

## CHAPTER XX

# Social Networks, Social Media, Social Change

*Jürgen Pfeffer, Kathleen M. Carley*

Carnegie Mellon University  
Pittsburgh, USA

[jpfeffer@cs.cmu.edu](mailto:jpfeffer@cs.cmu.edu), [kathleen.carley@cs.cmu.edu](mailto:kathleen.carley@cs.cmu.edu)

### ABSTRACT

Social Media has been touted as the way in which social networks form and can be mobilized to effect social change. In particular, the Arab Spring 2011 has been heralded as a social media based social transformation. On the other hand, scientists and journalists concede social media a less important role in the spate of revolutions and uprisings. In this article we make a contribution to this discussion by analyzing structural change in Arabic written Twitter data and by comparing the results with data collected from traditional English written news sources. We orient this discussion by pointing out, using a simulation example, the socio-demographics internet penetration and literacy are critical to the impact of any media. Then, to illustrate our cross-cultural analysis approach we use news articles and tweets discussing Syria in December 2011. We are able to show that online and offline data is significantly correlated but the change can be measured earlier in online networks. If we assume that social media can react immediately to forerunners of social change and if we have the ability to identify structural change in social media networks in real-time, we get a step closer to predicting social change.

**Keywords:** dynamic network analysis, social media, text analysis, change detection, rapid assessment.

## 1 NETWORKS TO DESCRIBE SOCIO-CULTURAL SYSTEMS

On January 26, 2011, Hasan Ali Akleh from Syria set himself on fire to protest against the Syrian government. More than two months earlier the similar death of Mohamed Bouazizi from Tunisia marked the beginning of the uprisings in countries in Northern Africa and the Middle East which are known as the Arab Spring. The

time and occasion for the beginning of the uprisings in Syria are similar to other countries; however, the progress of the escalation was different and took longer due to other factors such as disagreement of the veto power wielding nations in the United Nations Security Council and some specific local conditions. In this article we are not discussing these conditions or the global or local political decisions which lead to the situation in Syria at the end of 2011. Instead, we examine the structural aspects of the conflict using an approach for rapidly analyzing the raw text data that can be applied to both social media data and data from *classical* media sources such as newspapers.

Users of social media and regular media interact with each other and are engaged in creating and discussing content. These activities form various networks (Wasserman and Faust, 1995). Network text analysis (Carley, 1997) can be used to extract these networks from texts. The resulting networks are dynamic meta-network representations of complex socio-cultural systems and can be analyzed by making use of Dynamic Network Analysis (Carley, 2002). Since the situation rapidly changes when countries are in conflict rapid assessment is critical (Harris et al., 1997; Carley et al., 2011). Pfeffer and Carley (forthcoming) showed how to rapidly model and analyze networks extracted from pre-structured news articles. However, the methods developed for these rapid assessment approaches can be used for both social media data and traditional news article data.

In this article we compare the analysis of the Arabic tweets with English written news articles of the same time period for Syria. We are interested in the way social media content reacts to social change and how we can measure and visualize this change. In section 2 we set the context, geo-spatially and technologically. In section 3 we describe the data that we used for this work. Section 4 describes the steps of analysis and the results are discussed in section 5.

## **2 POTENTIAL IMPACT OF SOCIAL MEDIA ON THE ARAB SPRING**

Digital communication technologies enable people to connect globally and create communities independently from their actual geographical location (Harasim, 1993). This results in less need for an overlap of the geographic and the social spheres (Meyrowitz, 1985), supports coordination and mobilization, and changes the way people communicate and engage in political dialogue (Seib, 2008). Consequently, new communication technologies and social media are touted as critical to the recent uprisings and revolutions in Northern African and Middle East countries (Seib, 2008; Campbell, 2011). Apart from the question “are the revolutions internet revolts or not?” (Zhou et al., 2011), the belief in the power of the internet was such that in some cases embattled leaders turned off access, e.g., Egypt and Syria.

