Aggregating Information Based on Geolocated Twitter Data

Brian Mitchell
Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA 56267
mitch927@morris.umn.edu

ABSTRACT

Twitter users have the option to include location in their profile and in tweets. In this paper, I describe three examples: disaster management using the CrisisTracker tool, migration patterns, and a smile index for measuring societal happiness. The latter two examples also use face and smile detection. Although data from Twitter and tools for analyzing that data introduces bias making it less useful at a small scale, these data points can be compared to overall changes in populations for showing more accurate large-scale patterns. Having access to high-volume data is crucial for studying how people react to situations, and these studies offer examples of how Twitter data can be leveraged to supplement other more expensive and less timely data sources.

Keywords

Twitter, geolocation, image processing, bias

1. INTRODUCTION

As Twitter has grown to hundreds of millions of active users, data from the service can provide valuable information in ways that traditional sources, such as surveys, cannot. Its distributed users provide immediate, real-time, and high-value data. The geo and streaming features of the Twitter API and the popularity of using smartphones with GPS units allows for a large-scale source of geolocated messages.

Some locations provide a specific coordinate, while others contain a bounding box around an area. Location can also be extracted from the text of a tweet, or from the location field in a user's profile. Twitter posts can be accurately and efficiently grouped, organizing them into events that can be tracked over time. Object detection and other image processing schemes can be used on attached images to recognize faces and other attributes. This can provide valuable information such as demographic characteristics about a user or their experiences that can add to the text and location data.

Data from Twitter can be used in many ways. In this paper, I describe three examples: disaster management using the CrisisTracker tool in Section 3.1, migration patterns in Section 3.2, and a smile index for measuring societal happiness in Section 3.3. I conclude with some discussion on issues with these approaches for using Twitter as a data source.

UMM CSci Senior Seminar Conference, April 2016 Morris, MN.

2. BACKGROUND

2.1 Twitter

Twitter is a microblogging social network. Posts made on Twitter, often referred to as "tweets," are composed of up to 140 characters. Tweets can include optional embedded information such as location, timestamps, images, polls, and links. Twitter has two APIs for viewing tweets, REST and Streaming. In this paper, I focus on the Streaming API. This API is one that keeps an HTTP connection open for extended periods of time, during which new data is sent to the client when it is available. It is best for following specific users or topics, or for other data mining. [10]

Twitter provides users the option to include a location with a new tweet. When users opt-in to use this feature, a location is stored as a `place_ID`, with the option to include a precise coordinate (a latitude and longitude pair) in addition. A `place_ID` contains a bounding box of coordinates around the area. It can be as specific as a neighborhood, or as broad as a whole state, but is generally a city or neighborhood. A point of interest (POI) can also be a `place_ID`, for example "Golden Gate Park" in San Francisco. Due to the popularity of smartphones that contain GPS units, it is easy to include a coordinate when using a smartphone. Twitter will automatically turn this coordinate into a `place_ID`, and keep it embedded in the tweet. In addition to storing a location in a tweet, a Twitter user has the option to set a location that is displayed on their profile. This location is not normalized or validated, so it can range from a normal looking place such as "Morris, MN" to something completely unrelated to a location like "long suffering mets fan." [10]

The population of users on Twitter does not match the general population. It has been shown that there is an urban bias in volunteered geographic information (VGI) in services such as Twitter, Flickr, and Foursquare [3]. It has also been shown that in VGI from Twitter, Flickr, and Swarm, only about 75% of it is "local", meaning that only 75% of the locations closely corresponds with home locations of users [6]. Correlations have been made that suggest that well-educated people in occupations of management, business, science, and arts are more likely to include location in their tweets and photos [7]. These are all important factors to consider when using VGI and must be accounted for when making estimations about the general population.

2.2 Locality-Sensitive Hashing

Locality-sensitive hashing (LSH) is "a randomized technique that dramatically reduces the time needed to find a nearest neighbor in vector space" [8]. This paper covers an extended version of LSH that Petrović et al. have adapted so that as new documents are added, searching happens in constant time. They found that applying pure LSH and first story detection (FSD), where a story is a collection of clustered documents around an event, was slow and yielded a high variance in results. In this paper, LSH is used to group tweets into stories for use in the disaster management tool CrisisTracker as described in Section 3.1.1. [8]

2.3 Haar-like Features

Haar-like features are pieces of information about a digital image that are relevant for object recognition. The name comes from the similarity of Haar basis or wavelet functions and was used in the first real-time face detector as designed by Viola and Jones [11]. These features are organized into a classifier cascade, which is the product of "combining increasingly more complex classifiers" to quickly discard background regions so more computational power can be dedicated toward detecting the desired object-like regions of an image (such as faces). This system converts images to grayscale and then compares changes in contrast across regions of the image, rather than more traditional approaches such as looking for pixels with colors near that of skin tones. Regular characteristics of a human face, such as the upper-cheek and nose regions being lighter than the eye region, can be matched using Haar-like feature classifiers. These provide a fast, efficient, scalable, and accurate algorithm for detecting objects in an image. This is used for facial detection as mentioned in Section 3.3.2. [11, 13]

3. APPLICATIONS

The following three applications are examples of using Twitter as a data source for data analysis and processing.

