by labeling them. Labeling Tweets makes it easier for future references by making them appear in the electronic searches. The stability and consistency of the number of average hashtags used per Tweet in Egyptian uprising could be because of the existence of potential norm regarding the use of hashtags which puts a limit on the maximum number of hashtags in a single tweet. Especially because of the character limitation in each single tweet (140 characters limitation), the hashtags should be used with caution to make the best use of this limitation. In our future study, we will analyze the types of hashtags being used over time during the Egyptian movement, to look for potential correlations between the types of hashtags and the triggering events.

<table>
<thead>
<tr>
<th>Critical Point</th>
<th>Incident Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 January 2011 (&quot;the Day of Rage&quot;)</td>
<td>Ten of thousands protested throughout Egypt</td>
</tr>
<tr>
<td>28 January 2011 (The &quot;Friday of Anger&quot;)</td>
<td>Hundreds of thousands demonstrated after Friday prayers. Clashes broke out in Tahrir Square</td>
</tr>
<tr>
<td>2 February 2011 (Camel Battle)</td>
<td>Mubarak supporters rode camels in Tahrir square as a sign of escalation of violence</td>
</tr>
<tr>
<td>11 February 2011 (&quot;Friday of Completion&quot;)</td>
<td>Mubarak resignation and nationwide celebration</td>
</tr>
</tbody>
</table>

**Table 1. Description of the critical points in the life time of Egypt 2011 revolution (Causey, 2012)**
Figure 2. Pattern of original tweet post (a) pattern of url inclusion (b), and hashtag usage (c) during the lifetime of the Egypt 2011 movement compared to the total number of original tweets.

Conclusion

This paper introduces a longitudinal approach for studying the norms and the process of norm formation in an online social environment. Considering the context of online crowds, this study makes the effort to extend the application of structuration theory beyond the scope of institutionalized groups with enduring relationships and norms systems. The results of the first step of our methodology suggest that the online crowd follows patterns of behavior consistent with the external triggering events. The further studies are required to extend the indicators of normative behavior considered in this study and also account for the potential effects of confounding factors such as the demographic information of the users and the peer social influence. Our future study plans also include the analysis of other content characteristics, such as linguistic characteristics of the tweets which will be identified using the LIWC linguistic tool. Then the aggregated values of each of the tweet characteristic measures should be calculated for each role at each time period. The tweets will be categorized based on each identified role. All the tweets are time-stamped so we are able to track the pattern of changes in tweet content for each role over time.

There are some challenges and limitations in our study. These challenges will be resolved during the study and experiment. But at this point we just acknowledge them and leave the potential resolutions for future. First, external influential factors such as those in physical environments will have effects on network influence. These factors are difficult to control. Media is difficult to be detached from the online network, especially when the Social Media accounts of the news media are core for online information flow. Second, only reposting content from someone else may not fully reflect the influence. And analysis of click-through data may help resolve this limitation.

The results of our initial step confirm that the online crowd formed around the context of the Egyptian uprising in 2011 is not acting randomly. Instead, the collective behavior follows a pattern towards a normative stage. The next step of our study will focus on identifying how the norm is reached and the process underlying the norm formation in online crowds.
REFERENCES