Media usage requires both access to the media and for some media a certain level of literacy. The impact of messages sent out via a particular media will depend on not just the media, but these access constraints. In general, socio-demographic

factors will play an underlying role in determining access. Hence, the impact of media on what people learn, believe, and their behavior will vary widely across cities or countries with different socio-demographic and access profiles. This is illustrated in Figure 1 using the Construct simulation where messages were sent to the public to deter negative behavior via different media. The point is not the level of negative behavior but that the same message sent via the web and via the newspaper can have, depending on the underlying demographics and access constraints exactly the opposite impact – as is the case in Maine.

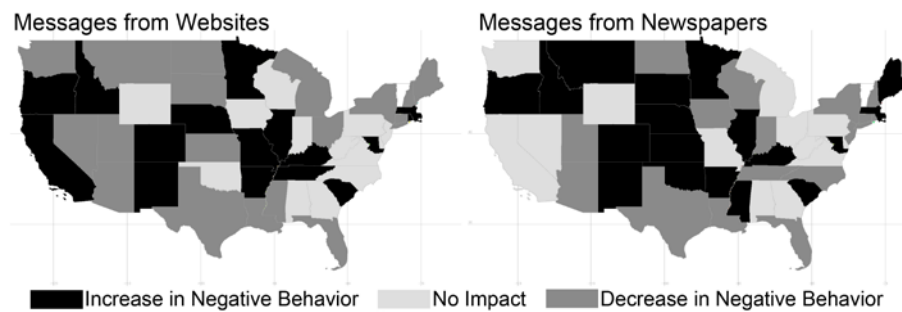


Figure 1 Construct simulation of impact of messages sent via the web and via standard newspapers on change in negative behavior by state based on networks data inferred from socio-demographics obtained from census data.

If we consider Syria, we find that social media usage is probably in the province of only the educated and economic elite. The profile of Syria in terms of media access, based on data from the world fact book is as follows. Overall, 80% of the population is literate suggesting that the vast majority can learn information through the traditional media. In contrast, only 43% of the population have cellphones and 20% have access to the internet. Although we do not have empirical data subject matter experts on the region suggest that most of the illiterate are not in the major cities and that most of those with cellphone and internet access are in or near the major cities. This suggests that it is the intellectual elite that is more likely attending to social media; whereas, the overall population attends to traditional media.

In a related study we examined the news paper data for the 18 countries involved in the Arab Spring over ten months. This study provided evidence that traditional media just draws the connection to social media in case of a successful revolution and not in the case of failed revolutions (Pfeffer and Carley, forthcoming). In addition, it was found that concern with the internet was not related to revolution – see Figure 2. Further, a simple regression exploring the extent to which major topics predicted revolutionary activity controlling for literacy, internet penetration, cellphone usage revealed that the only factors that were significantly related to revolution at greater than a .05 level were concern with human rights and international relations – see Table 1.

**Table 1. Multiple Regression for Revolutionary Activity for Arab Spring**

Variable	Coefficient	Significance
Intercept	0.0025	0.71
Economic News	0.0042	0.98
Human Rights	0.7246	0.00
Facebook	0.3622	0.44
War & Conflict	0.0375	0.42
Terrorism	-0.0295	0.32
International Relations	-0.0572	0.04
$R^2 = .759$ $F=5.79$ Significance = 0.01		



Figure 2 Prevalence of focus on the internet and facebook, such that the darker the country the more concern there was with social media and the internet.

### 3 TWEETS & NEWS ARTICLES

To analyze the questions raised in the introduction, we apply quantitative approaches which make use of publically available information from two very different sources. We collected tweets from Twitter as well as news articles for the same time period. The time period of interest is December 1 to December 18, 2011. The language used in the two data sets is different, the Twitter is Arabic/English and news is English. We now describe the gathering and handling of the data sets.

*Tweets* were extracted using tweet tracker. We selected tweets having “Syria” as a hashtag in the text written in English or Arabic. In addition, Arabic written keywords related to humanitarian crisis, e.g. violence, death, food, shelter, etc. were used as filter to reduce the number of tweets. During the time period of December 1 and 18, 61,633 Arabic written tweets were gathered from 10,186 users with this approach. These tweets include 2,588 different hashtags. For the following analysis we focus on the pre-structured network created by users posting specific hashtags in their tweets.

