Evolution of Sentiment in the Libyan Revolution

Working Paper for the NSF Minerva Project: *Modeling Discourse and Social Dynamics in Authoritarian Regimes, NSF 0904913 at the University of Texas at Austin*

VERSION OF OCTOBER 30, 2011

Christopher Brown
The University of Texas at Austin

Joey Frazee
The University of Texas at Austin

David Beaver
The University of Texas at Austin

Xiong Liu
Intelligent Automation, Inc.

Fred Hoyt
University of New England

Jeff Hancock
Cornell University

1 Introduction

In the last ten days, Libya has changed, with effects on millions of Libyan's lives. How do they feel about it?

Until recently, such a question would have been laughably unanswerable. But the advent and recent development of social media has brought us to a position where we can not only sensibly ask such questions, but begin to answer them. We will do just that. It is, of course, only a beginning, and for several reasons. Here are two. First, social media is in its early days and is by no means universally accessible for the people of Libya or any other country. Second, when people are able to express themselves and do so using some medium, we still cannot be sure what they mean. Even if we knew each and every blogger, facebook-status updater, and tweeter personally, we still would be faced with a complicated interpretation problem. So there are challenges. But the new technologies also provide a historically unrivaled opportunity, an opportunity to take a look at what hundreds of thousands of people think about a developing political situation in realtime as it unfolds. And that is our goal.

This is a preliminary report on results from an ongoing study of language use in the Arab spring. In the broader study, we are looking at both the language of leaders in public speeches, and the language of the populace in social media. Here we concentrate on just one country, Libya, and focus only on the language of the general public, as seen in tweets.

2 Analysis

We now present preliminary results for a first set of studies. The bulk of the results below are drawn from 5,842 unique Libyan tweets produced in the eight days from 10-15-2011 to 10-22-2011, though we also report on data that extends later than this, and on earlier data collected several months prior to Gaddafi's death. The tweets were collected by keyword using the Twitter "search" API (rather than the "fire-hose", which indiscriminately provides the user with huge amounts of Twitter data in realtime, and from which tweets of particular interest would then have to be filtered). We include only Arabic language tweets (as marked by Twitter) within an 800 mile radius of Waw an Namus (24.92, 17.77). The keywords included the English words "libya", "qaddafi", "gaddafi", "gadhafi", "sirte", "saif", "nato", "tripoli", "benghazi", and "misrata" along with their Arabic counterparts.

