Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery

Paige Maas
Facebook
pmaas@fb.com

Shankar Iyer
Facebook
shankar94@fb.com

Andreas Gros
Facebook
andreasg@fb.com

Wonhee Park
Facebook
wonheepark@fb.com

Laura McGorman
Facebook
lmcgorman@fb.com

Chaya Nayak
Facebook
cnayak@fb.com

P. Alex Dow
Facebook
adow@fb.com

ABSTRACT

After a natural disaster or other crisis, humanitarian organizations need to know where affected people are located and what resources they need. While this information is difficult to capture quickly through conventional methods, aggregate usage patterns of social media apps like Facebook can help fill these information gaps.

In this paper, we describe the data and methodology that power Facebook Disaster Maps. These maps utilize information about Facebook usage in areas impacted by natural hazards, producing aggregate pictures of how the population is affected by and responding to the hazard. The maps include insights into evacuations, cell network connectivity, access to electricity, and long-term displacement.

In addition to descriptions and examples of each map type, we describe the source data used to generate the maps, and efforts taken to ensure the security and privacy of Facebook users. We also describe limitations of the current methodologies and opportunities for improvement.

Keywords

crisis mapping, crisis informatics, GIS, social media

INTRODUCTION

As social media and messaging apps continue to be important communication tools in people's everyday lives, they have also come to play an important role in how people prepare for, respond to, and recover from disasters (Palen and Anderson 2016; Castillo 2016). They are used by people affected by a crisis event, people responding to it, and people following and observing the event from afar (Olteanu et al. 2015). They can be used for individual and mass communication, information seeking, and gaining situational awareness. A significant body of research in crisis informatics has focused on studying these behaviors, and on developing techniques and tools for harnessing social media and other data sources for improved crisis response.

Many of these tools seek to extract information about what is happening in a crisis from the text of Twitter posts (Reuter et al. 2018), but a sparsity of location metadata can make it difficult to associate the extracted information with explicit locations. Alternatively, crisis maps collect relevant data on a map, frequently through the work of digital volunteers sifting through posts to both social and traditional media (Okolloh 2009), or by accessing geospatial data from official sources. The scale of social media suggests a potential for extracting unprecedented insights into how populations prepare for and react to hazard events, the impact those events have on these populations, and what needs these populations have that are not being met. Unfortunately, a dearth of precise location data can make much of this potential difficult or impossible to realize.

In the course of providing services to their users, many smartphones and smartphone apps regularly collect precise location information. In the case of Facebook, people have an option of whether or not to provide this information to Facebook (Facebook 2019). Location data is used to provide a myriad of services, including helping people find nearby friends, information about nearby Wi-Fi hotspots, and location-relevant ads. This data also enables targeting of AMBER alerts and prompts to check-in as "safe" after a hazard event. In addition to powering Facebook product features, this location data, when aggregated and anonymized, can provide insights about how populations are affected by hazard events as they happen.

While data from phones and apps has enormous potential, it also comes with notable risks for individuals. Location data is sensitive, and misuse could compromise the privacy and safety of individuals and communities. Any attempt to generate insights and share them with humanitarian responders must first address privacy and security to ensure that people are protected.

This paper presents Facebook Disaster Maps, a collection of methods for processing Facebook data into dynamic maps that highlight several key factors of how populations are preparing for, impacted by, and coping with natural hazards. The maps make use of anonymized and aggregated data, including current and historical location data, information about cell site connectivity, and data on phone battery charging. While the raw data for the maps remains available only to Facebook, the aggregated maps, with privacy and security protections like adding random noise and dropping small counts, are shared with humanitarian organizations on an ongoing basis in the days and weeks following a hazard event. The maps are meant to address specific needs that these organizations face in formulating and executing a response, therefore the research and development that goes into the maps has been (and continues to be) informed by discussions between Facebook researchers and individuals at the humanitarian organizations.

We first discuss existing research in the field of crisis informatics that has informed this work. Next, we survey five distinct categories of Facebook Disaster Maps, each of which aims to answer different questions that arise in formulating humanitarian response to hazard events. We discuss the methodology used in building each map and how privacy and security issues are addressed within that methodology. Next, we discuss current limitations on the utility of the maps and ongoing research that aims to overcome some of these limitations. Finally, we give examples of how each of these maps has been utilized by humanitarian organizations in the field.

RELATED WORK

In order to formulate an effective response to a crisis situation, first responders and relief organizations must combine many, diverse sources of information into "situational awareness," or a coherent, big-picture understanding of the situation. In a crisis, this situational awareness includes information about "the status of the hazard agent, damage done to buildings and infrastructure, the location of evacuation centers, and the number and location of injured people and/or animals" (Vieweg 2012), among other things. More generally, Sarter and Woods (1991) describe situational awareness as "all knowledge that is accessible and can be integrated into a coherent picture, when required, to assess and cope with a situation." While situational awareness is essential for some crisis response roles, the focus on a single "big picture" can be insufficient in some cases. It misses the fact that different crisis response roles (from a first responder to a digital volunteer to a humanitarian worker) have different information needs. Zade et al. (2018) argue for a focus on "actionability," which means understanding the different needs of people in different roles and focusing on getting them the information that they can act on.

Much existing crisis informatics research deals with enhancing situational awareness by identifying, extracting insights from, and generating predictions based on social media posts, mostly on Twitter (Vieweg et al. 2010; Cameron et al. 2012; Vieweg 2012; Imran et al. 2013). This is based on the idea of social media users acting as "citizen sensors" (Sheth 2009) who report on conditions as they experience them through their social media posts. Due to the high volume of social media posts during crisis situations, researchers have deployed myriad algorithmic and statistical methods for processing the streams, including natural language processing (NLP), semantic analysis, supervised machine-learned modeling, unsupervised clustering, and others (Castillo 2016).

