EMBERS at 4 years: Experiences operating an Open Source Indicators Forecasting System

Sathappan Muthiah[1], Patrick Butler[1], Rupinder Khandpur[1], Prithwish Chakraborty[1], Saurav Ghosh[1], Parang Saraf[1], Nathan Self[1], Liang Zhao[1], Jose Cadena[2], Chang-Tien Lu[1], Bert Huang[2], Anil Vullikanti[2], Bryan Lewis[2], Achla Marathe[2], Madhav Marathe[2], Theodoros Rekatsinas[3], Elaine O. Nsoesie[4], John S. Brownstein[*5], Sumiko R Mekaru[*], Lise Getoor[○], Khoa Trinh[†], Aravind Srinivasan[†], Kristen Summers[**], Graham Katz[‡], Andy Doyle[‡], Jaime Arredondo[¶], Dipak Gupta[‖], David Mares[¶], Naren Ramakrishnan[*]

[1] Virginia Tech, 900 N.Glebe Rd., Arlington, VA 22203
[2] Biocomplexity Institute, Virginia Tech, Blacksburg, VA 24061
[○] University of California at Santa Cruz, Santa Cruz, CA 95064
[¶] University of California at San Diego, San Diego, CA 92093
[‖] San Diego State University, San Diego, CA 92182
[†] University of Maryland, College Park, MD 20742
[4] University of Washington, Seattle, WA 98195
[5] Harvard Medical School, Boston, MA 02101
[3] Stanford University, Stanford, CA 94305
[§] BASIS Technology, Herndon, VA 20171
[*] Epidemico, Boston, MA 02101
[‡] CACI Inc., Lanham, MD 20706
[**] IBM Inc.

ABSTRACT

This paper gives an account of the successful forecasts made by the EMBERS system and reasons out cases where EMBERS didnt perform well.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

1. INTRODUCTION

Anticipatory intelligence.

EMBERS.

Many regions, but here only LA.

Civil unrest, but also diseases.

There are other regions and other event classes we don't consider.

Contributions are:

1. We are honest characters. We give both successes and failures.

2. Rather than quantitative metrics, we talk about value to an analyst.

3. One more - what?

2. CIVILUNREST FORECASTING

The following section details out some of the successful and not so successful forecasts made by EMBERS over the past few years in Latin America.

2.1 Successful Forecasts

Described below are some examples of successful civilunrest forecasts made by the EMBERS system during 2013-2015.

Brazil Spring (June 2013): These protests were the largest and most significant protests in Brazil's recent history that caught worldwide attention. Millions of Brazilians took part in these demonstrations, also known as the Brazilian Spring or the Vinegar Movement (inspired from the use of vinegar soaked cloth by demonstrators to protect themselves from police teargas), which were catalyzed by an increase in public transport fares from $R\$3$ to $R\$3.20$ by the government of President Dilma Rousseff.

As shown in Figure. 1 EMBERS, while missing the initial uptick, captured the increase in the order of magnitude of the protest events during the Brazilian Spring and also captured the spatial spread in the events, in addition to forecasting that this would be a "General Population" protest.

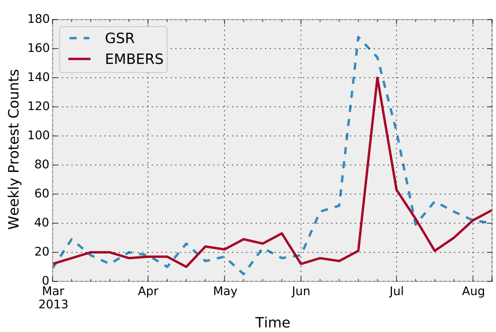


Figure 1: Brazil Spring

Around 68% of EMBERS Brazil alerts resulted from detected planned protest which can be explained from the fact that the social networking sites (Twitter and Facebook) and conventional news media played a key role in organization of these uprisings. Although, initial protests were mostly against the bus fare increase soon these grew from general population's deep dissatisfactions to include wider issues such as - government corruption, over-spending and police brutality. Also, demonstrators made calls for political reforms. In response, President Rousseff proposed a referendum on widespread political reforms in Brazil, but was later abandoned. EMBERS models were able to capture such discussions on Twitter (see Figure. 2), and followed these stories as they evolved through June.