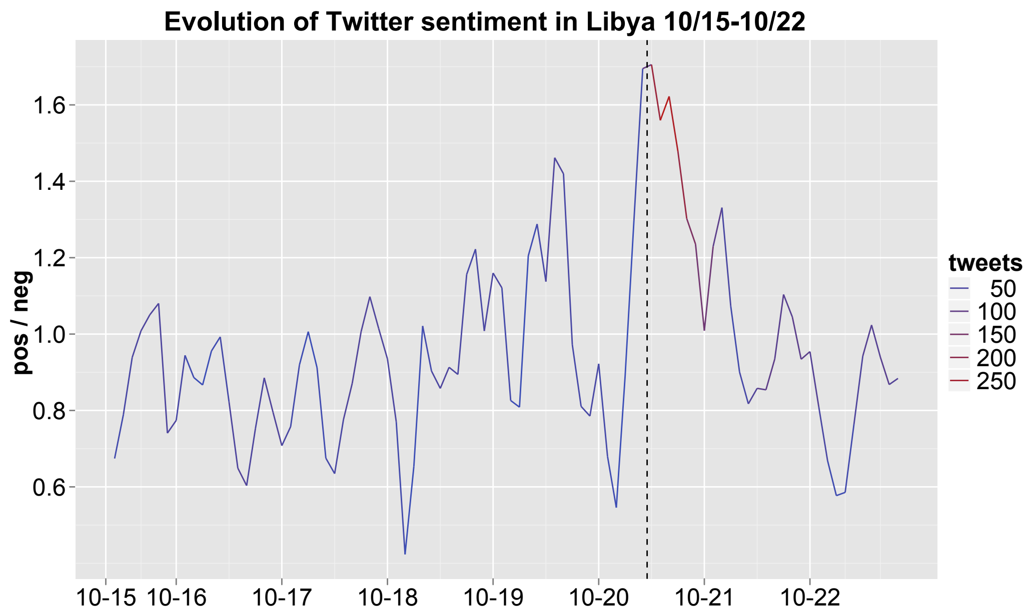

Figure 1 Sentiment during the last week of the Libyan revolution. Note that the peak corresponds with the time that Gaddafi's death became official (mid-morning on Thursday, 10-20-2011). The y-axis is the "sentiment ratio" — the ratio of positive-emotion words over negative-emotion words. Higher values indicate a bias towards positive emotion words. Values around 1 indicate a balance of emotion. The color of line shows the volume of tweets in the dataset.

To identify location, we use metadata provided by Twitter. In some cases this will be a geocode created at the time the user tweeted, for example if this information was provided by a cellular device used for the tweet. Twitter uses a fallback mechanism for geocoding so that in absence of a geocode it returns tweets for users whose profile location matches the search criteria.

Once the tweets were gathered, we removed all re-tweets and non-Arabic tweets. The remaining Arabic tweets were subsequently translated into English using Google translate. The output from the translation procedure was then evaluated using Linguistic Inquiry & Word Count (LIWC; Pennebaker, Booth & Francis 2007a, Pennebaker, Chung, Ireland, Gonzales & Booth 2007b). LIWC provides a dictionary which can be used to classify

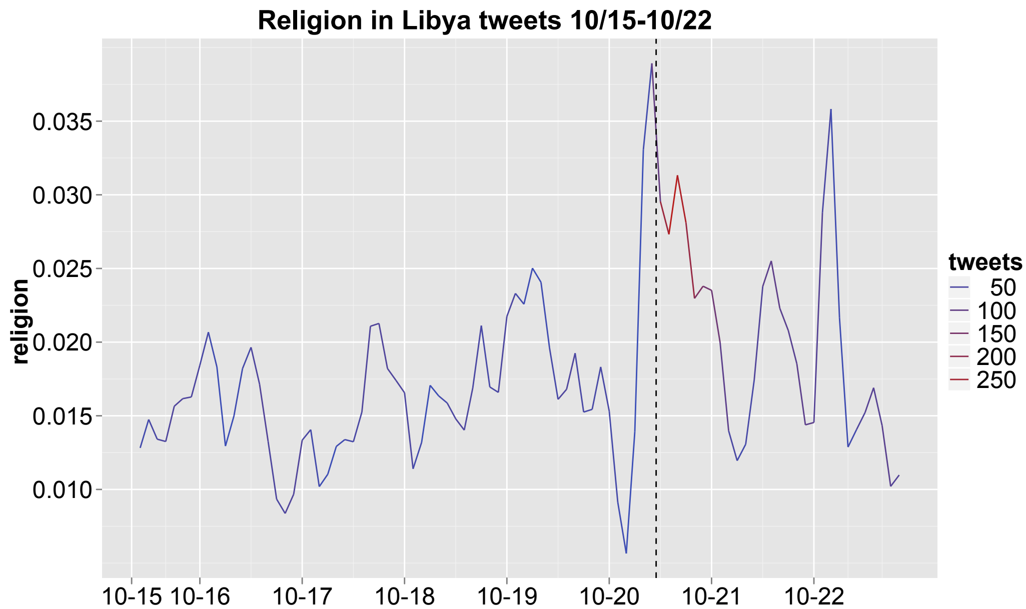

Figure 2 Religion during the Libyan revolution. The y-axis is the rate of religion-related terms (e.g., "god", "allah", "sacrifice", "gospel"). The number of occurrences of terms like these is normalized against the total number of words in a given time period. Again, the color of line shows the volume of tweets in the dataset.

words into linguistically, socially and psychologically meaningful categories. For example, there are classes of positive emotion terms like "good" and "wonderful", and classes of religion-related terms like "prayer".