For the set of *news articles* used for the analysis of this article, we used the LexisNexis Academic archive (<http://www.lexisnexis.com/hottopics/lnacademic/>). The selection “Major World Publications” includes 600+ English written news sources (newspapers, press agencies, and magazines). For the selected time period in December 2011, 7,763 articles were collected and downloaded on February 16, 2012. For the analysis in section 3 we focus on the subject networks which were extracted using the rapid network text analysis approach by Pfeffer and Carley (forthcoming).

Figure 3 shows an overview of the number of Tweets and the number of articles

by day. The x-axis represent the day in December 2011. The y-axis is the number of articles by day. Because the magnitude of the data is different, we use two different y-axes; the left one is for the articles and the right is one for the tweets. The minima for the news articles on days 3 and 4, 10 and 11, as well as 18, mark the weekends (Saturday and Sunday) with less media newspapers and, therefore, less media coverage.

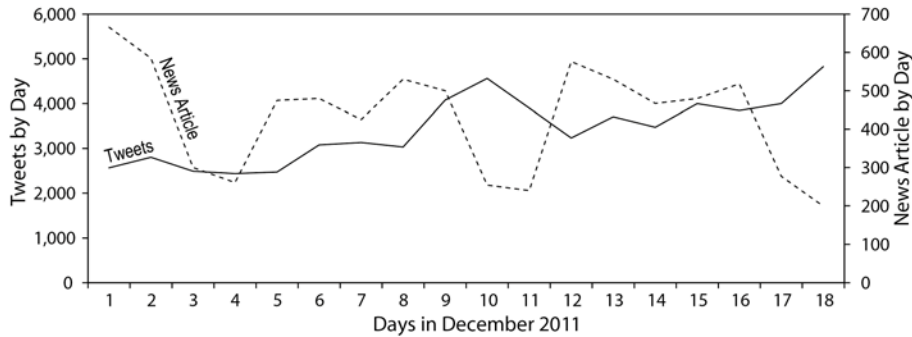


Figure 3 Number of tweets by day (left y-axis, solid line) and number of news articles by day (right y-axis, dashed line) in December 2011.

In Figure 3, despite daily oscillations, the news data trend is essentially flat; whereas, the trend in the number of tweets is a rise, but with an early peak from December 9 to 11. In this sense, the traditional media shows an overall *stability* despite the fact that the underlying situation in Syria during this time period is *unstable*.

### 3 ANALYZING CHANGE

We now consider what changes occurred in the content and the structure of the collected data. When we look more closely at the index items of the news article and at the hashtags of the tweets, we find that the city of *Homs* is among the most frequent terms. Therefore, the first analysis deals with the importance of Homs in our data over time. We use this case to look at the correlation between tweets and news articles. The second analysis takes an overall look at the change of the structure of the tweets by comparing the users, the hashtags, and the network structure created between users and hashtags for every day of the observation period. Finally, we take a closer look at whether the structural change we have observed in the tweets can be retraced by analyzing the change of topics of traditional media coverage.

#### 3.1 HOMS: Change in a Single Topic

With a population of 800,000+ the city of Homs is the third largest city in Syria

Homs is located in western Syria about 100 miles north of Damascus, the capital of Syria. Homs was told to be the center of the Syrian uprisings 2011/2012 and was under siege from the Syrian Army since the beginning of May 2011. To analyze the importance of Homs for social media and for news coverage we compute the proportion of documents in both sources containing Homs (in English or Arabic).

The result visualized in Figure 4 shows a strong connection between the two lines over time. During the first 7 days there is little difference in the attention to Homs in the two media. On day 8 the proportion of tweets with the keyword Homs triples. The same happens for news articles within the following two days. The decrease from these maxima takes about one week. The tremendous change in the prevalence of Homs in both media is due to the escalation in violence and social unrest in the city as dozens of deaths occurred within 24 hours all related to a planned nationwide strike. The Syrian Army issued the people in Homs an ultimatum to end the protest within 72 hours (The Daily Telegraph, 2011). These and other related incidents resulted in increased media coverage on December 9 and even more on December 10. The more interesting result in Figure 3 is that the maximum in the tweet data is reached on December 8, one day before the documented escalation.

