
Characterizing Social Response to Urban Earthquakes using Cell-Phone Network Data: The 2012 Oaxaca Earthquake

Benyounes Mounni

EPFL
Lausanne, Switzerland
benyounes.mounni@epfl.ch

Vanessa Frias-Martinez

Telefonica Research, Madrid,
Spain
vanessa@tid.es

Enrique Frias-Martinez

Telefonica Research, Madrid,
Spain
efm@tid.es

Abstract

The data generated by pervasive infrastructures, and specially cell-phone networks, has been used in the past to improve responses to emergency events such as natural disasters or disease outbreaks. However, very little work has focused on analyzing the social response to an urban earthquake as it takes place. In this paper we present a preliminary study of the social response using the information collected from a cell-phone network during the 2012 Oaxaca earthquake in Mexico. We focus our analysis on four urban environments located between 100-200km away from the epicenter of the earthquake. The social response is analyzed using four different variables: call volume, call duration, social activity and mobility. Initial results indicate a social response characterized by an increase in the number of calls, a decrease in call durations, a moderate increase in the number of people contacted by highly connected citizens and a moderate increase in the mobility.

Author Keywords

Call Detail Records, Big Data for Social Good, Earthquake

ACM Classification Keywords

K.8.1 [Personal Computing]: Miscellaneous.

General Terms

Algorithms, Experimentation, Measurement.

Introduction

In recent years, cell phones have become a pervasive technology with users carrying them at almost all times. From this perspective, the ubiquity of such platforms has transformed cell phones into one of the main sensors of human behavior. Cell phone traces have been used across various disciplines including the analysis of human and social response to natural disasters or epidemic spreading. For example, Bengtsson et al. [2] used cell phone traces from the main carrier in Haiti to trace population movements on a post-earthquake scenario and to model disease spread. Similarly, the government of New Zealand used cell phone traces to characterize short-term population movements after the 2011 Christchurch earthquake [8]. Bagrow et al. analyze the collective response to large-scale emergencies such as bombings, blackouts, earthquakes or big festivals [1]. On the other hand, Frias-Martinez et al. [4] used cell phone traces to characterize and model the 2009 H1N1 outbreak in Mexico and Tatem et al. [9] used cell phone traces to understand the spread of malaria.

Although other data sources such as GPS [6] and twitter [7] have been used to study human behavior during natural disasters, we consider that, specially in developing economies, the use of CDRs avoids the biases introduced by such data sources.

An important goal of these studies is to gain an understanding of social responses, which could be useful to implement or improve policies for future emergency situations. For example, Bengtsson et al. [2] found that after the Haiti earthquake citizens in the capital city,

where the earthquake mostly struck, left urban areas and went to the places where they typically spent Christmas; and Frias-Martinez et al. [5] found that the total closure of non-essential activities during an epidemic spread (as recommended by the WHO) provoked that citizens flocked to the airport twice as much before the closure was in effect.

In this paper, rather than analyzing the post-emergency events, we focus our study on the social response during an emergency in urban settings. Our objective is to understand, from a behavioral perspective, the immediate response of citizens to emergencies in urban scenarios. Such information can be of high interest to both sociologists studying how citizens cope with stressful situations as well as to policy makers responsible for setting up policies during emergencies.

We focus our analysis on the social behavioral changes that took place during the 2012 Oaxaca earthquake in Mexico. For that purpose, we gather cell phone network data from four mid- to large urban areas located at 100-200km from the earthquake epicenter and analyze the behavioral changes that take place in terms of communication (call volumes and call durations), social activity and mobility. We focus in urban environments (and urban earthquakes) given that cities are the most vulnerable settings for seismic damage [3].

2012 Oaxaca Earthquake

The earthquake had a 7.4 magnitude and struck at 12:02 local time on Tuesday, 20 March 2012. Its duration was approximately between 40 and 300 seconds. It had a focus of 15 to 20 km and its epicenter was near the border between the states of Guerrero and Oaxaca. The earthquake caused strong shaking over a large area

along the Oaxaca and Guerrero border and the adjacent Pacific coastline. Significant tremors were felt in areas up to several hundred kilometers away, including Mexico City. Two people were killed and over 30,000 houses were damaged or destroyed. Around 800 aftershocks were reported in the Guerrero and Oaxaca region, the strongest of them with a magnitude of 6.0. Although the magnitude of the earthquake was considerable, damages on infrastructures were small due to the high standards of Mexican building regulation. As a result the cellular infrastructure was not damaged and communications happened during the earthquake.