The tweets were then sorted into two hour blocks between 10am EET on 10-15-2011 and 8pm EET on 10-22-2011 (EET = Eastern European time, which is the timezone of Tripoli) and a six hour moving average was calculated for various LIWC-based metrics. The graphs in Figures 1–4 illustrate the resulting time-series. The first graph shows changes in overall sentiment, calculated as the ratio of positive to negative emotion words. When the messages in aggregate contain more positive words, sentiment trends up and when they contain more negative words, sentiment trends down. This way of representing sentiment has been used previously to link the Twitter stream to traditional public opinion polls (O'Connor, Balasubramanyan, Routledge & Smith 2010). The remaining graphs each focus on a single LIWC category: *religion*, *anger*, and *death*.

In interpreting figure 2, it should be born in mind that many high-positive exclamatives in Arabic are religious formulas. The fact that the figure shows a spike around the time Gaddafi was killed is thus consistent with the separate observation of highly positive sentiment at that time. It cannot be assumed that the large amount of religiously oriented language reflects the writers' degree of piety *per se*.

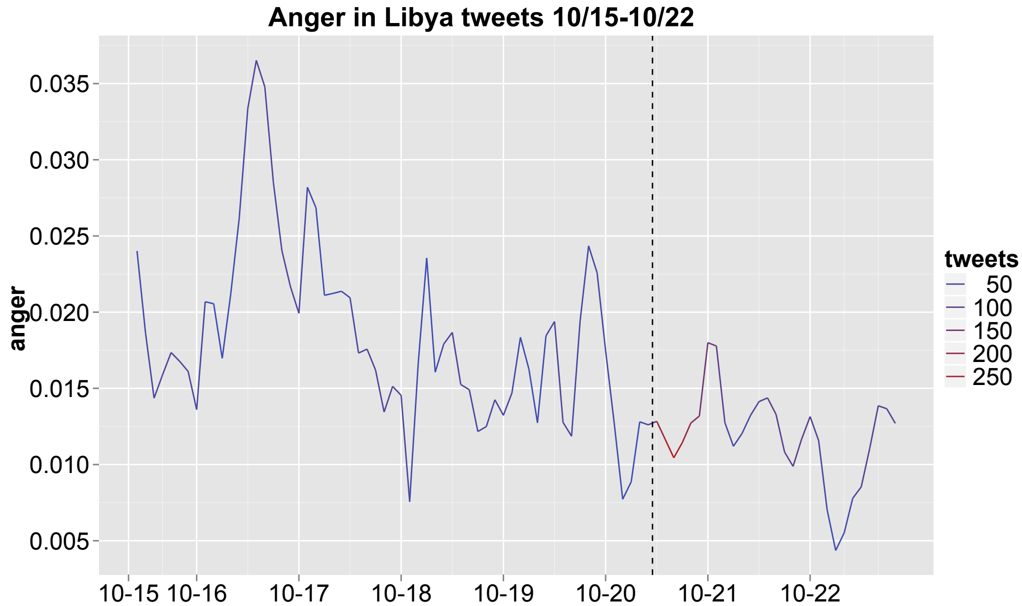

Figure 3 Anger during the Libyan revolution. The y-axis is the rate of anger-related terms (e.g., “brutal”, “danger”, “assault”, “revenge”).

Figure 5 plots positive and negative emotion separately over a slightly extended time period: the black line is the time at which AP tweeted this (in English): “BREAKING NEWS: Witnesses say Libyan fighters overrun last positions of Gadhafi loyalists in Sirte, city falls”.

For comparison with the October 2011 data, we also give one graph, 6, of data from tweets in an earlier period of the Libyan revolution, (March 2011–May 2011): this graph also utilizes automatically translated tweet data, though the collection methodology and geographical distribution of the tweets differs from those that make up the bulk of the current report.

3 Discussion

One complication in analyzing this data is that Twitter volume (rate of tweets) fluctuates dramatically, but predictably, throughout the day. As might be expected, the rate of tweeting climbs throughout the day low. While the general shape of volume throughout the day is consistent from day to day, the volume increases substantially throughout the day following the announcement of Gaddafi’s death, and remains quite high over the following several days (Figure 7).