A statistical analysis of the correlation between the two proportional lines of Figure 4 reveals a significant correlation if we allow a time-lag for the news articles. The Pearson correlation is 0.43 with a weak p-value of 0.07 when we compare the two time lines as they are. The small inset picture in Figure 3 represents the correlation between the proportion of the news articles and the proportion of the tweets including the city of Homs with different time-lags of the news articles. The largest correlation with 0.64 (p-value: 0.006) can be measured with a time-lag of +1 day. Therefore, the number of news articles referring to Homs is most highly correlated to the tweets referring to Homs at a delay of one day. As newspapers normally print yesterday's news, this result is not surprising. However, the extent of the correlation is astonishing regarding the two very different data sources.

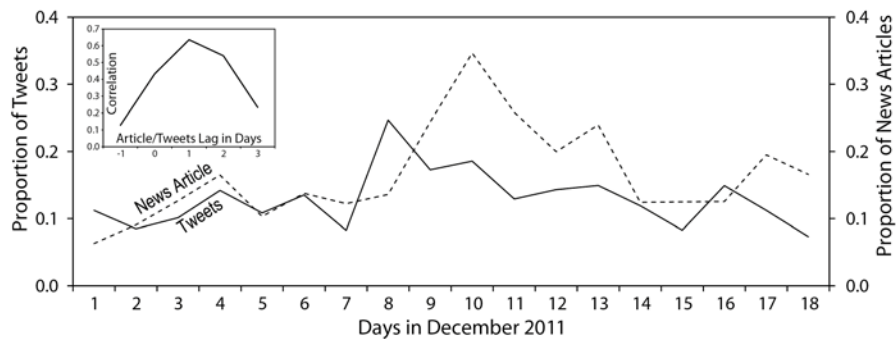


Figure 4 Proportion of news articles (dashed line) and tweets (solid line) including "Homs" in English or Arabic language. Small picture: Correlation between the proportion of news articles and the proportion of tweets with a lag in days for the news articles.

### 3.2 Change in Online Structure

In the previous sub-section we were able to show a numerical change in tweets as well as a change of importance of the city of Homs. Now we consider whether there is also an overarching structural change in the tweet network in addition to the change in frequency. The tweet network consists of 10,186 users, 2,588 hashtags, and the network created by users posting the hashtags. To compare the 18 networks representing the 18 days of the observation period, we create three mathematical objects for every time point: A vector for every day including every user with the number of tweets s/he sent at this day, a vector for every hashtag with the number of tweets the hashtags are part of, and a user/hashtag network (matrix) containing how often every hashtag is used by every user. To get the similarities and dissimilarities between the 18 time points we calculate the distances of the three objects of every pair of time points resulting in an 18 x 18 distance matrix.

Applying multidimensional scaling (MDS) on this matrix and using the first eigenvector for the axes of a scatter plot results in Figure 5. Every point represents a single day of tweet data. The space itself is the two-dimensional representation of high-dimensional data. The position of a single data point has no meaning but the overall arrangement of the points does have meaning. Two data point are drawn near each other in case they have similar values in the lines and columns of the distance matrix. Consequently, two days are near each other in Figure 5 if similar twitter users use similar hashtags and two days are distant if different users post different hashtags.