Calling Data

Cell phone networks are built using a set of base transceiver stations (BTS) that are in charge of communicating cell phone devices with the network. Each BTS tower has a geographical location typically expressed by its latitude and longitude. The area covered by a BTS tower is called a cell. Whenever an individual makes a phone call, the call is routed through a BTS in the area of coverage. CDR (Call Detail Record) databases are generated when a mobile phone connected to the network makes or receives a phone call or uses a service (e.g., SMS, MMS, etc.).

For this study we gathered a dataset containing a total of two weeks of CDRs before and during the earthquake for four urban areas that were affected by the earthquake with more or less intensity: Oaxaca and Veracruz at approximately 100km away from the epicenter, and Mexico City and Acapulco at approximately 200km from the epicenter. For analytical soundness, we also collected CDRs from the city of Monterrey located at more than 1000km away from the epicenter and which was not affected by the earthquake. From all the data

contained in a CDR, our study uses the encrypted originating number, the encrypted destination number, the time and date of the call, the duration of the call, and the latitude and longitude of the BTS tower used by the originating cell phone number and the destination phone number when the interaction happened. Note that no information about the exact position of a user in a cell is known. In order to preserve privacy, all the information presented is aggregated and original records are encrypted. No contract or demographic data was considered or available for this study and none of the authors of this paper collaborated in the extraction of the original data.

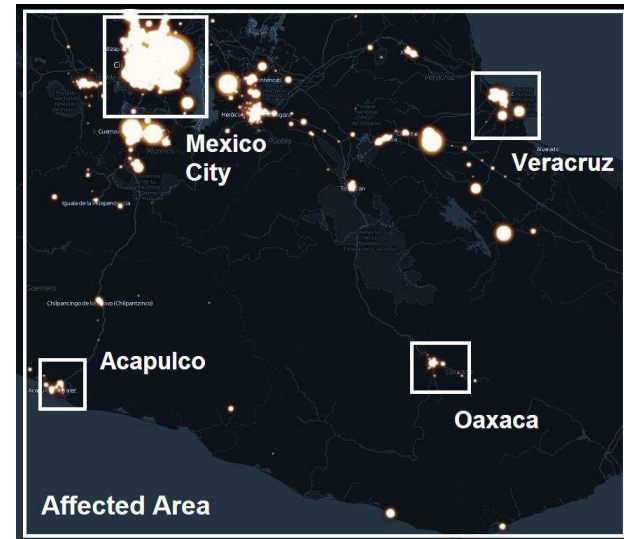


Figure 1: Geographical area affected by the earthquake and urban settings studied.

Methodology

In order to understand the social response to the earthquake, we analyze behavioral changes in five urban areas namely Mexico City, Oaxaca, Veracruz, Acapulco and Monterrey. Figure 1 presents the area affected by the earthquake and the cities we study (with the exception of Monterrey). Our objective is to evaluate how call volume, call duration, social activity and mobility change in the immediate moments after the earthquake.

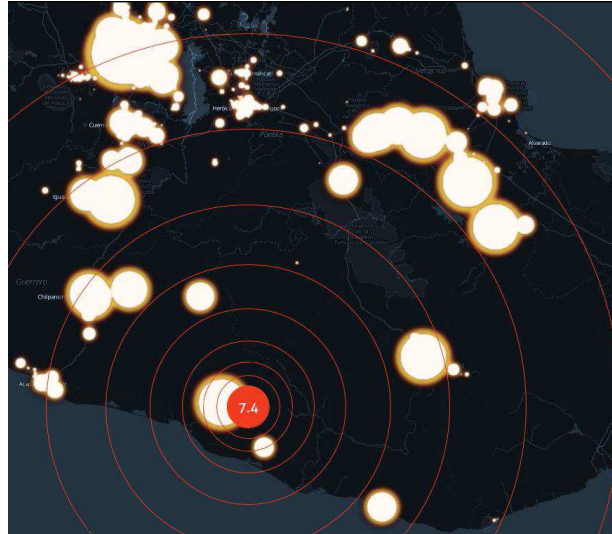


Figure 2: Call volume during earthquake.