Some of the data points appear out of proportion because it is difficult to convey sentiment ratios alongside volume. For example, very early on October 17th, ‘Anger’ and

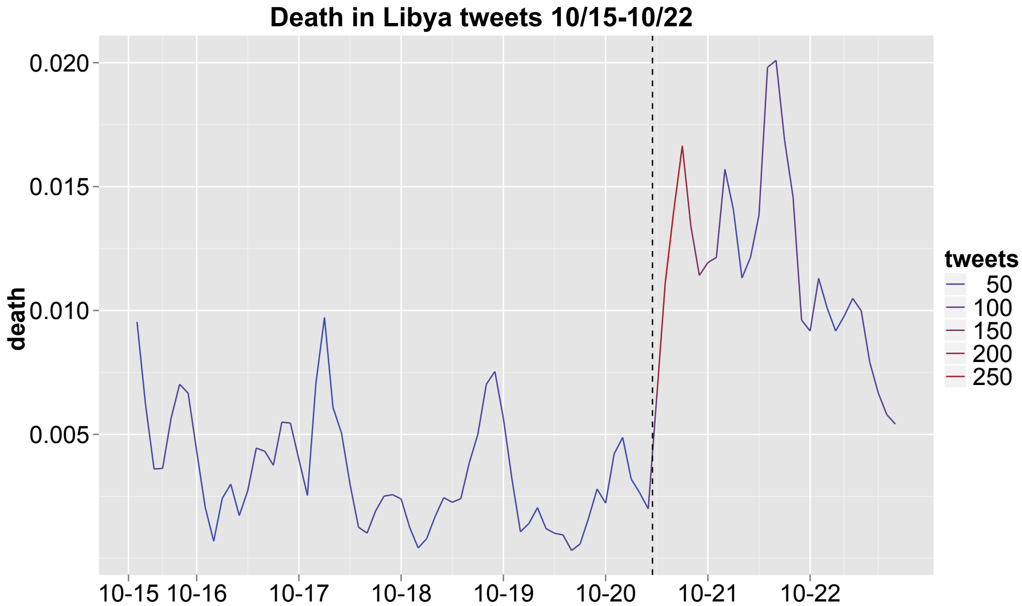

Figure 4 Death during the Libyan revolution. The y-axis is the rate of terms relating to the concept of death (e.g., “death”, “dead”, “slaughter”, “burial”).

‘Death’ peak at the same moment. There were eleven tweets in this period, and four are retweets of this one:

- “. . . Khamis Qaddafi killed for the 17th time . . . Shorouk_News channel (opinion) in favor of the Qaddafi regime to confirm the killing of his son Khma” (5:32 am).

The other tweets talk about thunderstorms in Tripoli, war, and internet connectivity. The problem is not the off-topic tweets, it is the sparseness of the data. Obviously, the disproportionately high number of mentions of *death* and *anger* words is not significant. Conveying this visually is tricky.

Many analyses of Twitter data aggregate millions of tweets spanning many timezones, smoothing out the sorts of idiosyncrasies inherent in tweets. We do not have this luxury. Our attention is focused on a short period of time and small geographical area so there are no more than a few thousand relevant tweets available from each day.

Our solution has been to bin the tweets into six hour blocks of time, smoothing out the periods of extremely low volume. This means that the height of the line in Figures 1–4 does not reflect the underlying variation in the rate of tweeting, since all the graphs represent ratios rather than absolute numbers of words in particular LIWC categories. For Figure 1, this is a ratio of two separate LIWC categories. For Figures 2–4, the plot represents the ratio of words in a particular LIWC category occurring in a set of documents to the total number of words occurring in that set of documents.