3.1 Disaster Management

Having access to real-time social media during times of crisis is crucial for accurate reporting. One such piece of software, CrisisTracker [9], tracks keywords and creates stories based on the similarities of these keywords. A report of the pilot deployment during an eight day period in 2012 focused on the Syrian civil war. It is one of the first tools to combine crowdsourcing with automated analysis. [9]

3.1.1 CrisisTracker

CrisisTracker works by collecting tweets from Twitter's Streaming API. The API returns a collection of tweets that are tagged within any geographical region that at least partially lays in a specified bounding box. Some tweets are filtered out to keep the data relevant and of higher quality. For example, tweets that contain fewer than two words are discarded (such as, "@username thanks!"). [9]

New tweets are compared with previously collected tweets as a weighted set of words. Tweets are fed into a similarity metric that groups similar tweets. Tweets are then run through an extended version of locality-sensitive hashing as described in Section 2.2. This creates a set of stories that can then be viewed in CrisisTracker. Stories contain automatically extracted metadata such as timestamps, keywords, and number of users. Users of CrisisTracker can rank stories by their size (number of users who mention the story). Limiting each story to the top 5,000 users tweeting about the given disaster was found to work for keeping detailed updates while omitting jokes, opinions, and summary articles. Users can also tag a story to a map, merge similar stories that were not automatically merged, remove misclassified content, and hide irrelevant stories (for example, a cooking recipe named after a location). [9]

Ikawa et al. has taken CrisisTracker and added the ability to infer locations from similar messages and classify messages based on the location. In the modification, locations are broken into four categories: locations in text, focused locations, user's current location, and user's location profile. Locations in text are place names or POIs included in the text of a message. These locations can be used to see the geographic distribution of each message. However, not all locations in text places are relevant; for example, a message containing the places "London" and "Syria", where "Syria" is the location of the content and is the relevant location, and "London" is the location of the reporting and is not relevant. In contrast, focused locations are just the relevant locations in a message. In the previous example, "Syria" would be a focused location as it is where the main content is located. A message can contain several relevant locations. Focused locations are used to more appropriately locate information on a map. User's current location refers to the location where a user posted the message. This can be in the form of a geotagged location, or inferred from the text of the current and previous messages. Finally, the user's location profile is the home location of a user. This location can be entered on the user's profile or be found using algorithms for inferencing locations from past messages. [5]

The modification uses GeoNames (a geographical database that contains over eight million names and coordinates) as the database for the system [2]. Location Name Recognition and Toponym Resolution are the two components that make up the system. The former detects locations in the given string. It attempts to filter out proper nouns, such as "Mr. Paris," however some words may still pass through, such as "Obama," which is both the name of the US President and the name of a Japanese city. The outputs of locations are passed to the Toponym Resolution, which assigns coordinates to the location. It also attempts to match a location name that represents multiple location instances with its actual location using population data from GeoNames. [5]

A confidence score is calculated for every possible location instance. A score for Location Popularity is given based on the population of the location, and a Region Context score is given based on the locations that are included in the message. The Region Context score will be higher if the message includes a location instance that is in the referenced country from the message. The confidence score is made by multiplying the Location Popularity and Region Context scores. The location with the highest confidence score will be used for Toponym Resolution. [5]

3.1.2 Results

The evaluation of the modified version of CrisisTracker used a sample of 182 tweets that mentioned the Syrian civil war as collected by the original CrisisTracker. A subset of the GeoNames database was used: cities in Syria with a population over 15,000. To get the best results for ideal conditions, all of the place names were extracted by hand from the messages. [5]

The results of the evaluation of Location Name Recognition and Toponym Resolution can be seen in Table 1. The #appearance row refers to the total number of locations while the #unique row refers to the number of locations after the removal of duplicate elements from the dataset. The

Table 1: Modified CrisisTracker evaluation results [5]

	Country	State	City/ Town	Village	Total
#appearance	250	39	41	12	342
#unique	20	7	11	8	46
Precision	0.996	1.000	1.000	0.917	0.994
Recall	0.992	1.000	0.927	0.750	0.977

precision and recall are used to measure relevance. Precision is how successful the technique is at finding known relevant data, while recall is how completely the technique finds relevant data. The results show that the system performs well for major place names (all locations except for villages on Table 1), and performs fairly well for villages. Ikawa et al. found that many of the locations from tweets were inferred from text. For example, a bombing near a police station matched the entire city was when the actual event was in fact much more localized. The lower recall on villages is due to the GeoNames database not having some of the small villages in its record, often because of mistranslations from Arabic to Latin characters or ambiguity in the database. Adding the ability to infer locations in CrisisTracker allows for a larger amount of useful location aware messages that can make the tool provide more information for managing disasters. [5]

