Using news media and Twitter to provide a quantitative overview of factors related to the outbreak, successes and failures of the "Arab Spring"

Kenneth Joseph [*1], Matthew Benigni [†1], Wei Wei [‡1], Kathleen M. Carley[§1], and Mia Bloom[¶2]

[1]Computation, Organization and Society Program, Carnegie Mellon University, 5000 Forbes Ave. , Pittsburgh, PA, USA
[2]Center for Terrorism and Security Studies, University of Massachusetts, Lowell, Lowell, MA, USA

April 10, 2015

Abstract

1 Introduction

Story concepts:

how do discussions change over time and space in the arab spring?

this allows us to compare public focus and news focus

which drives which and how much

in intro: problems with count data, e.g. by showing diffrences in A, other biases

Work to do:

1. Fake data create 1 that is positive, does it recover random walk data

[*]kjoseph@cs.cmu.edu
[†]mbenigni@cs.cmu.edu
[‡]weiwei@cs.cmu.edu
[§]kathleen.carley@cs.cmu.edu
[¶]mia_bloom@uml.edu

2. Stop word data

3. Re-estimate model with this

4. Is A term significant? (Raw MLEs) Hyp A term is diff from 1; compare with boring data. Expected: A is negative over time; generally decreasing interest

5. news is fairly stable, Twitter is not. Also some countries with more noise less stable (bigger sigma?)

6. where do news and twitter diverge? lag plot and just straight time series correlations; maybe think about indices for freedom of press

7. Overall, some high level comparison - event data; corr matrix w/ static data; "event/outlier detection" (or, better, figure out how to get away from this)

8. Case study, Matt

9. lit review

Ideas not for this paper: correlate with network changes over time, compare/correlate categories,

Goldstone's (2011) recent assessment of the drivers of instability and revolutionary success during what has come to be known, for better or worse Gelvin (2015), as the Arab Spring, suggest that a recurrent causal story played a role in the early revolutionary successes in Egypt, Tunisia and Libya. Goldstone states that "virtually all successful revolutions were forged by cross-class coalitions...pitting society as whole against the regime". He then provides qualitative evidence that the revolutions in three countries were no different. In Egypt, Tunisia and Libya, previously disparate social groups combined under a united "cross-class coalition" of protesters and revolutionaries, which made the revolutions significantly more likely to succeed.

In later work, Goldstone (2013) suggests that several additional predictors were important in determining whether or not a particular government was ultimately overthrown. Chief amongst these predictors, he argued, was the structure of the overarching regime had a strong impact on

whether or not the government was ultimately overthrown. Goldstone defines a *personalist* regime as one in which a single individual – who may have begun as an elected leader, or head of a military or even party regime – takes total or nearly total control of the national government. He then provides qualitative evidence for his believe that "the single best key to where regimes in MENA have been overturned or faced massive rebellions is where personalist regimes have arisen".

Goldstone's critiques are two of many articles to consider what may have led to the revolutions that spread throughout the MENA region and whose effects still continue to reverberate globally today. One factor of primary interest was the use of social media, which was originally suggested by popular media to be the sole cause of the spread of the revolution. At this point, it is almost common knowledge that this emphasis on social media as a *cause* of the revolutions is overblown (Bruns, Highfield, & Burgess, 2013; Comunello & Anzera, 2012; Goldstone, 2013). However, many scholars have pointed to the fact that the use of social media may have aided certain aspects of the revolutions in important ways for different people (Gallé, Renders, & Karstens, 2013; Lotan et al., 2011; Starbird & Palen, 2012; Tufekci & Wilson, 2012). This fits with the current understanding of the causal structures involved, as scholars now suggest that a breadth of heavily intertwined issues came together in a unique and unpredictable fashion to provide the conditions necessary for the spread of both successful and unsuccessful revolt and protest throughout the region.

In contrast to work considering the causal role that social media played in the revolution, recent work has suggested that data from social media data may help to inform this broader understanding of how the revolutions grew and spread (). This use of social media data takes as a given that these tools were used in some way during the revolution and goes beyond this in assuming, via the well-worn notion of media multiplexity (), that causal processes of the revolution which existed "offline" are likely to have played out "online" as well. Such work has shown that, for example, X, Y, Z. Similarly, recent work has suggested that data from news media during the time of the Arab Spring

also may be of use in better understanding these processes (Joseph, Carley, Filonuk, Morgan, & Pfeffer, 2014; Pfeffer & Carley, 2012). Combined, efforts with social and news media data from the Arab Spring have shown the value in using this data to a) provide evidence for or against prior hypotheses presented in the literature using quantitative metrics on large sets of data, and b) use the available data to generate new or more nuanced understandings of the incredibly complex web of factors that have been implicated in a causal role.