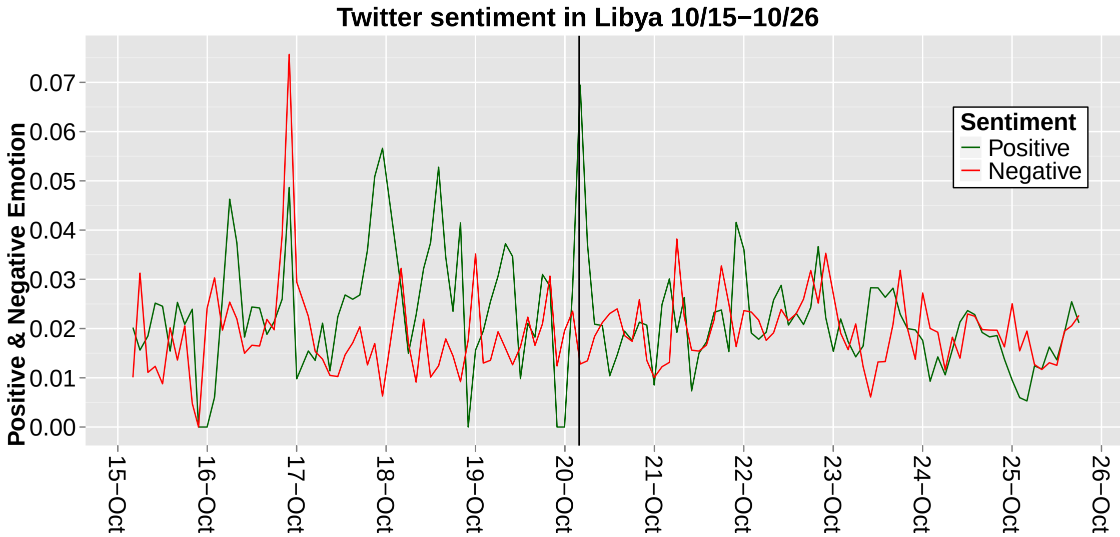

Figure 5 Plot for an 11 day period in October 2011 with one vertical marking the first news of the rebel victory at Sirte.

Despite these problems, we can see definite trends and changes that seem to correspond closely to the time of Gaddafi's death. In 1, particularly, there is a significant boost in positive emotion in the segment of tweets analyzed by our filter.

Here are few of the significantly positive tweets from that period:

- "God is great, God is greater for the largest agency Elaallah and God is great and thank God for the freedom of free people of Libya and Libya's pride lived Libya is free" (4:12 pm).
- "Allahu Akbar Allahu Akbar Allahu Akbar of the agency Elaallah and God is great and all praise to the Libyan people freedom, dignity, free of Libya Libya lived free independent" (3:14 pm) ('Allahu Akbar' basically translates to 'God is Great')
- "Libya's Muammar free I'm Matt Libby a free and proud Libya is free" (7:16 pm)

These were automatically translated from the original Arabic by Google Translate. Some portions of the translations are questionable, "The likes of Muammar al-Muammar, the dead!! It was like God, **God is alive to senior underwriter** .... God is great and thankfully lived Libya is free...", but the LIWC classifier is able to capture the general sentiment. While many details will be lost in translation, our analysis does not require perfectly faithful, grammatical translations. LIWC is based solely on vocabulary (unigrams), and machine translation is much better vocabulary than grammar. Logistically, though, machine translation is basically the only choice; it is fast and cheap; translating 57,381 tweets via Google's Translate API costs just $51.00 and takes a matter of hours.