Figure 2: Word cloud from data extracted by EMBERS models

The protests intensified in late June (see Figure. 1), which were captured by EMBERS, as they also coincided with FIFA 2013 Confederations Cup matches. This was an important factor due to which protests gained momentum as they were covered by world media. Majority of protests occurred in those cities, which were hosting (FIFA) soccer matches. EMBERS submitted most of its alerts for these host cities (see Figure 3)— Rio de Janeiro, São Paulo, Belo Horizonte, Salvador and Porto Alegre, among others. For example, on 27th June during Confederations cup semi-final in Fortaleza, around 5000 protestors clashed with the police near the Castelao stadium - In this case EMBERS sent out an alert the day before. Then later on 30th June, when the last games of the confederation cup took place in Rio de Janeiro and Salvador that were also plagued by mass protests – EMBERS predicted these events and submitted multiple alerts for Rio on 28th and 29th June and one for Salvador on 29th June.

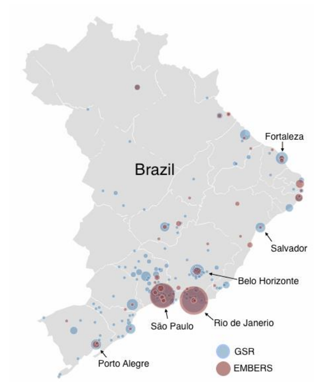


Figure 3: Geographic overlap of Groud Truth protest events and EMBERS alerts for Brazil June 2013.

Venezuelan Spring (Feb-March 2014): EMBERS captured some of the first 'calls to protest' for the trigger city of San Cristobal and its nearby surrounding areas and correctly forecast the population (Education) and that the protests would turn violent. Over the next days, EMBERS closely forecast the spike in the number of events as shown in Figure. 4 and the spread of the protests to additional cities as shown in Figure. 5.

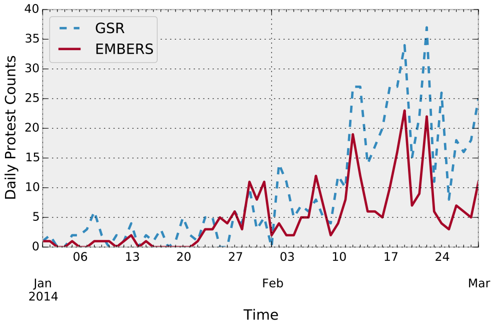


Figure 4: Venezuelan Spring

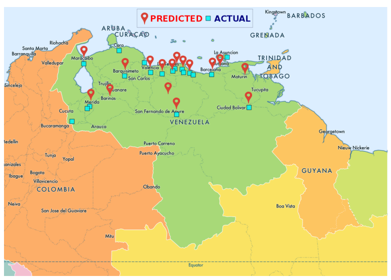


Figure 5: Venezuelan Spring

Mexico Protests (October 2014): EMBERS , as shown in Figure. 6 forecast an uptick of Mexico protests during early October 2014 stemming from students and teachers demanding action on the 43 students missing case, with a lead time of about 3 days. It also generated a series of alert spikes coinciding with the first large-scale nationwide protests between October 5th to 8th.