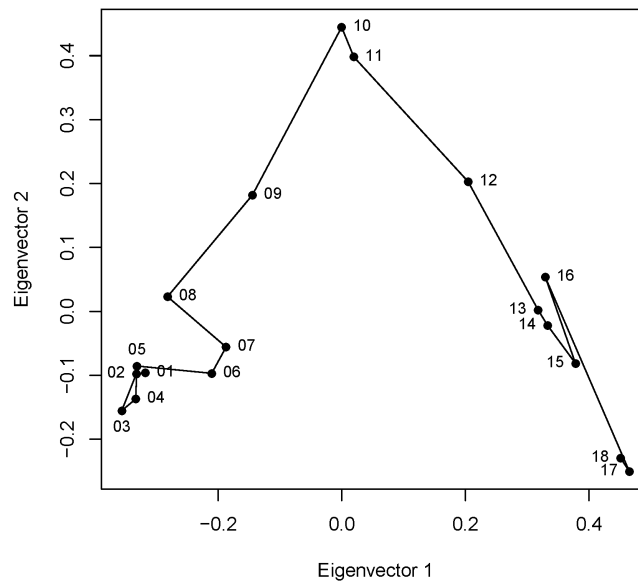


Figure 5 Structural change of the tweet network over time. Multidimensional scaling of similarities and differences of users, hashtags, and the user/hashtag network over time.

The first 5 days are rather stable from this perspective because the points are located near each other. Starting at day 6 the distances between two consecutive points increase as well as the distance from the starting area of the structural space. Day 9 and 10 as well as day 12 and 13 bring the user/hashtag network in a complete new area where the network stabilizes for some days. The last 2 days of the observation period are, again, in a different area implying an altered set of users writing different stories. One interpretation for the structural change in the tweets can be offered by looking at the weekdays connected to the change. Days 2, 9, and 16 are Fridays and Friday is the organized protest day in Syria. Even more, every Friday is assigned with a specific label which should help the protesters to focus their protest, e.g. December 9 - "Friday of the Dignity Strike", December 16 - "Friday of the Arab League Is Killing Us". Our analysis shows evidence that these events have the power to change the social media networks by changing the users and topics they discuss. In future research we will analyze longer time periods to see if the results stabilize.

### 3.3 Change in Media Coverage

We now contrast the change in the structure of the tweet network with the activity in the (English written) news articles. We compute the most important index items of media coverage for the *stable zones* of Figure 5, e.g. the consecutive days 10 and 11, and 13 and 14 can be considered as stable because of structural similar tweet networks. Table 2 shows the index items with the highest occurrences for days in different areas of Figure 4. We include the one day delay suggested by the analysis in section 3.1. Therefore, days 11 and 12 of media coverage should correspond to days 10 and 11 of the social media data.

The first column is dominated by international politics and the discussions about embargoes and sanctions against Syria. The following two time periods are dominated by Bashar Al-Assad, the president of Syria as well as by mentions of the protest and escalating conflicts. At the end of our observation period media coverage seems to focus on both protests and international politics. Therefore, this high level analysis provides evidence that the topic structure of news articles is changing even as the structure of the social media is changing

**Table 2 Top 5 index items in news articles of selected days**

#	December 1+2	December 11+12	December 14+15	December 18
1.	International Relations	Bashar Al-Assad	Bashar Al-Assad	Protests & Demonstrations
2.	State Departments	Protests & Demonstrations	Protests & Demonstrations	Talks & Meetings
3.	United Nations Institutions	International Relations	International Relations	International Relations
4.	Embargoes & Sanctions	Armed Forces	Armed Forces	State Departments
5.	Bashar Al-Assad	War & Conflict	United Nations Institutions	Heads Of State & Government



## 4 DISCUSSION

Social media as a technology is evolving rapidly. As such, predicting its role in effecting social change and social networks is to an extent speculation. Current trends, though, suggest that the technology is being altered in three key ways: specialized support for topic/task based social networks, real-time news, and integration with geo-location tools. Our findings, keeping in mind these trends, suggest that in a social media world it will be possible to support the mobilization of larger groups, to enable faster coordination, and to provide information and focus attention on more specific causes.

In this article we used data from the uprisings in Syria in December 2011 to explore the role of social media in social change. We created networks from Twitter data and from news archives. We were able to show high correlation for specific topics between social media and traditional news articles even if the data sources are written in two different languages. For the case of Homs we were able to show that online and offline media is highly correlated with a one day delay. In particular we showed evidence that the Friday demonstrations change the structure of the online networks and the topic networks of media coverage. Future work will discuss the interrelations between social media and media coverage in more detail.