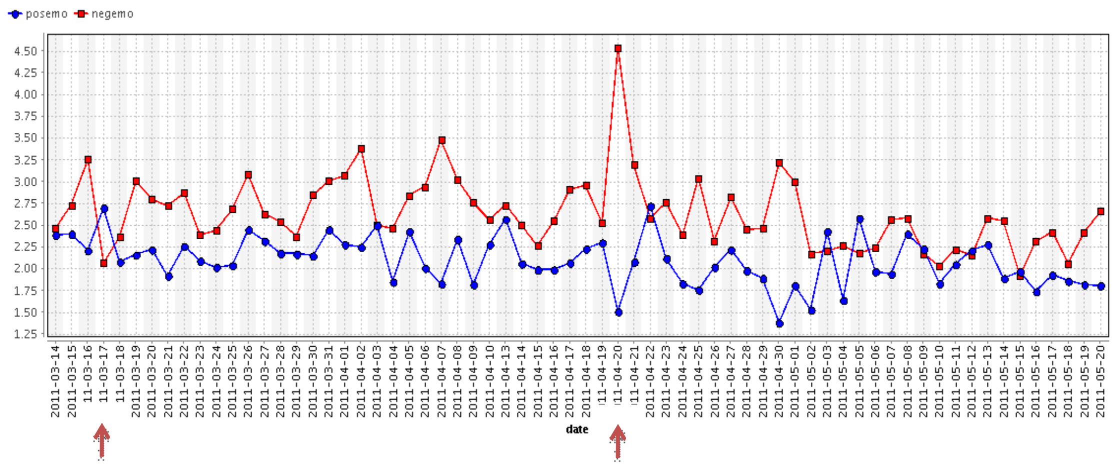

Figure 6 Positive and negative emotion in Libyan tweets, March–May 2011

While a relatively small proportion of Twitter data is geocoded with precise latitude and longitude coordinates. Figure 8 overlays some of the tweets on a map of northern Africa. In this static format, the data is not very informative. But an interactive or moving visualization of trends over time could make it possible to observe public response to geopolitical change.

Another strategy that may help with the problem of fluctuating Twitter volume is to model that fluctuation and account for that in weighting and displaying sentiment ratios. As we can see in an autocorrelation analysis of the volume in Figure 9, the volume is actually surprisingly predictable.

It should be noted that in this preliminary study, we have not discriminated according to whether the tweet is from an organization (e.g. a news organization) or from a private individual, though the bulk of tweets fall in the latter category. Identifying the main news organizations tweeting, and filtering those tweets out, would be relatively straightforward. However, the situation is complex, since individuals commonly retweet or quote from news sources such as Reuters, al-Jazeera, etc. One might reasonably wonder whether passing along a news quote is an expression of an individual's relatively non-emotive attitude to an event to the same extent that tweeting "Allah akbar!" five times in a row would typically be an expression of a rather highly emotive attitude. By not distinguishing corporate messages or part-corporate messages from individual messages, we implicitly assume that it is.

Our use of a lexicon of terms designed for English with translated Arabic tweets is quite obviously problematic. This leads to two questions: how problematic is it, and what can be done about it?

As regards the question of how problematic the methodology is, we have performed

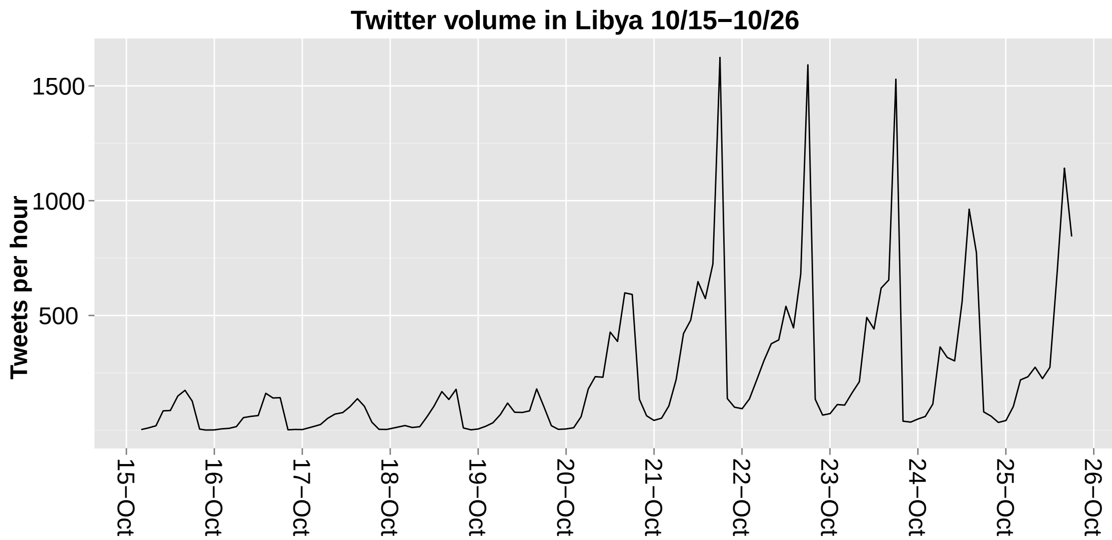

Figure 7 Total Twitter volume in our Arabic-only dataset. Gaddafi's death corresponds closely with a significant boost in volume that begins on October 20th and continues throughout the following days.