A prominent approach to organizing and displaying crisis data for increased situational awareness is in the form of a crisis map. Crisis mapping involves the collection, analysis, and geospatial visualization of data. The data involved may come from a variety of sources, including social media and other citizen reports, traditional news reports, and official sources. For example, Google Crisis Maps[1] pulls together and displays public information, including storm paths, flood zones, evacuation routes, and more on top of Google Maps. Crowdsourcing and digital volunteers are used to collaboratively construct crisis maps and other datasets for improved situational awareness. Ushahidi (Okolloh 2009) is a popular and flexible crowd-sourced crisis mapping platform, which allows volunteers to, among other things, pin reports relevant to a crisis onto a shared map. In the immediate aftermath of the January 2010 earthquake in Haiti, volunteers used Ushahidi to collect reports about conditions on the ground from Twitter and other sources on a map (Starbird and Palen 2011; Meier 2012). Another notable achievement was volunteers using satellite photos to improve OpenStreetMap in the affected areas (Soden and Palen 2014). This later gave rise to the Humanitarian OpenStreetMap Team[2], which uses open mapping to support humanitarian action and community development more broadly. Castillo (2016) includes a detailed list of crisis mapping platforms and other tools for analyzing and visualizing insights from social media crisis data.

In their survey of research on social media in crisis management, Reuter et al. (2018) describe several types and aims of existing work related to the work described in this paper. In particular, they discuss the processing of social media data for monitoring ongoing events, and visualizing the results of analysis in order to make these results more salient and useful. These methods and insights frequently culminate in the design and building of tools and systems that address problems that arise in the course of crisis response. Some researchers employ participatory design, where professional crisis responders are involved in the design process of these systems (Hughes 2014; Hughes and Shah 2016). As Reuter et al. (2018) discuss, evaluation procedures for these systems vary widely and are often limited. Hughes and Shah (2016) evaluated the effectiveness of their Twitter-based crisis monitoring application by sitting with and observing public information officers using their tool while responding to wildfires. This allowed them to identify concrete contributions their tool made to the analysis, documentation, and reporting on social media produced by the participants.

Several of the data processing and mapping methods described in this paper highlight how population densities and movement patterns in crisis situations differ from pre-crisis situations. Human mobility has been shown to be highly predictable at city (Yan et al. 2014) and regional/national scale (Simini et al. 2012) in non-crisis times. Some research has found that even though mobility (e.g., travel distances and destinations) does change during crises, it can remain highly predictable (Lu et al. 2012). Wang and Taylor (2016) affirm this over a wide variety of hazard types and impacted regions, though they also found that some of the most severe crises can impact mobility patterns such that they are no longer predictable using pre-crisis states. This suggests that near-real-time insights into human mobility can be useful for gaining situational awareness and administering humanitarian assistance. To that end, Ciravegna et al. (2018), built a system for incorporating location tracking into mobile phone apps, processing the collected location data, and generating near-real-time insights into human movement patterns that may be useful for crisis response and recovery. For example, their visual analytics interface displays the frequency that different routes are taken, which, when compared across crisis and pre-crisis time periods, may highlight dynamic road access constraints. This paper, while using different data sources, builds on this line of work.

DISASTER MAPS DATASETS AND METHODOLOGY

In this section, we describe how each of the five Facebook Disaster Maps are produced. While we start with a brief overview of each type of map, we then describe concepts, privacy protection mechanisms, and computations that are shared by some or all of the map types. Once we have explained each of these, we discuss each type of map individually, including walking through examples based on real events.

Overview of the maps

As mentioned above, there are five distinct map types that make up Facebook Disaster Maps. Here we list each of them and state a fundamental question that each map type is intended to help answer. The first four listed maps rely on a shared calculation procedure where we compare counts of certain events during a crisis to baseline expectations from a pre-crisis period. In most cases, we produce these maps at regular intervals for two weeks after a crisis begins. Over these relatively short time scales, it is plausible that the main causal driver of population-level changes is the crisis event itself. The fifth map type, Displacement maps, relies on a different methodology.

[1]https://support.google.com/crisismaps/
[2]https://www.hotosm.org/

- **Facebook Population:** Where are there more or fewer Facebook users than we would expect based on pre-crisis levels? This can indicate areas that are affected by the crisis, or where evacuations are occurring.
- **Movement:** Which pairs of places are Facebook users moving between more or less often than we would expect based on pre-crisis levels? This can provide signal on large-scale population movements.
- **Power Availability:** Where are Facebook users charging their mobile phones more or less often than we would expect based on pre-crisis levels? In aggregate, this charging behavior can serve as a proxy for the state of the power grid.
- **Network Coverage:** Where does the usage of mobile phones indicate that there are more or fewer cell sites actively serving a location than we would expect based on pre-crisis levels?
- **Displacement:** In the weeks and months following a crisis event, where has the affected population resettled? Unlike the other map types, Displacement maps do not rely on a pre-crisis baseline computation, which can introduce additional complexities and caveats when interpreting the data. This will be discussed in detail in a later section.

Common concepts

There are several concepts relevant to understanding the maps. First of all, we construct maps using two different methods of identifying locations: tiles and administrative polygons. The **Bing Maps Tile System** defines a series of grids at different resolution levels over a rectangular projection of the world (Schwartz 2018). Each level is constructed by dividing the previous level into fourths. We typically use Bing tile levels 13 through 16, where level 13 results in tiles that are about 4.9 x 4.9 km at the Equator. The other method we use for identifying a location is **administrative polygons**, which define the political and geographic boundaries of countries, states, provinces, counties, cities, and more.

When generating a map for a crisis event, we specify a rectangular **bounding box** around the most directly affected area. The different map calculations, described in the following sections, are done relative to this region, and, for most of the maps, only data within this region is included.

Most of the map types are based on counting events that occur within a **time interval**, which is frequently 8 or 24 hours. The time interval determines what data is included in a calculation as well as the minimum frequency with which new maps are generated.