The present work provides an example of how social and news media data can be used in each of these two ways. First, we extract quantitative evidence for the assertions of Goldstone's (2011). More specifically, we consider the following two claims made by Goldstone and consider the extent to which they played out in social and news media across Egypt, Libya and Tunisia as well as fourteen other nations in the MENA region:

- **H1**: *If a country has more disparate groups, it is more unstable*
- **H2**: *If these groups align under a revolutionary cause, the revolution is more likely to succeed*

These two hypotheses are directly quantifiable via a combination of social media and news data and external indicators based on expert opinion. **todo Kenny: [[In order to evaluate H1, we consider ... In order to evaluate H2, we ...]]**

In the second part of the paper, we provide a characterization of how social media use may have differed in nations with personalist regimes as opposed to nations with other types of regimes discussed by Goldstone. We combine a host of metrics measured on both news and social media and use a variety of statistical techniques to extract differences. **todo Kenny: [[We find XYZ. This leads to an interesting extension/correlary/etc of Goldstone in that XYZ...]]**

2 Related work

The immense number of complexities and historical artifacts that played a role in the beginnings, successes and failures of the Arab Spring prohibit a full exploration here. Instead, we provide only a summary of some of the more widely accepted factors, and ones that readers should be aware of and keep in mind throughout our analysis. For a more detailed overview, we direct the reader to Gelvin's (2015) recent book. In this section, we also discuss recent work utilizing social and/or news media during the Arab Spring, detailing how the methods and data used here are both similar and different.

2.1 Causal factors of the Arab Spring

Two broad classes of causal factors can be considered - those that prompted revolution, and those which affected the success or failure of revolutions. We briefly review each set of factors here.

2.1.1 Causes of protests and their spread

A host of historical factors led to the conditions in 2011 that made many nations in the Arab world ripe for protest Gelvin (2015). One long-standing issue was the increasing extent of economic problems caused by ineffective, corrupt and state-run economies. These issues led to high levels of unemployment and inflation Dewey, Kaden, Marks, Matsushima, and Zhu (2012) as well as to both food shortages and huge hikes in food prices, all of which contributed to high levels of civil unrest Comunello and Anzera (2012); Goldstone (2011). The effect of high unemployment rates was particularly a problem because its effect was particularly strong on well-educated youth populations, individuals who had often been promised that their education efforts would be rewarded with jobs Dewey et al. (2012); Gelvin (2015). This, combined with a "youth buldge" in which a disproportionate percentage of the population was between the ages of 15-29 in many of the MENA

region countries, provided a fodder of civil unrest that required only a spark to ignite and a gust of wind to spread throughout the region.

On December 17th, 2010, a spark came in the form of Mohamed Bouazizi, who immolated himself in Sidi Bouzid, Libya in response to harassment from both a local policewoman and local municipality officers. Bouazizi's case resounded with others who were tired of being harassed and vicitimized by a corrupt government. These individuals took to the streets in protest. Although early protests were relatively small and were met with violence from government forces, social media sites like Twitter, Facebook and YouTube were used to record these events and display them to the broader public.

This access to social media, or more aptly, the existence of communication infrastructures that supported all forms of new media(Tufekci & Wilson, 2012; Wolfsfeld, Segev, & Sheafer, 2013), served as a causal element in the spread of the protests from Tunisia, to Egypt, to Libya and on to several other nations in the Arab world. Even before the protests, the existence of internet connectivity allowed individuals within the Arab region to observe the democratic processes existant in other regions of the world, stoking their desire to live in that type of environment Hussain and Howard (2013). During the early moments of the revolutions, these tools allowed both delibrate diffusion processes, defined as those "carried out via the conscious sharing of tactics and frames by activists who are linked by networks that may be transnational" and demonstration diffusion effects, " ' the power of precedent' " to occur, rapidly engulfing an increasing number of actors in an increasing number of nations in protest. Through both logics, a further cause of spread was the strengthening of an underlying Arabic identity, which united protesters in their unified goal of empowerment and the ending of corruption.