a qualitative study of the errors introduced in our sentiment analysis, but have not yet performed a systematic statistical error analysis. We studied a sample of 300 tweets in the original Arabic, drawing the samples from one set classified as positive, and one set classified as negative. Qualitatively, we found that while tweets classified as positive were largely indeed positive (as measured by judgment of an Arab speaker on our team, Fred Hoyt), there was a significant number of tweets that our automatic analysis had classified as negative, when to an Arab speaker they were clearly positive. Of 150 tweets that we had automatically classified as negative, 26 were found to be clearly positive as judged by a human reader. However, some of the problems we discovered do not amount to straightforward "bugs" in the system. Rather, the question is one of what we want to measure: many terms that are commonly associated with negative emotions (terms to do with *death* being particularly salient here) can be used by people who are in a heightened state of positive emotion.[1]

Accepting that the current methodology, though it allowed for a rapid rough-and-ready analysis of current events, is flawed, let us move onto the question of what can in principle be done about it. There are two obvious directions. First, given a systematic human sample of the Arabic language data, the results could be adjusted to reflect any systematic biases. Second, we could make the move to direct analysis of Arabic texts. A separate strand of

1 Without wishing to trivialize an event of great historical import, consider, by analogy, the spectacle in a children's classic of colorful munchkins singing "Ding dong, the witch is dead". The song is clearly joyous, but by contemporary US standards somewhat dark. Our analysis is simply too course-grained at the moment to capture the subtlety of joy in the midst of inherently dark subject matter.

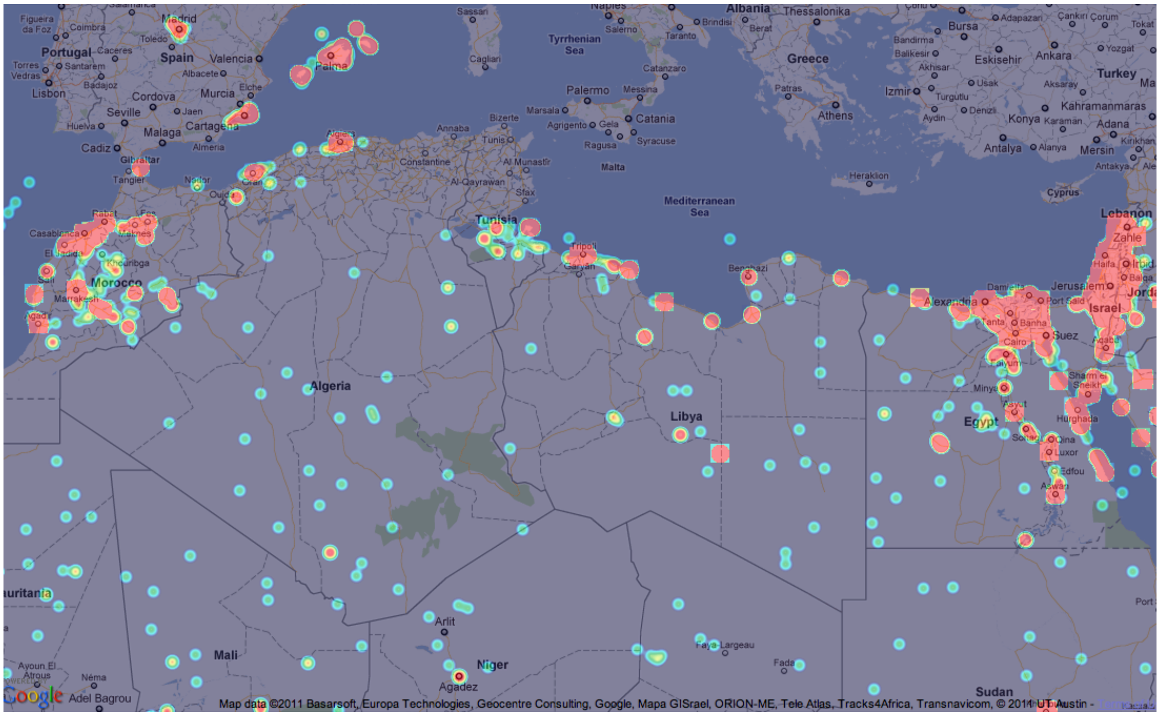

Figure 8 Tweet volume for about 100,000 tweets originating in northern Africa and nearby, over eleven days centering on October 20th.

work being undertaken by our broader project team involves development of Arabic LIWC Hayeri, Chung, Booth & Pennebaker (2010). Arabic LIWC is now available, although it does not cover the full range of categories of English LIWC, and thus requires further development before it could be applied to duplicate the analyses here. Note that there is prior work involving creation of an Arabic sentiment lexicon for machine classification, e.g. Farra, Challita, Assi & Hajj 2010. It would be worthwhile to compare performance of both work based on an Arabic sentiment lexicon and work based on a translation methodology such ours against a human-annotated gold standard. There can be little doubt that a carefully constructed Arabic lexicon would yield better results, but how much better?