We find it critical to draw a distinction between cause and coordination. Our analysis of news media relative to the 18 countries involved in the Arab Spring over the course of 10 months revealed that the pattern of spread of the revolutions was not related to the pattern of social media usage, nor was the concern with or use of social media per se related to the advent of protests. In other words, the social media did not cause the revolutions. Rather, what these two studies Syria of and the Arab Spring suggest is that social media is a coordination mechanism that enables and mitigates activity rather than a force causing activity.

We suggest that social networks are not static structures; rather, they are continuously evolving socio-cultural systems that evolve as individuals interact and communicate. The detection and assessment of change in these dynamic meta-networks is in the focus of our future work. Our analyses did not locate actual events in the context of the tweets but they did provide strong evidence that the change in online media happens in real-time or maybe *earlier* than some actual event activities. The Homs example suggests that change in online networks may presage the occurrence of socially coordinated real-world incidents. This could be possible if social media users are reacting to critical causal events and then using social media to coordinate the social response. If this is the case, then future research should attempt to characterize events in terms of the coordination level needed, and the complexity of the causal events, as well as creating new metrics for better describing and assessing change in social media networks. Further, given the access constraint issues, future research should consider how the nature of social media as a coordination tool changes its effectiveness as the socio-economic profile of those with access to the social media changes.

## ACKNOWLEDGMENTS

This work is supported in part by the Office of Naval Research (ONR), United States Navy (ONR MURI N000140811186, ONR MMV N00014060104). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research or the U.S. government.

## REFERENCES

- Campbell, D. G. 2011. *Egypt Unshackled: Using social media to @#:) the System*. Amherst, NY: Cambria Books.
- Carley, K. M. 1997. Network Text Analysis: The Network Position of Concepts. Text Analysis For the Social Sciences, in: *Methods for Drawing Statistical Inferences from Texts and Transcripts*, ed. C. W. Roberts, Lawrence Erlbaum Associates, Mahwah, NJ, USA.
- Carley, K. M. 2002. Smart Agents and Organizations of the Future, in: *The Handbook of New Media*, eds. L. A. Lievrouw, S. Livingstone, Sage Publ Inc, Thousand Oaks, CA, USA.
- Carley, K.M., Bigrigg, M., Papageorgiou, C., Johnson, J., Kunkel, F., Lanham, M., Martin, M., Morgan, G., Schmerl, B., and van Holt, T. 2011. Rapid Ethnographic Assessment: Data-To-Model. *Proceedings of HSCB Focus 2011: Integrating Social Science Theory and Analytic Methods for Operational Use*.
- Harasim, L. M. 1993. *Global networks: computers and international communication*, Massachusetts Institute of Technology Press, Cambridge, MA.
- Harris, K. J., Jerome, N. W., and Fawcett S.B. 1997. Rapid Assessment Procedures: A Review and Critique. *Human Organization*, 56 (3): 375-378.
- Meyrowitz, J. 1985. *No Sense of Place: The Impact of Electronic Media on Social Behaviour*. Oxford University Press.
- Pfeffer, J. and Carley, K. M. forthcoming. Rapid Modeling and Analyzing Networks Extracted from Pre-Structured News Articles. *Working Paper*, CASOS/ISR, Carnegie Mellon University, Pittsburgh, PA.
- Pfeffer, J., and Carley, K. M. forthcoming Rapid Network Text Analysis of the 2010/11 Arabic Conflicts and Revolutions. *Working Paper*, Carnegie Mellon University, Pittsburgh, USA, .
- Seib, P. 2008. *The Al Jazeera Effect: How the New Global Media Are Reshaping World Politics*. Dulles, VA: Potomac Books Inc.
- The Daily Telegraph, Sherlock, R. 12/9/2011. Syria Crisis: Assad Regime Forces Issue Deadline for End to Protests in Homs. London. Retrieved: 2/29/2012.
- Wasserman, S., and Faust K. 1995. *Social Network Analysis, Methods and Applications*, Cambridge University Press, Cambridge, UK.
- Zhuo, X., Wellman, B., and Yu, J. 2011. Egypt: The First Internet Revolt? *Peace Magazine*, 27/3, pp. 6-10.