Privacy protection mechanisms

We employ a suite of privacy protection mechanisms in order to obscure the identity and actions of individuals and small groups while preserving the population-level insights that are useful for humanitarian response. Some of these are only applicable to a subset of the map types, which we will indicate in the sections below.

- **Random noise:** A small amount of random noise is added to count data to ensure that it is not possible to ascertain precise, true counts for sparsely populated locations.
- **Spatial smoothing:** We average counts for a location with those of surrounding locations using inverse-distance-weighted averaging. This gives more weight to closer locations and less weight to further locations.
- **Dropping small counts:** Locations with small counts are dropped from the final datasets. In cases where there are both baseline and crisis-time counts (as described in the next section), if either is less than a threshold value then both are dropped.

Common computations for maps with baselines

As mentioned above, four of the map types (Facebook Population, Movement, Power Availability, and Network Coverage) use a shared procedure for computing a pre-crisis baseline that can be compared to observations during and after a crisis. A baseline is computed for each location in a map (which can be a tile, an administrative polygon, or, as we will see for Movement maps, pairs of tiles or polygons). The duration of pre-crisis data available for computing a baseline differs for different source datasets, and is dependent on technical capacity and data retention policies. While these can change over time, all the map types described here use baselines derived from 5-to-13 weeks of pre-crisis data.

When computing a map for a given time interval after a crisis, we account for normally occurring daily and weekly patterns by computing a baseline using only data from the same time-of-day and day-of-the-week in the period preceding the crisis. Therefore, for a given location and time interval, the baseline dataset is composed of a set of counts from the same location over the same time interval on the same day-of-the-week for multiple weeks preceding the crisis.

Once we have collected the baseline dataset, we eliminate extreme values using winsorization. We do this by computing the mean and standard deviation of the pre-winsorization distribution, identifying the 2.5th and 97.5th percentiles of a Gaussian with that mean and standard deviation, and setting values outside those bounds to the lower and upper bound values if they are anomalously low or high, respectively.

With extreme values eliminated from the baseline dataset, we next compute the baseline mean (μ_{baseline}) and standard deviation (σ_{baseline}) from the remaining data. Finally, we compare what is observed in the current crisis time interval (c) to these baseline statistics. The first comparison is the percent difference between crisis and baseline:

$$\frac{c - \mu_{\text{baseline}}}{\mu_{\text{baseline}} + \epsilon} \tag{1}$$

where ϵ is a small value, usually 1. The second comparison is the z-score:

$$\frac{c - \mu_{\text{baseline}}}{\max\left[\sigma_{\text{baseline}}, \sigma_{\text{min}}\right]} \tag{2}$$

were $\sigma_{\text{min}} \approx 0.1$ is introduced to handle the case where there is no variance in the baseline distribution. The z-score highlights the areas on the map with the most significant differences between what is being observed during the crisis and what is typically seen during the baseline, so unless we note otherwise, this is the value presented in the example maps used in this paper. When sharing data with humanitarian organizations, we also include the percent difference, the baseline means, and the crisis counts. All of this data is modified for purposes of privacy preservation using the methods described in the section above, including dropping counts from locations where either the crisis or baseline value is less than a threshold.

Facebook Population maps

These maps show statistics about the aggregate number of people observed in a location (tiles or administrative polygons) in 8-hour intervals following a crisis compared to a pre-crisis baseline period. The counts include people with location services enabled on their mobile device. If the same person appeared at multiple locations in a time interval we only count their most frequent location, choosing the latest of their most frequent locations in the event of a tie. All of the described privacy protection methods are used on these maps to ensure that the locations of individuals or small groups cannot be identified.

Figure 1 shows an example of the Facebook Population maps[3] in the aftermath of Cyclone Gaja, which affected part of South India in November of 2018. The time interval is the 8 hours ending at 13:30 IST on November 17. The red and blue portions of the map show tile-level Facebook Population z-scores, which have been clipped to a range of -3 (dark red) to 3 (dark blue) for clearer interpretability. The eastern coastal region of the map, around the city of Nagapattinam, is where the cyclone made landfall and shows significant drops in the number of observed people compared to the baseline. These drops could be attributed to evacuations, loss of power, loss of network connectivity, or some combination of all three. Many parts of the west coast, on the other hand, demonstrate increases in the population when compared to the baseline.

Movement maps

These maps present statistics about aggregate movement between pairs of locations in subsequent time intervals. As with the Facebook Population maps, we use the baseline computation procedure to measure pre-crisis movement patterns, and compute statistics for comparing this with what is being observed in the post-crisis period.

Figure 2 shows selected movement vectors between administrative regions in the area impacted by Cyclone Gaja. Each line on the map represents a pair of locations for which there has been an observed change in the number of people moving during the baseline and the crisis time intervals. This map displays only the locations of movement

[3]Note that the methods described in this paper are under active development and considered a work in progress. Therefore, the figures reflect the state of our mapping algorithms at the time that they were generated, namely during the two weeks after the crisis event. As such, they do not always reflect the latest state of our algorithms at the time of publication.