We model each city with a set of distributions, one per behavioral parameter. Specifically, for each city and behavioral parameter we generate two signals: an 'earthquake' behavior signal and a baseline behavior signal. The former is computed using CDRs from the day of the earthquake, Tuesday March 20th for each of the cities under study. The latter is generated using CDRs from

two Tuesdays before the earthquake, Tuesday March 13th and Tuesday March 6th. We posit that averaged parameters from these two Tuesdays represent a baseline of normal behavior.

In order to understand changes in behavior, we will report quantitative differences between the earthquake and baseline behaviors together with statistical tests to understand the significance of the differences. Throughout, we apply the Lilliefors test to test for normal distributions followed by t-test or Wilcoxon rank sum and Kolmogorov-Smirnov (K-S) tests respectively.

Result Analysis

Next, we describe our analytical results for each of the four behavioral variables under study.

Call Volume

We define call volume for each urban setting as a time series containing the normalized number of calls initialized each minute. For our analysis, we compute two time series, one for the day of the earthquake and another one with the baseline, the two with a 1-minute resolution. Figure 2 shows the call volume per BTS and city exactly at 12:03, one minute after the earthquake, on March 20th. Compared to Figure 1, which represents the call volume at 12:03 during an average normal Tuesday (baseline), we can observe an important increase in the activity right after the earthquake takes place.

In order to get a better understanding of the impact of the earthquake at an urban level, we compute for each city, the difference between the normalized call volume time series during the earthquake and during the baseline. Figures 3 and 4 present the normalized differences for the cities of Oaxaca and Mexico City respectively.

We can observe that both cities show a positive spike right after the earthquake happens and it lasts around 5 minutes (12:02 to 12:08). However, the increase in call volume is five times larger for Oaxaca than for Mexico city. Running a similar analysis for Veracruz and Acapulco, we observe that cities that are closer to the earthquake epicenter show higher peaks in the call volume than cities farther away. This increase is followed by a negative spike with call volumes under the baseline the following two hours for Oaxaca and Acapulco and the following hour for Mexico City and Veracruz. Finally, the call volume increases until reaching normality at approximately 18:00.

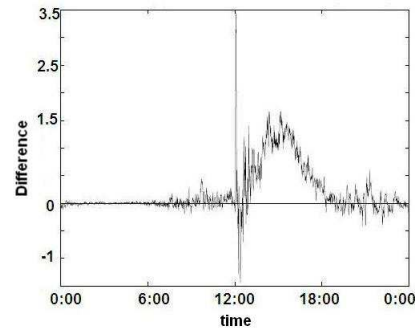


Figure 3: Differences in call volume between earthquake and baseline behaviors for Mexico City.

In general, we observe that the earthquake provokes an initial increase in the call volume with larger increases associated to cities closer to the epicenter. This period of panic is followed by a decrease in call volumes until normal behaviors are recovered approximately six hours after the event took place. We observed these behaviors across the four cities under study. On the other hand, no behavioral changes were observed in Monter-

rey as expected, which is more than 1000km away from the epicenter.

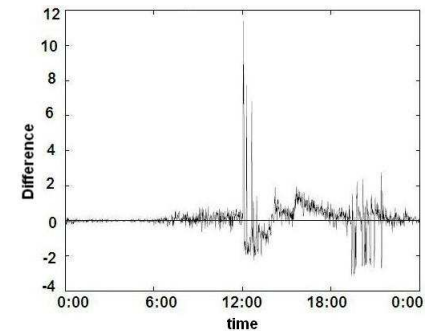


Figure 4: Differences in call volume between earthquake and baseline behaviors for Oaxaca.