In conclusion, what we have is noisy data, and a method which introduces yet more noise, but the analysis provides an indication of local exhilaration around the time of Sirte's overthrow and Gaddafi's death, just as our data from the earlier period of the Libyan revolution reveals positive emotion at the time of the declaration of the no-fly zone, and negative emotion at the time of a widely reported shelling incident in Misrata.

References

Farra, Noura, Elie Challita, Rawad Abou Assi & Hazem Hajj. 2010. Sentence-level and document-level sentiment mining for arabic texts.

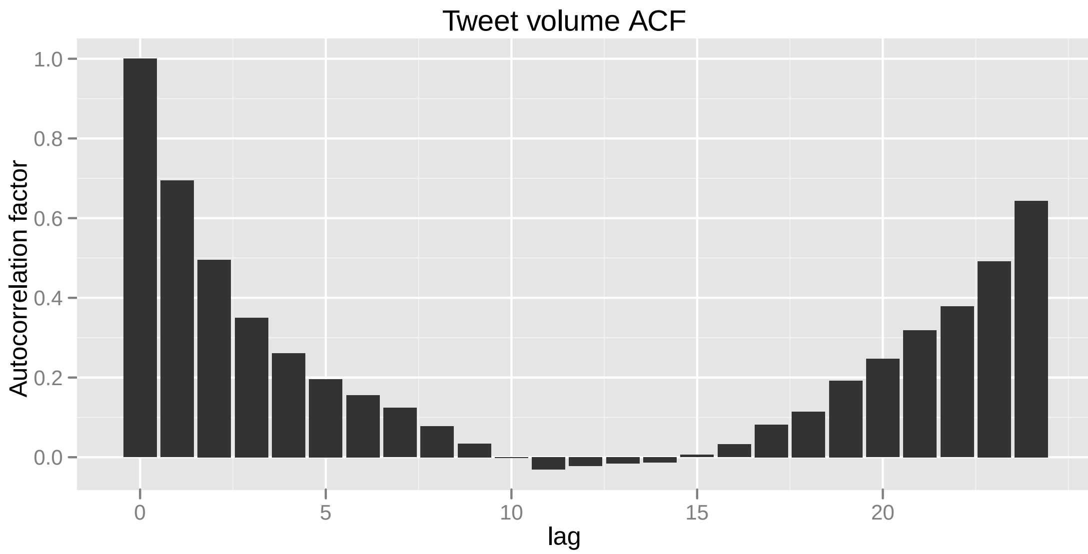

Figure 9 Lag correlation is very regular. This could be modeled to smooth down areas of extreme sentiment ratios but low volume.

Data Mining Workshops, International Conference on 0. 1114–1119. doi:http://doi.ieeecomputersociety.org/10.1109/ICDMW.2010.95.

Hayeri, Navid, Cindy Chung, Roger J. Booth & James W. Pennebaker. 2010. LIWC for Arabic texts. Austin TX: www.LIWC.net.

O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge & Noah A. Smith. 2010. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the fourth international AAAI conference on weblogs and social media*, .

Pennebaker, James W., Roger J. Booth & Martha E. Francis. 2007a. Linguistic Inquiry and Word Count (LIWC2007): A text analysis program. http://www.liwc.net.

Pennebaker, James W., Cindy K. Chung, M. Ireland, A. Gonzales & R.J. Booth. 2007b. The development and psychometric properties of LIWC2007. http://homepage.psy.utexas.edu/homepage/students/Chung/Publications_files/PennebakerChungIrelandGonzales&Booth2007_LIWC.pdf.