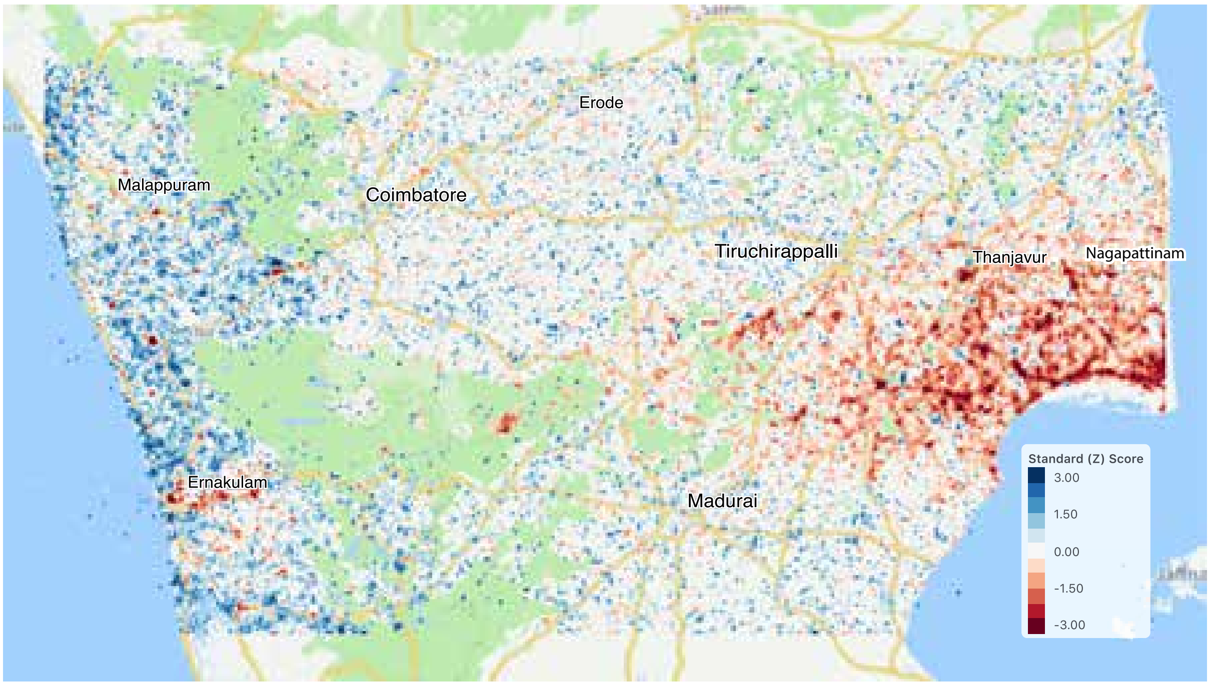

Figure 1. Tile-level Facebook Population map for part of South India on November 17, 2018, in the aftermath of Cyclone Gaja.

vectors. In an interactive visualization, a user of the maps could click on a vector or access the data in tabular form in order to see the actual values. Table 1 shows the percent differences from baseline to crisis for some of the major cities in the map. These differences are based on movement that occurred between eight-hour time intervals ending and beginning at 5:30 IST on November 19, 2018. Because of data sparsity, we only show vectors between larger population centers. While the data in this map does not show a complete picture of what is happening on the ground, it can be useful when combined with other data sources. For example, this map shows a potential increase in travel from Pattukkottai to Thanjavur, and, likewise, from Thanjavur to Tiruchirappalli.

Power Availability maps

For some people with Android devices, it is possible to observe when they connect to a power source. While that power could come from places like cars, generators, or external batteries, when taken in aggregate, it can provide a proxy for the state of the power grid in an area. For each location, these maps count power connection events for devices with location services enabled. We use the baseline computation procedure outlined above to process these counts and produce statistics that highlight where there appear to be significant changes after a crisis when compared to the baseline period. We also use all the privacy protection mechanisms described above in these maps.

Figure 3 shows two more maps related to Cyclone Gaja. The left side of the figure shows the baseline means for connections to power from the 5 weeks preceding the crisis (darker blue indicates more connections), while the right side shows the post-crisis z-scores for the 24-hour period ending November 18, 2018 at 5:30 IST (red indicates negative z-scores). By examining these maps side-by-side, we can see which regions typically show power connections and which of those seem to be experiencing significant drops during the crisis period. These include some of the larger, heavily-impacted cities, such as Nagapattinam, Thanjavur, and Pudukkottai. On the other hand, Tiruchirappalli and Kumbakonam do not appear to have experienced large-scale decreases in connections over this time period.

Network Coverage maps

Cell sites are points of connection in a cellular network made up of antennas with their basestations. A cell tower usually holds multiple such antennas, each of which has a unique identifier. For Android devices with location services enabled, we can use these cell site identifiers to infer network coverage. For each cell site, we draw an estimated coverage area describing the locations of the devices that are accessing that site to obtain cellular connectivity. These inferred coverage areas may overlap, and we count the number that overlap with any tile at any given point in time. These maps are only created for Bing tile locations, not administrative polygons. This

Figure 2. Movement map for Cyclone Gaja on November 19, 2018, highlighting where there were notable differences in movement between pairs of locations when compared to the baseline period.

Starting Location	*Ending Location*	*Percent Difference*
Kumbakonam	Thanjavur	19%
Pattukkottai	Thanjavur	69%
Pudukkottai	Tiruchirappalli	-19%
Thanjavur	Tiruchirappalli	21%
Thanjavur	Kumbakonam	-20%
Tiruchirappalli	Thanjavur	-4%
Tiruchirappalli	Pudukkottai	56%