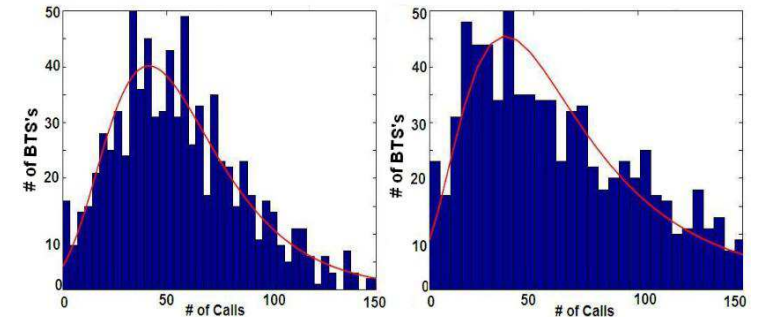


Figure 5: Number of BTSs (Y-axis) that handle each volume of calls (X-axis) in Mexico City during the (left) baseline (normal day) and (right) during the earthquake.

A complementary analysis is shown in Figure 5 which represents the number of BTSs that handle different volumes of calls between 12:00 and 12:10 in Mexico city. The left panel represents the volumes during the baseline and the right panel during the earthquake day.

Similarly to our previous analysis, we also observe an increase in the call volume as well as in the number of BTSs in the cellular infrastructure that were used to give such increased service. Finally, it is important to highlight that all four cities rejected the Wilcoxon Rank-Sum and K-S tests indicating that the increase in call volumes across these cities is statistically significant.

Call Duration

We define call duration for each city as a time series containing the average duration of the calls initialized at each minute. Similarly to call volume, we compute two time series, one for the day of the earthquake and another one with the baseline, the two with a 1-minute resolution. To understand the impact of the earthquake at an urban level, we compute the difference between the normalized call volume time series during the earthquake and during the baseline. Figure 6 shows such comparison for Mexico City.

We can observe a drastic reduction in the duration of the calls during the first ≈ 80 seconds after the earthquake. This might be due to the fact that citizens that started a phone call before the earthquake hung up immediately after they felt the earthquake tremors. However, we observe that the call duration is consistently shorter than the baseline during approximately three hours after the earthquake. After that, and up till the end of the day (from 15:00 to 23:59) there is an increase in the length of the calls. Additionally, similar results are obtained for all urban areas, with the exception again of Monterrey. From these results, we can hypothesize that the first reaction upon an earthquake might be to make short calls to social contacts to confirm they are sound and safe. Once citizens feel that the situation is normal again, they make longer calls than usual, probably

to hold discussions about the event. Distance to the earthquake epicenter does not seem to have an impact on the duration of the calls.

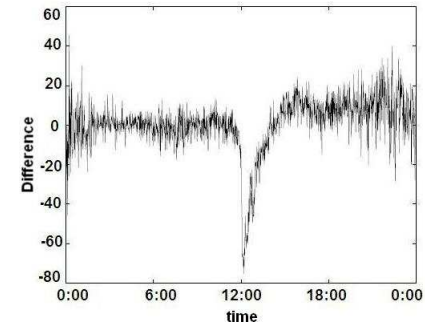


Figure 6: Differences in call duration between earthquake and baseline behaviors for Mexico City.

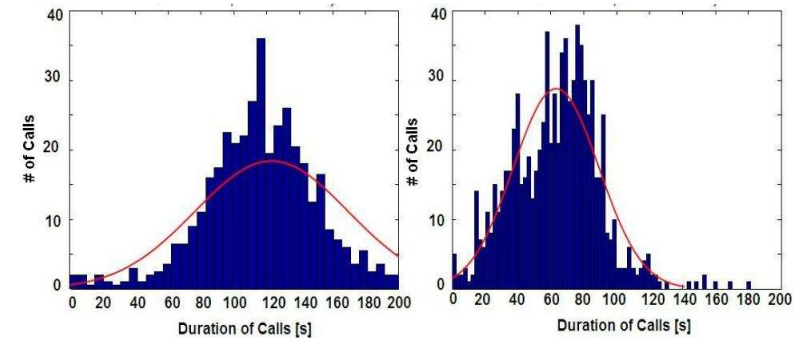


Figure 7: Number of calls (Y-axis) for each duration (X-axis) in Mexico City: (left) baseline signal and (right) earthquake signal.