2.1.2 Causes of Varying Outcomes

A consistent take on the primary factor influencing the success or failure of a particular revolution was the structure of the ruling regime Bellin (2012); Comunello and Anzera (2012); Goldstone (2013) . Goldstone's (2013) work argues that personalist regimes were the most susceptible because their power was tied to their ability to provide the necessary economic and political incentives to their constituency, particularly in nations that depended on oil production. While Comunello and Anzera (2012) notes that personalist, or as he refers to them, neopatrimonial states, had controlling arms that made it difficult to organize any sort of formal protest, three other factors led to conditions in which such formal protests could arise.

One such factor that has already been discussed was the economic conditions under which the revolution occurred. While these economics played a role in bringing about revolution, they also prevented personalist regimes from being about to "buy their way out" of the protests, and thus also had a role in the revolutions' successes. The second factor that played into the success of personalist regimes in the face of revolution was the relationship of the regime to the military Battera (2014); Comunello and Anzera (2012). This relationship can be boiled down to one decision the military made - whether or not to shoot at the protestors Bellin (2012)[1]. In countries where the military made the decision to quell protests with violent force, protesters were able to flood the streets without fear of the full wrath of the state. This lack of impunity led to stronger protests that eventually led to the downfall of the Tunisian and Egyptian regimes. In contrast, in **todo anyone: [[A,B and C]]**, where the military opted to side with the regime and use force to quell protests, regimes managed to maintain power.

[1] While regimes certainly had other forces willing to deal with the protestors, the extent and breadth of the protests eventually came to a point at which it "is sufficient to look at the character of the military and its capacity and will to repress in order to reckon the immediate chances of regime survival." Bellin (2012)(pg. xx).

In making the decision of whether or not to fire on protestors, one important factor was the extent to which previously disparate social groups formed a cross-class coalition in their protests and revolutionary efforts Goldstone (2011) . A unified coalition of protestors made it more difficult for the military to justify the use of force in their response to the protest for two reasons. First, the combination of various social groups lessened the military's ability to claim that violent actions were a response to a particular out-group in the interest of protecting the "nation". Second, the sheer size of such a coalition would relegate military action to being viewed as "illigitimate slaughter" Bellin (2012).

The development of these unified coalitions across various social groups also provided an opportunity for the news media to characterize protestors under a national identity, rather than as protests via one specific subgroup. As we discuss shortly, the portrayals of the revolutions by news media thus had an important impact on revolutionary outcomes in that their coverage ultimately influenced a final factor in revolutionary success, which was the extent and type of involvment in the revolutions in the various nations from international powers Comunello and Anzera (2012); Goldstone (2013). This applies both to the actions of the West during the revolutions as well as the actions taken by Arab nations themselves.

Social media, of course, had much the same effect, increasing both the level of information spread and the extent to which individuals felt compelled to participate in protests Bellin (2012); Tufekci and Wilson (2012); Wolfsfeld et al. (2013). As insinuated above, it is generally agreed upon that social media, and new media more generally, played at least some role in the spread and success of certain revolutions that occurred during the Arab Spring. There do exist, however, many scholars who feel this effect has been overstated, or that no such effect exists at all. Comunello and Anzera (2012) give an overview of a significant amount of work focusing on the relationship between social media and the Arab Spring. The authors consider how, across 8 different properties of socail

media, technological determinists (people who believed social media played *the* causal factor) and the techno-realists (people who believed social media played *no* causal role) differ on their opinions regarding the effect of social media. Additionally, Hussain and Howard (2013) provide a unique analysis of the role that communication infrastructures played in coordination with other effects on the revolutions.

2.2 Using new media to study the Arab Spring

Having given our stance on the question of whether or not social media played a role in the revolution, we now turn to how data from social media has been used to better understand the processes inherent to the revolution as a whole. In particular, we focus on Twitter, as this is the data available to us in the present study. Of the many studies which have focused on Twitter in the context of the Arab Spring, we here highlight X facts

Bruns et al. (2013) work considers usage patterns in tweets that had the hashtags #egypt or #libya over the course of ten months in 2011.

One of the earliest works to focus on the use of Twitter during the Arab Spring was that of Lotan et al. (2011). This observation

Information structures differed across countries

Lotan et al. (2011) The authors stress the strong interplay between Twitter and news media, and showed that the relationships between mainstream media outlets, activists, journalists and bloggers differed in datasets collected from Tunisia and Egypt.

Bruns et al. (2013) Further, they observed that Egypt and Libya displayed differing amounts of tweets in the two languages, with Egypt having significantly more tweets in Arabic than English, and Libya the other way around.