Table 1. Selected movement data from Cyclone Gaja on November 19, 2018

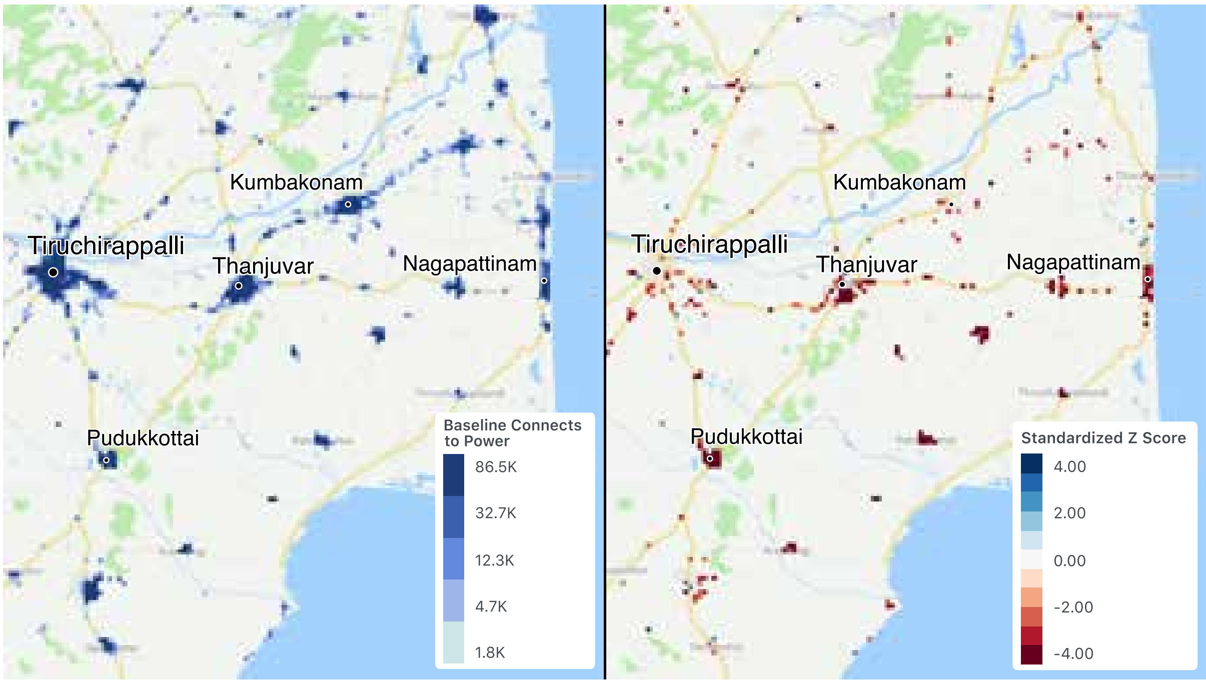

Figure 3. Tile-level Power Availability maps for Cyclone Gaja on November 17, 2018. The map on the left shows the mean power connections from the baseline time period, while the map on the right shows the crisis-time z-scores.

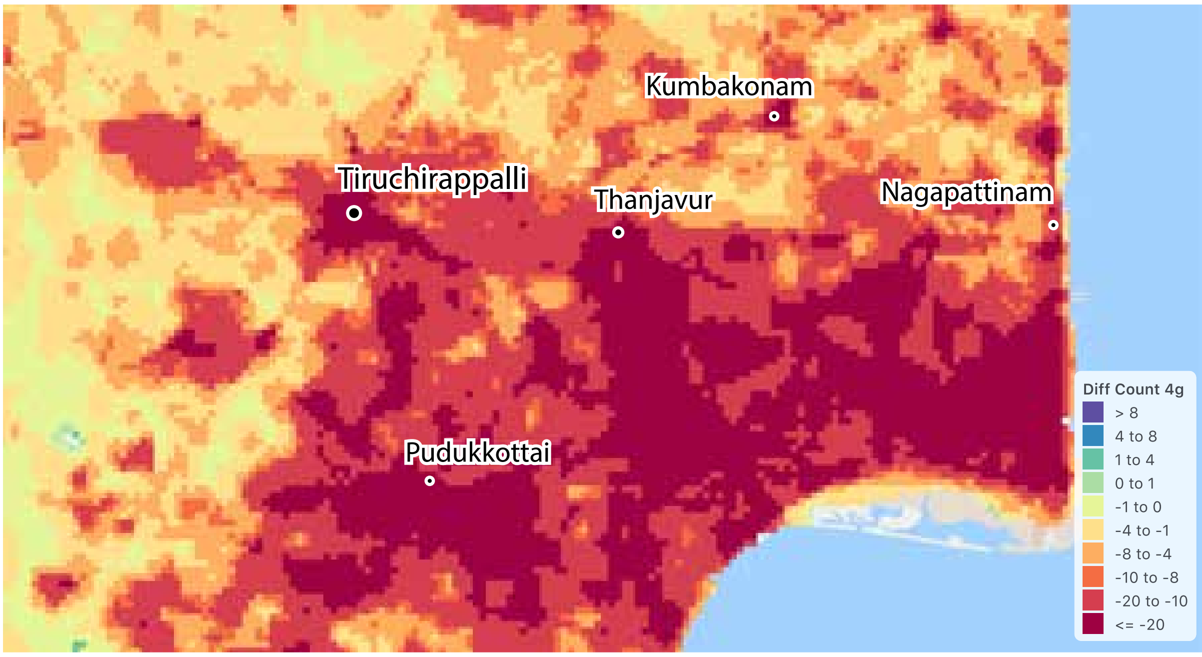

Figure 4. Tile-level Network Coverage map for Cyclone Gaja on November 17, 2018.

count is then compared before and during the crisis using the baseline computation procedure outlined above. We safeguard user privacy by aggregating this data across users and cell sites for each location. Additionally, we drop small counts in the same manner as for the other map types.

Figure 4 shows the a Network Coverage map for a detailed region impacted by Cyclone Gaja based on the 24 hour period ending November 18, 2018 at 5:30 IST. The colors on the map indicate decreases in the number of 4G network connections from a given tile in that time period, compared to the mean connections observed in a three-month baseline period. The patterns observed here echo those in Figure 1, where we saw many fewer people located in the coastal region around the city of Nagapattinam. This highlights an issue where these maps are unable to clearly differentiate between evacuations, loss of power, and loss of connectivity. When considered together, and in combination with other data sources, these maps will provide better, but not complete, clarity about the state of cellular network coverage during a crisis.

Displacement Maps

The final class of Facebook Disaster Maps are long-term displacement maps. The fundamental question that these maps address is the following: in the weeks and months following a hazard event, where has the affected population gone? These maps are based on locations, typically at the city-level, inferred from a person's internet connection.