Figure 7 shows the number of calls handled by the cellular network between 12:00 and 12:10 and their durations, during the first minutes after the earthquake for

Mexico City. The left panel shows the distribution during the baseline and the right panel during the earthquake. Confirming our previous analysis, we observe that the BTSs in Mexico City handle calls that are, on average, twice longer on a given normal day than on the earthquake day (right panel). Wilcoxon Rank-Sum and K-S tests rejected the null hypothesis for all cities under study, with the exception of Monterrey indicating that the differences between baseline and earthquake signals were statistically significant for all four cities affected by the earthquake.

Social Activity

As stated earlier, we have observed a significant increase in the call volume across all cities affected by the earthquake, especially between 12:00 and 18:00. In this section, we want to understand whether such increase corresponds to people reaching to friends outside their typical contacts or on the contrary whether the usual contacts are contacted multiple times during and after the earthquake.

For that purpose, we measure social activity for each urban environment as the number of individuals that communicated with a given number of contacts between 12:00 and 18:00. Specifically, we compute two distributions for each urban environment: one modeling the social activity during the earthquake and one modeling the social activity during the baseline. Figure 8 presents the histograms for both social activity distributions normalized by the number of users for Mexico City. On the other hand, the inset in the Figure shows the difference between both distributions. We can observe that users that on a given normal day communicate with four or more contacts, show a small increase in the number of different people contacted during the earthquake. How-

ever, callers that typically have three or less communications on a normal day, appear to reduce the number of people they contact during the earthquake. We hypothesize that highly connected users tend to expand their communications and reach to contacts they do not typically call whereas low-connected users do not appear to extend their communications beyond their usual contacts. Although similar results were observed across affected cities, the hypothesis could not be statistically proved given that the K-S test did not reject the null hypotheses in any of the urban environments under study.

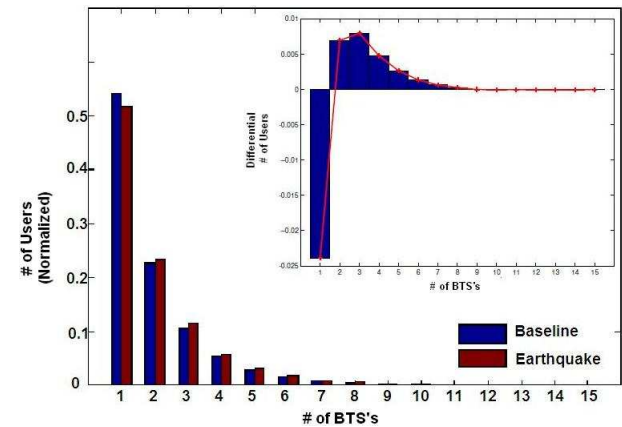


Figure 8: Normalized distribution of the number of different contacts called (X-axis) by users and (inset) the difference between both distributions.

Mobility

We measure mobility as the number of different BTS towers used by the individuals within a city between 12:00 and 18:00, hours when the communication activity is most affected by the earthquake. To understand the impact of the earthquake on the mobility of the population in the cities under study, we compute

mobility distributions both for the baseline and during the earthquake and compare values. In general, larger numbers of different BTS towers imply a higher mobility but not necessarily larger distances travelled. Although, intuitively, distance traveled can be considered a better proxy for mobility, we decided against it because the low granularity of BTS towers in certain parts of the cities implied that distances introduced a lot of noise in the real value of the metric. Finally, in order to make sure that we have sufficient individual information to model mobility, we build the mobility distributions with users that have at least five calls during both the baseline and the earthquake periods.

Figure 9 presents both distributions normalized by the number of users for Mexico City, and the inset in Figure 9 shows the difference between both distributions. We observe a small increase in the number of citizens that use two of more BTS towers during the earthquake when compared with the baseline. However, the number of citizens using just one BTS tower during the earthquake decreases. We can hypothesize that, in general, citizens that use various BTS towers during a normal day increased the number of BTS towers used (and thus its mobility) during the earthquake *i.e.*, citizens with high mobility move even more during the earthquake. However, citizens that do not typically move during the baseline (they only use one BTS) move even less during the earthquake. However, the hypothesis was not confirmed by the K-S statistical test.