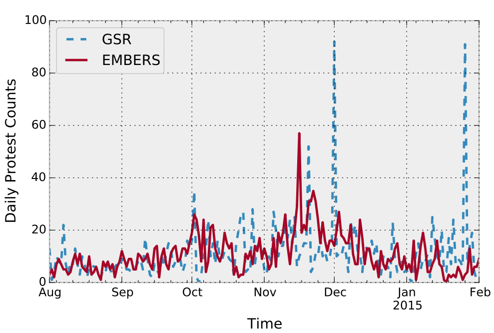


Figure 6: Mexico Protests

Figure. 7 provides a timeline of GSR events and EMBERS alerts for Mexico during this period

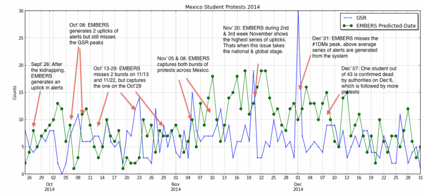


Figure 7: Mexico Protest timeline.The plot indicates the comparison between counts of GSR events and EMBERS alerts' for unique states/admins per day.The EMBERS series' as can be seen is more pronounced during Nov-Dec'14,

Colombia Protests (December'14 -March'15): EMBERS successfully forecast the uptick in the number of events during the middle of December 2014 and also the increase in protest counts during February 2015 as shown in Figure. 8, though in the latter case EMBERS over predicted the counts. The uptick in December 2014 was led by the opposition leader Alvaro Uribe against impunity. Whereas the increase in protest counts in February 2015 was due to trucker's strike against increase in fuel price.

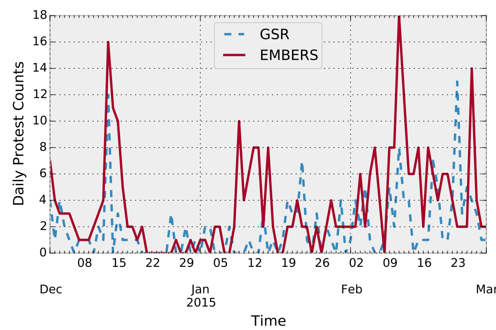


Figure 8: Colombia Protests

Paraguay Protests (February 2015): EMBERS forecast the uptick in number of protest events in Paraguay during mid February 2015 as shown in Figure. 9. The events were mainly due to the lack of opportunity and basic needs and against the introduction of new public-private partnership law.

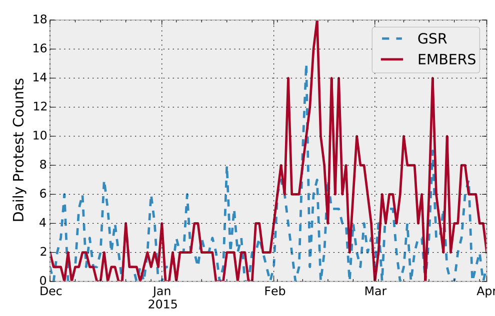


Figure 9: Paraguay Protests

2.2 Events missed by EMBERS

: The following section details out certain events that EMBERS failed to capture properly.

Brazil Protests (March 2015): EMBERS predicted the accurate increase in the rate of events but does not capture the true counts. The protests were mainly targeted against president Dilma Rousseff due to increasing corruption.

During this period there was a significant architectural change in the EMBERS processing pipeline. EMBERS had moved to Heideltime temporal tagger from the previously used TIMEN temporal tagger due to Heideltime‘s support for more languages and active development cycle as opposed to TIMEN. Heideltime had no support for portuguese (the main language used in Brazil) and EMBERS team had extended Heideltime to portuguese by translating the resources for spanish to portuguese. It turned out that simple translation of rules from spanish to portuguese were not sufficient and this affected the recall of one of our main models for Brazil - Planned Protest - as it depended on the quality and recall of date extraction from text.

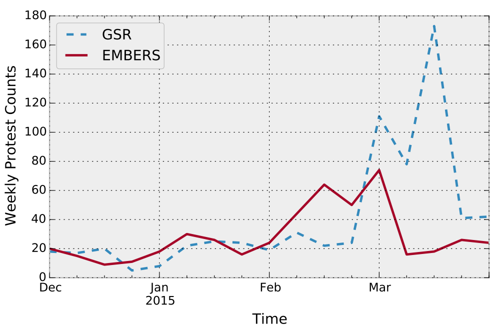


Figure 10: Brazil 2015 Protests

Mexico December 1, 2014 Protests: EMBERS missed the huge single day spike on December 1st when people turned out in huge numbers in different parts of Mexico demanding Pena Nieto‘s ouster. EMBERS predicted nationwide events for December 1st but failed to capture the its spread.