Lotan et al. (2011) For example, journalists and activists in both cases appeared to serve as

primary information sources, with their efforts being retweeted by a variety of other types of actors. These interested parties fall into roughly three categories: 1. People directly connected to an incident, either as residents or expatriates that want to know about dangerous conditions and the state of their homes and families, or who are experiencing a crisis event firsthand; 2. MSM who want to learn about developments on the ground so that they can provide upto- date coverage across media channels and hold audience attention; and 3. General interest readers who want to know about events as they happen.

Bruns et al. (2013) They found, consistent with a variety of previous work, that Arabic and English speaking users mostly interacted amongst themselves, and thus that there existed only a few actors who bridged the two communities. This suggests that international interests varied and that within these nations there was varying levels of use of Twitter. Finally, Bruns et al. (2013) observed that these patterns changed over time, with the English-speaking world gradually losing interest.

Borge-Holthoefer, Magdy, Darwish, and Weber (2014)

We apply our methodology on a set of nearly 6 million Arabic tweets crawled between June 21 and September 30, 2013. From these, we reconstructed the network of Twitter users who authored tweets in the collection, and we crawled all meta information for 120,000 users along with their latest 3,200 tweets prior to December 2013. All the tweets and Twitter user IDs in our collection are publicly available[41], so that researchers can replicate the exact dataset we used for possible future studies.

We look at switching between Secularist and Islamist camps and between pro-military and anti-military camps. Our network and content analyses indicate that less than 5% of users switched sides. Instead, the main narrative seems to be one of pro-military intervention and Secular users being dominant in terms of volume leading up to July 3, and anti-military intervention and Islamist

users gaining in volume afterwards. Furthermore, in contradiction to the dominating narrative in mass media, the correlation between being a secular and a supporter of military intervention is far from perfect. However, some correlation was noticed between being an Islamist and against the military intervention.

we focus on Arabic tweets

Many, many other studies on Twitter. Methods discussed in their context as appropriate.

2.3 The role of the news media

(hatem)

Difference between really biased newspapers ... who said nothing was really happening ... and the international news who could cover it reasonably

Egyptians relied heavily on rumors ... when you have a situation where everyone knows the news is propaganda .. you rely on rumors more

Also when the media is obviously lying ... I'm looking at this streety as its being recorded ... the typical egyptian deligimizes the news

they also get a lot of international news from satellite TV ...

Hussain and Howard (2013) international news organizations played in giving them the global exposure to help stave off overtly violent reactions from security forces.

(Goldstone, 2011)

The impact of public media in favor of the protestors is also greater if media representation shows protestors as representative of the whole society, rather than as one particular group seeking partisan advantages for itself. For these reasons, virtually all successful revolutions

3 Conclusion

Our analysis is cowith the acknowledgement of the rash of recent claims over the methodological trapdoors that exist within this type of data Joseph, Landwehr, and Carley (2014); Morstatter, Pfeffer, Liu, and Carley (2013); Ruths and Pfeffer (2014); Tufekci (2014). While these issues are not to be ignored, the present work utilizes cautious, nuanced analysis techniques which account for, or at least admit, these possible biases.

References

Battera, F. (2014, October). Perspectives for change in tunisia, egypt and syria: the military factor and implications of previous authoritarian regimes. *Contemporary Arab Affairs*, *7*(4), 544–564. Retrieved 2015-01-29, from `http://dx.doi.org/10.1080/17550912.2014.948312` doi: 10.1080/17550912.2014.948312

Bellin, E. (2012). Reconsidering the robustness of authoritarianism in the middle east: Lessons from the arab spring. *Comparative Politics*, *44*(2), 127–149. Retrieved 2015-01-29, from `http://www.ingentaconnect.com/content/cuny/cp/2012/00000044/00000002/art00002`

Borge-Holthoefer, J., Magdy, W., Darwish, K., & Weber, I. (2014, October). Content and network dynamics behind egyptian political polarization on twitter. *arXiv:1410.3097 [physics]*. Retrieved 2015-01-14, from `http://arxiv.org/abs/1410.3097` (arXiv: 1410.3097)

Bruns, A., Highfield, T., & Burgess, J. (2013). The arab spring and social media audiences english and arabic twitter users and their networks. *American Behavioral Scientist*, *57*(7), 871–898. Retrieved 2014-04-21, from `http://abs.sagepub.com/content/57/7/871.short`

Comunello, F., & Anzera, G. (2012, October). Will the revolution be tweeted? a conceptual framework for understanding the social media and the arab spring. *Islam and Chris-*

tian–Muslim Relations, *23*(4), 453–470. Retrieved 2015-01-29, from http://dx.doi.org/10.1080/09596410.2012.712435 doi: 10.1080/09596410.2012.712435