The calculation procedure for the Displacement maps is as follows:

1. **Home city estimation**
   (a) For everyone who used Facebook from within the bounding box 1-to-5 weeks before the crisis, we identify the most common city from which they accessed Facebook. This city is treated as the person's "provisional" home city.
   (b) For each person identified in step a, we confirm that their most common city in the week immediately preceding the crisis was the same city as in step a. This is meant to filter out people that live in a city but were not present when the crises event occurred. The people that remain make up the effective population for a given home city H.
   (c) We filter to only include the population whose home city, H, is in the bounding box.
2. **Destination city estimation**: For each week following the crisis, calculate the most common city for the population identified in step 1. This destination city, D, does not need to be within the bounding box that defines the affected region.

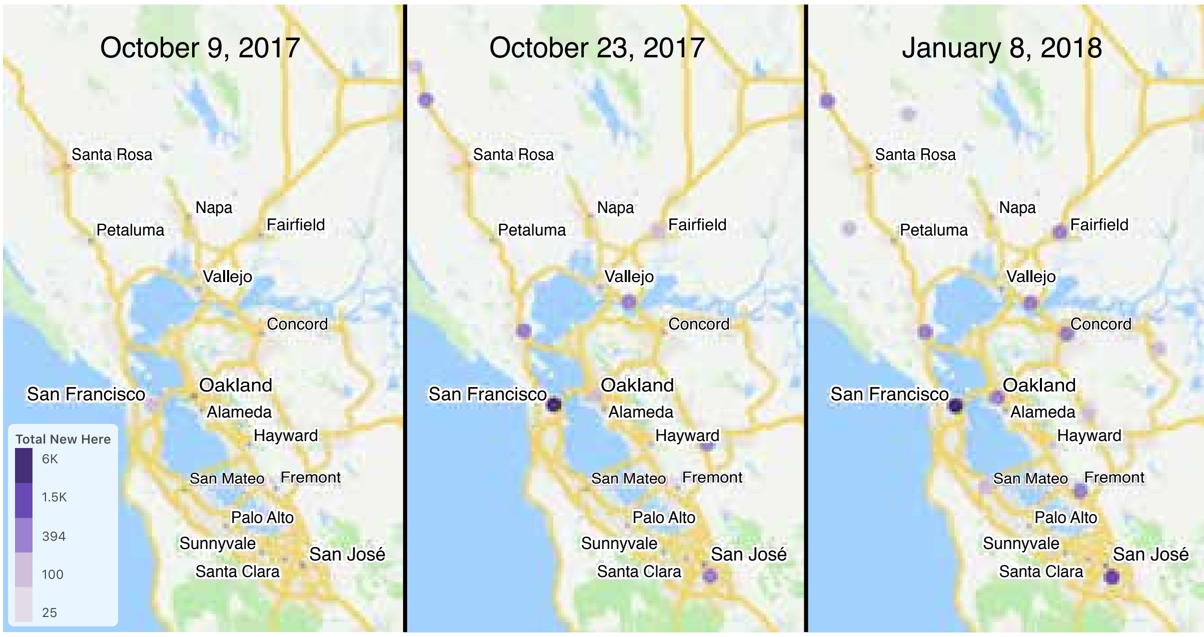

Figure 5. Displacement maps for the Tubbs Fire in the vicinity of Santa Rosa, California at three different points in time. The fire started on October 8, 2017. The maps place purple dots on cities outside of the immediately affected region that have been the destination for at least 100 people who lived within that region.

3. **Home-city-to-destination-city transition estimation**: Aggregate home city-destination city pairs and measure transition counts: of those people from home city H, how many are now in destination city D?

Figure 5 shows examples of the output of these calculations for the Tubbs Fire, which affected Sonoma County, California, in the vicinity of Santa Rosa, in October 2017. The maps place dots on cities outside of the immediately affected region where at least 100 people from the affected region migrated. Darker purple dots indicate the presence of more possibly-displaced people. The fire started on the evening of Sunday, October 8, and we see that on the following day, San Francisco is the only city with a dot. Two weeks later, we see that sufficiently large populations have moved to several cities around the Bay Area. The figure shows only Northern California, but by October 23, more than 100 people had gone to Los Angeles as well. By January 2018, additional Northern California cities appear on the map, and Los Angeles and San Diego would also appear if we zoomed out to include Southern California.

It is not surprising that some people who were in Sonoma County in September 2017 (the time period used to establish the home city estimates) would be in California's major cities by the end of the year. This points to a current limitation of this methodology: Although these maps would ideally measure the population that was forced to migrate because of the crisis, their current state is largely descriptive and does not address that causal question. This may not be a problem in the immediate aftermath of a crisis if we assume that the hazard event is the main driver of mobility in that period. However, as time progresses, the counts of transitions between cities become confounded by several other factors, such as seasonal effects, normal population flows, and fluctuations in rates of Facebook usage by the fixed population under consideration. Unlike our other maps, the displacement maps do not currently include any comparison of the observed counts to "baseline" values that encapsulate what we ought to have expected in the absence of a crisis event. Performing such a comparison could help isolate the causal effect of the crisis itself, and we describe ongoing work in this direction in the next section.

LIMITATIONS AND FUTURE WORK

The maps described above have the potential to provide a picture of crisis situations with a novel level of detail, timeliness, and global coverage. That being said, this picture is imperfect and has a number of notable limitations. In this section we discuss these limitations and directions for future research to address them.

First, our data sources are not representative of the population affected by crises, especially in regions where Facebook penetration is low. Moreover, for most of our maps, there is the additional restriction that we rely on Facebook users who have location services enabled. Drawing inferences based on their behavior may lead to a skewed picture of what is happening on the ground. We attempt to mitigate the risk this poses by ensuring that partner organizations that utilize the maps understand these limitations and use them in combination with other sources of data in a process of triangulation.

In a similar vein, because our data sources depend on signals received by Facebook, different maps may conflate trends in people's movement, connectivity, and power availability. To give a concrete example, if the Facebook Population map shows a decrease in the number of people in a crisis region and the Network Coverage map shows a reduced number of active cell sites, we cannot be sure whether some people have evacuated, internet access has been disrupted, or both.