December 1st was picked by the protestors due to its historical significance - it was the day when Pena Nieto was sworn in as President in 2012 amidst much controversy and opposition from general public. Also another reason for the missed prediction was due to EMBERS inability in extracting dates mentioned using twitter lingo like "#1Dmx".

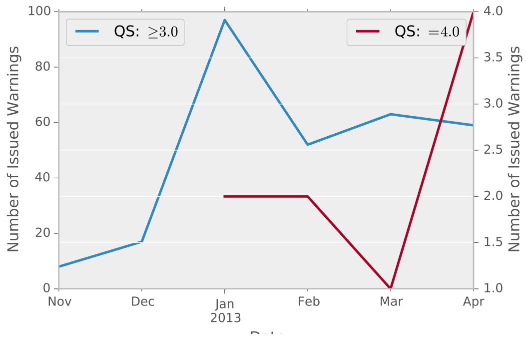


Figure 11: Number of High quality ILI warnings over time.

3. INFLUENZA-LIKE-ILLNESS FORECASTING

In this section we analyze the forecasts generated for Influenza-like-Illness or ILI events. For ILI, we concentrated on short-term forecasts for the first 2 years of EMBERS. We shited our focus to long-term forecasts for the subsequent years. We analyze both types of forecast here for general performance and analyze in details a few succesful forecasts which showcased the strength our system. There were a number of EMBERS ILI forecasts which significantly deviated from the target sources. We analyzed these scenarios and discuss in details some of the weakness of EMBERS. These scenarios also helped us to increase the robustness of our system.

3.1 Successes

Our models for ILI were tuned over time to incorporate our findings over time. Over the course of our project we sent a number of perfect warnings (QS$=4.0$) as well as significant number of high quality warnings (QS ≥ 3.0). For example, we sent several ILI case count warnings for Chile for event date 08/07/2013. The actual value for the said date was 626 while we sent a first warning for 581 (QS: 3.71) and a second update (QS: 3.95). Figure 11 shows a temporal history of high quality warnings. As can be seen, the number of perfect warnings sent increased over time without sacrficing the number of high quality warnings significantly.

We also show a distribution of the sent warnings, by combining updates for each event date, in figure **??**. As can be seen, we can see a dual peak distribution with the first peak around the mean score and the second one around the perfect score of 4.0.

During the course of the project we were allowed to update our warnings, if needed, before the event is published. Figure 12 showed the distribution of scores for warnings by taking the mean scores of all the updates issued. We further analysed our update scheme in Figure **??**. As can be seen we see a significant upward shift in our score from first update to last update. We can also see that if the best update could have been predicted our scores would have been higher.

For our long term forecasts, we were able to send perfect predictions for a number of countries such as Bolivia for a number of event types such as peak date for flu/SRV and Total Flu count.

3.2 failures

We also analysed our performance for significant low scores.

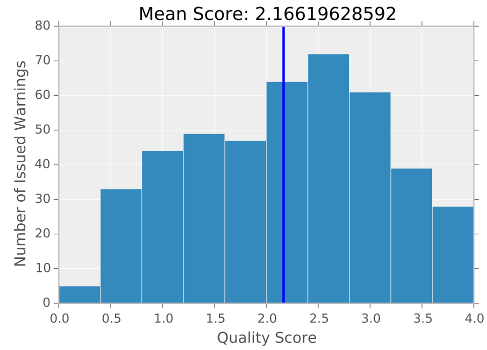


Figure 12: Histogram of Quality scores for ILI warnings combininig updates for each event date.