Conclusions

In this paper we have presented preliminary results to characterize the social response to an urban earthquake. We have focused our analysis on the effects of the 2012 Oaxaca earthquake on five different cities located at ap-

proximately 100km and 200km away from the epicenter. Although the earthquake was of considerable magnitude, it did not cause a lot of destruction which probably determined the social behavior we observed during the earthquake. We acknowledge that our results will vary across types of earthquakes depending on its destructive power and across geographical boundaries. However, we believe that there is value in this work, especially for earthquake and disaster managers in Mexico.

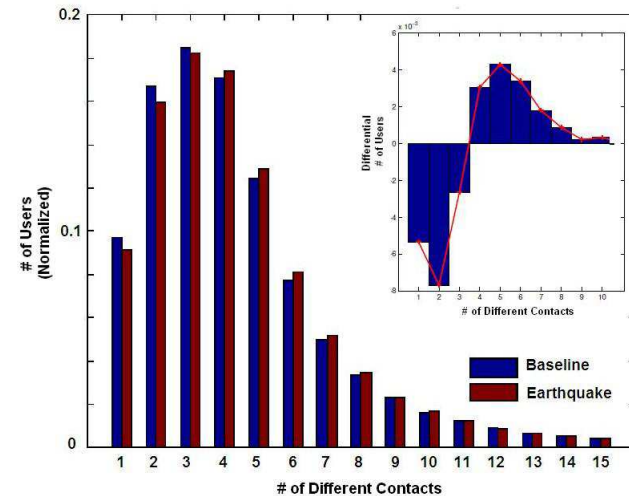


Figure 9: Normalized distribution of the number of different BTS towers used (X-axis) by callers and (inset) the difference between both distributions.

Our results indicate that the social response in terms of call and duration volumes can be divided in three stages: (1) the first stage, from the earthquake up to five minutes after the earthquake, consists of one spike of high activity characterized by a large number of short calls; (2) the second stage, which lasts between one

and two hours, is characterized by a reduced activity compared to the baseline and also short calls and (3) the last stage, that lasts around 5 hours, has a moderate increase in both call and duration volumes. These changes are observed across all four affected cities, although the increase in volumes tends to be larger as the city gets closer to the epicenter of the earthquake. We have also observed that highly connected users tend to expand their communications and reach to contacts they do not typically communicate with, probably to check that all their non-frequent contacts are sound and safe. Finally, we also report a small change in mobility whereas citizens appear to increase their mobility during the earthquake when compared to the baseline.

References

- [1] Bagrow, J.P. and Wang, D. and Barabasi, A.L. Collective Response of Human Populations to Large-Scale Emergencies. *PLoS ONE* 6, 3 (2010).
- [2] Bengtsson, L., Lu, X., Thorson, A., Garfield, R., and von Schreeb, J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in haiti. *PLoS medicine* 8, 8 (2011), e1001083.
- [3] Bilham R. Urban Earthquake fatalities - a safer world or worse to come? *Seismological Research Letters* (2004).
- [4] Frias-Martinez, E., Williamson, G., and Frias-Martinez, V. An agent-based model of epidemic spread using human mobility and social network information. In *2011 IEEE 3rd Int Conf Social Computing (SocialCom)*, IEEE (2011), 57–64.
- [5] Frias-Martinez, V., Rubio, A., and Frias-Martinez, E. Measuring the impact of epidemic alerts on human mobility. *Pervasive Urban Applications–PURBA* (2012).
- [6] Horanont, T., Witayangkurn, A., Sekimoto, Y., and Shibasaki, R. Large-scale auto-gps analysis for discerning behavior change during crisis.
- [7] Sakaki, T., Okazaki, M., and Matsuo, Y. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, ACM (2010), 851–860.
- [8] Statistics New Zealand. Using cellphone data to measure population movements experimental analysis following the 22 february 2011 christchurch earthquake new zealand government, isbn 978-0-478-37759-0 (2012).
- [9] Tatem, A. J., Qiu, Y., Smith, D. L., Sabot, O., Ali, A. S., Moonen, B., et al. The use of mobile phone data for the estimation of the travel patterns and imported plasmodium falciparum rates among zanzibar residents. *Malar J* 8 (2009), 287.