Recognizing these limitations, we are engaged in ongoing efforts to make the maps more useful and representative of the on-the-ground situations during crises. We are studying how Facebook users with location services enabled differ from the general population in different parts of the world. We seek to characterize this bias, understand how it affects insights drawn from the maps, and, ultimately, correct for it.

Apart from the representativeness issue, our maps may not always isolate the consequences of the hazard event from all other potentially confounding effects. This is particularly problematic for the Displacement maps, because they attempt to measure long-term effects. Many things can happen in the intervening time between the hazard event and map generation: there can be confounding effects from holidays, school schedules, other crises, etc. We are currently working on developing a baseline methodology for this map as well that could help isolate the causal effect of the hazard event.

REPORTS OF USAGE FROM THE FIELD

Since the launch of Disaster Maps in June of 2017 (Jackman 2017), our partners have grown to include more than 30 of the world's most significant nonprofits and UN agencies in disaster response, including the International Federation of the Red Cross, the World Food Programme, the United Nations Children's Fund (UNICEF), NetHope, Direct Relief, and others. Facebook has recruited partners for this effort through multiple channels, including a training partnership with the NetHope consortium, comprised of nearly 60 international NGOs with an interest in the use of technology for humanitarian response.

In the nearly two years since Disaster Maps launched, the maps have been used during major disasters in nearly every region of the world. When Hurricane Maria struck the island of Puerto Rico, NetHope and the American Red Cross used the maps to inform their deployment of nearly 100 Wi-Fi hotspots across the island to the areas that needed them most (Brinkhurst and Crowley 2018). During Hurricanes Florence and Michael, Humanity Road used the Facebook Population maps to monitor large-scale evacuations and determine where communities were still sheltering in place, sharing these insights with FEMA, the US Coast Guard, and state-response agencies (Waggoner 2018). During the Thomas, Carr, Mendocino Complex, and Camp Fires, Direct Relief used the maps to monitor how populations affected by wildfires were moving and then used this to guide distribution of hundreds of thousands of respiratory masks, as well as to coordinate with networks of health centers to alert hospitals that might see increased volumes (Snibbe 2018).

In August of 2018, Kerala, India experienced severe flooding that displaced over a million people. SEEDS India used the Facebook Population and Displacement maps to appropriately time their early recovery phase. After identifying when people appeared to begin to return home, SEEDS launched a nationwide collection drive for widely needed items (Gupta 2018).

When volcanic eruptions in Guatemala and a combined earthquake and tsunami struck the island of Sulawesi in Indonesia, UNICEF used the Network Coverage maps to determine how many people could be reached through U-Report, an outreach tool built on Facebook Messenger. In Guatemala, UNICEF reached 3,000 new people with information about what to do after the eruption and in Indonesia, U-report reached 3,500 people within 48 hours to understand how they were impacted (Brecha Cero 2018). As part of this deployment, UNICEF was better able to deliver on needs related to water, child protection, health and nutrition in collaboration with the Government of Indonesia and United Nations Office for the Coordination of Humanitarian Affairs.

CONCLUSION

We have described the data and methodology behind Facebook Disaster Maps, a collection of geospatial datasets meant to contribute to situational awareness during disasters. The maps are based on aggregated and anonymized data that is collected and employed in the course of Facebook usage, including current and historical location data, information about cell site connectivity, and data on phone battery charging. The privacy and safety of Facebook users is a primary concern, and a number of methods are employed to maintain privacy, including filtering and obscuring cases where data is sparse.

We describe five types of maps. The Facebook Population maps show regions where more or fewer Facebook users are present when compared to pre-crisis periods. This can highlight areas where people are evacuating from or to. The Movement maps show how movement between pairs of locations differs during and before a crisis. The Power Availability maps show where the crisis has been accompanied by a reduction in the number of people charging their phones, which can indicate disruptions in the power grid. The connectivity maps show regions with reduced connections from phones to particular cell sites, which can indicate connectivity loss for those sites. Finally, the Displacement maps show, in the weeks and months following a crisis, indications of the magnitude of the population that has been displaced from their home city and what cities that population has been displaced to.

In the section "Limitations and Future Work," we discussed a number of limitations of the maps that prevent them from being more useful to humanitarian responders. Our ongoing work aims to surmount several of these limitations, for example the problem of isolating the causal effects of the crisis on long-term displacement and the issue of representativeness across all five map types. One overarching goal of all of our research efforts is to better isolate the most actionable information in each map from the less relevant details. That goal will, however, never be perfectly attainable, given the uniqueness of each crisis event. Therefore, our partner organizations will always need to supplement our maps with other sources of data and with their own domain expertise to effectively formulate humanitarian responses.

A rigorous evaluation of the utility of Disaster Maps for humanitarian response is difficult, but should be attempted in future work. In lieu of such an evaluation, in the section "Reports of Usage from the Field," we have collected anecdotal reports of how the maps have been used in practice. While this does not substitute for more formal evaluation, it does provide evidence that the novelty of the data source, the spatial resolution, and the up-to-date nature of these datasets have provided concrete value for the organizations that have used them.

ACKNOWLEDGMENTS

We are indebted to the many people at Facebook who have developed and maintained the systems and infrastructure upon which Disaster Maps are built. We would also like to thank several specific colleagues on the Disaster Maps project, who have contributed to its creation and development and who have also provided helpful feedback on this paper. These collaborators include Molly Jackman, Chang Lang, Eugenia Giraudy, Sourav Chatterjee, Peter Tu, Lior Berry, Alex Pompe, and Sarah Vieweg. Finally, we'd like to thank each of the humanitarian organizations that have worked with us to make Disaster Maps a success.

REFERENCES

Brecha Cero (2018). *Por qué la información en las redes sociales marca una diferencia en los desastres naturales*. Brecha Cero. URL: http://brechacero.com/por-que-la-informacion-en-las-redes-sociales-marca-una-diferencia-en-los-desastres-naturales/.