Dewey, T., Kaden, J., Marks, M., Matsushima, S., & Zhu, B. (2012). The impact of social media on social unrest in the arab spring. *International Policy Program*. Retrieved 2015-01-29, from http://stage-ips.stanford.edu/sites/default/files/shared/2012%20Practicum%20Policy%20Brief%20SocialMedia.pdf

Gallé, M., Renders, J.-M., & Karstens, E. (2013). Who broke the news?: an analysis on first reports of news events. In *Proceedings of the 22nd international conference on world wide web companion* (pp. 855–862). International World Wide Web Conferences Steering Committee. Retrieved 2014-02-03, from http://dl.acm.org/citation.cfm?id=2488066

Gelvin, J. (2015). *The arab uprisings: what everyone needs to know* (2nd ed.). Oxford University Press. Retrieved 2015-02-03, from https://books.google.com/books?hl=en&lr=&id=8pDVBQAAQBAJ&oi=fnd&pg=PR3&dq=The+Arab+Uprisings:+What+Everyone+Needs+to+Know&ots=DCiavf98sQ&sig=sWNe3IkAVHGwD3SRfzyu-KYv5uI

Goldstone, J. A. (2011). Cross-class coalitions and the making of the arab revolts of 2011. *Swiss Political Science Review*, *17*(4), 457–462. Retrieved 2015-01-06, from http://onlinelibrary.wiley.com/doi/10.1111/j.1662-6370.2011.02038.x/full

Goldstone, J. A. (2013). Bringing regimes back in–explaining success and failure in the middle east revolts of 2011. *Available at SSRN 2283655*. Retrieved 2015-01-06, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2283655

Hussain, M. M., & Howard, P. N. (2013). What best explains successful protest cascades? ICTs and the fuzzy causes of the arab spring. *International Studies Review*, *15*(1), 48–66. Retrieved 2015-01-14, from http://onlinelibrary.wiley.com/doi/10.1111/misr.12020/full

Joseph, K., Carley, K. M., Filonuk, D., Morgan, G. P., & Pfeffer, J. (2014, December). Arab

spring: from newspaper data to forecasting. *Social Network Analysis and Mining*, *4*(1), 1–17. Retrieved 2014-03-30, from http://link.springer.com/article/10.1007/s13278-014-0177-5 doi: 10.1007/s13278-014-0177-5

Joseph, K., Landwehr, P. L., & Carley, K. M. (2014). An approach to selecting keywords to track on twitter during a disaster. In *Proceedings of the 11th international conference on information systems for crisis response and management.* State College, PA.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & Boyd, D. (2011). The revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International Journal of Communication*, *5*, 1375–1405. Retrieved 2014-01-31, from http://www.aish.es/files/laura/Revoluciones%20tuiteadas(1).pdf

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. *Proceedings of ICWSM*. Retrieved 2013-08-13, from http://www.public.asu.edu/~fmorstat/paperpdfs/icwsm2013.pdf

Pfeffer, J., & Carley, K. M. (2012). Rapid modeling and analyzing networks extracted from pre-structured news articles. *Computational and Mathematical Organization Theory*, *18*(3), 280–299. Retrieved 2013-05-01, from http://link.springer.com/article/10.1007/s10588-012-9122-1

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, *346*(6213), 1063–1064. Retrieved 2015-01-05, from http://emma.memect.com/t/8990484048dfe7fe71582552389de2c91f9fc7968100af273a59ee068737127e/Science-2014-Ruths-1063-4.pdf

Starbird, K., & Palen, L. (2012). (how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work* (pp. 7–16). ACM. Retrieved 2014-01-31, from http://dl.acm

.org/citation.cfm?id=2145212

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *ICWSM '14: Proceedings of the 8th international AAAI conference on weblogs and social media.* Retrieved 2014-04-15, from http://arxiv.org/abs/1403.7400

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication*, *62*(2), 363–379. Retrieved 2015-01-29, from http://onlinelibrary.wiley.com/doi/10.1111/j.1460-2466.2012.01629.x/abstract doi: 10.1111/j.1460-2466.2012.01629.x

Wolfsfeld, G., Segev, E., & Sheafer, T. (2013, April). Social media and the arab spring politics comes first. *The International Journal of Press/Politics*, *18*(2), 115–137. Retrieved 2015-01-29, from http://hij.sagepub.com/content/18/2/115 doi: 10.1177/1940161212471716