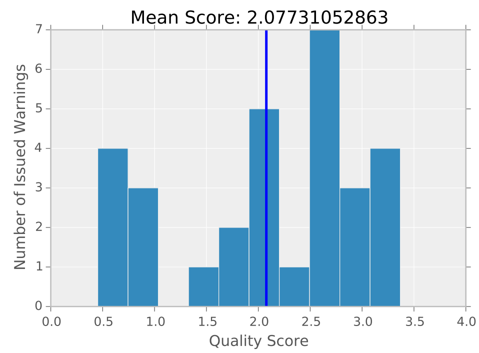
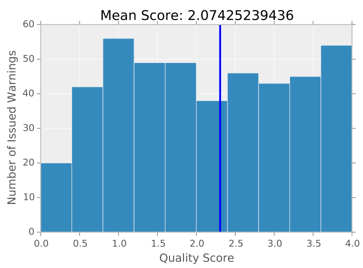
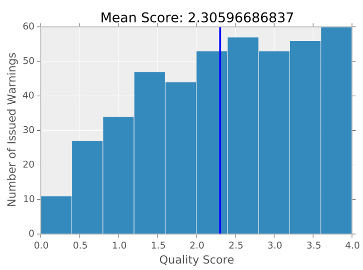


Figure 14: Histogram of Quality scores for ILI warnings for Argentina.

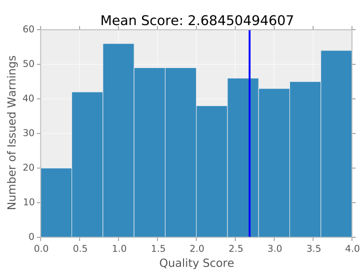
Figure 14 shows the distribution of scores for Argentina where we performed poorly w.r.t to other countries such as Bolivia. As can be seen the distribution is unimodal around the mean. The lack of our perfect warnings could be attributed to more unstable nature of ILI surveillance in the country as well as the physical size of the country where a single surveillance center may be composed of trends from several disparate regions.



(a)



(b)



(c)

Figure 13: Histogram of Quality Scores for ILI warnings for each event date. (a) Considering the first update, (b) Considering the last update, and (c) Considering the best update.

4. RARE DISEASE FORECASTING

One of the key event classes studied in EMBERS included forecasting outbreaks of four rare diseases (hantaviurs, cholera, yellow fever and machupo) in 10 countries of Latin America. The rare disease model employed a corpus of publicly available health-related news articles from HealthMap (cite) to extract topics about the mentioned rare diseases and their corresponding spatio-temporal patterns. The spatial and temporal distributions of rare disease topics were then utilized by 1-class SVM (cite) as features to predict the emergence of a rare disease outbreak at a future time point. This prediction is generated for each individual source where source refers to the publisher of the news articles, e.g. "www.biobiochile.cl" a prominent source reporting disease outbreak news in Chile. We had 798 different news sources extracted from the HealthMap corpus. To combine the predictions of multiple news sources, we used a multiplicative weights algorithm (cite).

4.0.1 Successful forecasts

EMBERS rare disease model successfully forecasted the hantavirus outbreaks in Chile and Argentina (2013 and 2014).

4.0.2 Failures

EMBERS rare disease model failed to forecast the cholera outbreak in Mexico during the month of October, 2013. One of the possible reasons is that this cholera outbreak spread to Mexico from its neighboring country Cuba, thus there was no prior signal about this outbreak in the HealthMap corpus. Alternative data sources, such as travel patterns could have helped us in forecasting this outbreak.

5. CONCLUSIONS

This paragraph will end the body of this sample document. Remember that you might still have Acknowledgments or Appendices; brief samples of these follow. There is still the Bibliography to deal with; and we will make a disclaimer about that here: with the exception of the reference to the LaTeX book, the citations in this paper are to articles which have nothing to do with the present subject and are used as examples only.

Acknowledgments

Supported by the Intelligence Advanced Research Projects Activity (IARPA) via DoI/NBC contract number D12PC000337, the US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the US Government.

6. REFERENCES