Brinkhurst, M. and Crowley, J. (2018). *Unlocking Insights from Data: Collaboration with Private Sector creates Cutting-Edge Maps for Disaster Response*. NetHope Blog. URL: https://nethope.org/2018/09/10/unlocking-insights-from-data-collaboration-with-private-sector-creates-cutting-edge-maps-for-disaster-response/.

Cameron, M. A., Power, R., Robinson, B., and Yin, J. (2012). "Emergency Situation Awareness from Twitter for Crisis Management". In: *Proceedings of the 21st International Conference on World Wide Web*. WWW '12. ACM.

Castillo, C. (2016). *Big Crisis Data: Social Media in Disasters and Time-Critical Situations*. 1st. New York, NY, USA: Cambridge University Press.

Ciravegna, F., Gao, J., Ingram, C., Ireson, N., Lanfranchi, V., and Simanjuntak, H. (2018). "Mapping Mobility to Support Crisis Management". In: *15th International Conference on Information Systems for Crisis Response and Management*. ISCRAM '18.

Facebook (2019). *Facebook Privacy Basics: Location*. URL: https://www.facebook.com/about/basics/manage-your-privacy/location (visited on 03/21/2019).

Gupta, M. (2018). *Shedding light on displacement trends in disasters through technology?* LinkedIn. URL: https://www.linkedin.com/pulse/shedding-light-displacement-trends-disasters-through-technology-manu/.

Hughes, A. L. (2014). "Participatory Design for the Social Media Needs of Emergency Public Information Officers". In: *11th International Conference on Information Systems for Crisis Response and Management*. ISCRAM '14.

Hughes, A. L. and Shah, R. (2016). "Designing an Application for Social Media Needs in Emergency Public Information Work". In: *Proceedings of the 19th International Conference on Supporting Group Work*. GROUP '16. ACM.

Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., and Meier, P. (2013). "Extracting Information Nuggets from Disaster- Related Messages in Social Media". In: *10th International Conference on Information Systems for Crisis Response and Management*. ISCRAM '13.

Jackman, M. (2017). *Using Data to Help Communities Recover and Rebuild*. Facebook Newsroom. URL: https://newsroom.fb.com/news/2017/06/using-data-to-help-communities-recover-and-rebuild/.

Lu, X., Bengtsson, L., and Holme, P. (2012). "Predictability of population displacement after the 2010 Haiti earthquake". In: *Proceedings of the National Academy of Science* 109.29, pp. 11576–11581.

Meier, P. (2012). *How Crisis Mapping Saved Lives in Haiti*. National Geographic Newsroom. URL: https://blog.nationalgeographic.org/2012/07/02/how-crisis-mapping-saved-lives-in-haiti/.

Okolloh, O. (2009). "Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information". In: *Participatory Learning and Action* 59, pp. 65–70.

Olteanu, A., Vieweg, S., and Castillo, C. (2015). "What to Expect When the Unexpected Happens: Social Media Communications Across Crises". In: *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing*. CSCW '15. ACM.

Palen, L. and Anderson, K. M. (2016). "Crisis informatics—New data for extraordinary times". In: *Science* 353.6296, pp. 224–225. eprint: http://science.sciencemag.org/content/353/6296/224.full.pdf.

Reuter, C., Hughes, A. L., and Kaufhold, M.-A. (2018). "Social Media in Crisis Management: An Evaluation and Analysis of Crisis Informatics Research". In: *International Journal of Human–Computer Interaction* 34.4, pp. 1–15.

Sarter, N. B. and Woods, D. D. (1991). "Situation Awareness: A Critical But Ill-Defined Phenomenon". In: *The International Journal of Aviation Psychology* 1.1, pp. 45–57.

Schwartz, J. (2018). *Bing Maps Tile System*. URL: https://docs.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system (visited on 01/07/2019).

Sheth, A. (2009). "Citizen Sensing, Social Signals, and Enriching Human Experience". In: *IEEE Internet Computing* 13.4, pp. 87–92.

Simini, F., González, M. C., Maritan, A., and Barabási, A.-L. (2012). "A universal model for mobility and migration patterns". In: *Nature* 484, pp. 96–100.

Snibbe, K. (2018). *How technology is helping responders save lives during disasters like the California wildfires*. The Orange County Register. URL: https://www.ocregister.com/2018/08/21/how-technology-is-helping-those-responding-to-disasters-like-the-california-wildfires-save-lives/.

Soden, R. and Palen, L. (2014). "From Crowdsourced Mapping to Community Mapping: The Post-earthquake Work of OpenStreetMap Haiti". In: *Proceedings of the 11th International Conference on the Design of Cooperative Systems*. COOP '14. Springer International Publishing.

Starbird, K. and Palen, L. (2011). ""Voluntweeters": Self-organizing by Digital Volunteers in Times of Crisis". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. ACM.

Vieweg, S. (2012). "Situational Awareness in Mass Emergency: A Behavioral and Linguistic Analysis of Microblogged Communications". PhD thesis. University of Colorado at Boulder.

Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. (2010). "Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. ACM.

Waggoner, M. (2018). *Facebook Tool Helps Charities Find Evacuees Who Need Help*. U.S. News & World Report. URL: https://www.usnews.com/news/healthiest-communities/articles/2018-09-18/facebook-tool-helps-charities-find-evacuees-who-need-help.

Wang, Q. and Taylor, J. E. (2016). "Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster". In: *PLOS ONE* 11.1, pp. 1–14.

Yan, X.-Y., Zhao, C., Fan, Y., Di, Z., and Wang, W.-X. (2014). "Universal Predictability of Mobility Patterns in Cities". In: *Journal of the Royal Society Interface* 11.

Zade, H., Shah, K., Rangarajan, V., Kshirsagar, P., Imran, M., and Starbird, K. (2018). "From Situational Awareness to Actionability: Towards Improving the Utility of Social Media Data for Crisis Response". In: *Proceedings of the 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing*. CSCW '18. ACM.