3.2 Migration Trends

Measuring migration flows is difficult due to inconsistent, outdated, and sometimes nonexistent data. Researchers generally rely on census data to estimate the movement of people. This limits data to census years and does not show recent trends. Additionally, since data must be made consistent between sources and countries, it may take a few years for it to become available. The lack of recent data can strongly affect migration projections and can make for larger errors in medium- and long-term projections. Data from online sources, such as geolocated data from Twitter, can supplement traditional data sources and improve the understanding of migration patterns. While migration was not explicitly defined in Zagheni et al.'s paper, I will define it as the temporary or permanent movement of people with a desire to settle. [14]

3.2.1 Data collection and pre-processing

Zagheni et al. used data from Twitter as one data source for both intranational and international migrations. Tweets from 500,000 users who posted between May 2011 and April 2013 were downloaded. Tweets and their respective users were mapped to countries until about 3,000 users were in each country. For the initial seed data, oversampling was done in countries with lower rates of mobility, while in countries with high mobility, relatively smaller samples were collected. For each country, they calculated the fraction of users who had geolocated tweets in at least one country outside of their home country. To get a similar sample size for each country, users were then sampled with a probability inversely proportional to this fraction. For example, if in one country, 50% of users posted tweets from a foreign country, and in a different country 5% of users posted tweets from a foreign country, the latter country would need a sample roughly 10 times larger than the former. [14]

Starting with the initial seed data, and continuing with the sampling procedure, at least one geolocated tweet for 500,000 users in member countries of The Organisation for Economic Co-operation and Development (OECD)[1] was obtained. For users who posted the selected tweets, all public tweets were downloaded. In the sample, about 345,000 users had at least 10 geolocated tweets, and on average, users posted 142 geolocated tweets. The distribution was fairly skewed, with the median number of geolocated tweets being 34. On average, there were 225 days between the first and last geolocated tweet, and 12 days between each tweet. For users who have posted at least 10 geolocated tweets, the average number of days between each tweet went down to 6 days. [14]

It was decided to select a smaller sample of users who have more detailed and consistent information over time versus a larger sample of users with less detailed information. Users that have information over a longer period of time are more likely to provide reliable information, and keep posting tweets in the future. The dataset was split into periods of four months, and only users that had at least 3 geolocated tweets for each period were considered. This reduced the sample size to about 15,000 users. [14]

3.2.2 Demographic characteristics

For each user, a unique identifier, the text of tweets, date of posts, and geographic coordinates of the tweets were available. To estimate the gender and age of a user, the facial recognition software, Face++[2], was used. Out of the sample of users with at least one geolocated tweet, 21,553 users with a profile photo could be evaluated with Face++. These photos do not necessarily mean that they are a photo of the given user, but the data can still be used for general purposes. The sample was skewed towards younger people and not representative of the general population, but was an interesting fit for migration data as migration is typically a transition that happens at a younger age. At an individual level, the estimates for age and gender were uncertain, but regardless of accuracy some useful patterns can be observed at the holistic level. The estimates of users' ages may be biased as some accounts will not have an up-to-date profile photo. [14]

3.2.3 Estimation of trends in out-migration rates with a difference-in-differences approach

For each user who has posted at least three geolocated tweets in each four-month period, a home country was estimated (the modal country). If there was a large uncertainty (the number of tweets from the modal country was not at least three times as high as the next most frequent country), the information for that user during that period was discarded. If the user's modal country changed between two four-month periods, the user was estimated to have migrated from the first to the second country during the eight months. [14]

The migration rates that were estimated could not be considered representative of OECD countries; they represented the inferred experiences of the Twitter users who frequently post geolocated tweets. Zagheni et al. proposed a difference-in-differences approach to estimate recent trends in mobility rates for the general population. Let m_c^t be the out-

[1] http://www.oecd.org/about/membersandpartners/list-oecd-member-countries.htm

[2] http://en.faceplusplus.com

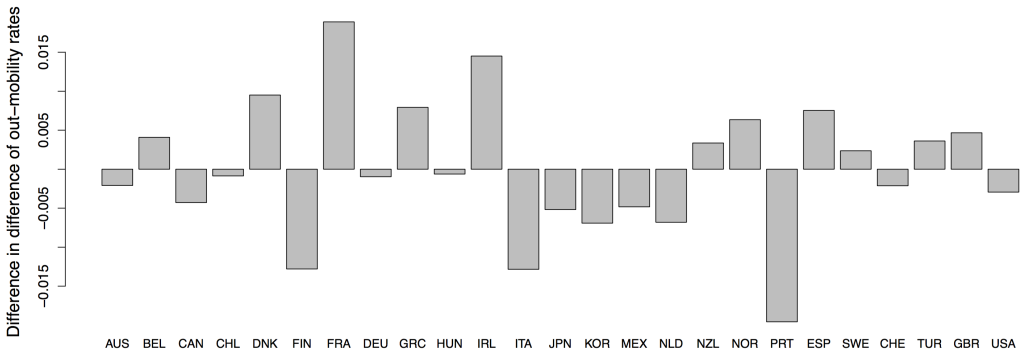

Figure 1: Values of the difference-in-differences estimator $\hat{\delta}$ for out-migration rates for OECD countries that consistently have a sample of at least 100 Twitter users for each four-month period. [14]

migration rate from country c to all other countries (number of users in country c/number of users who are migrants from country c), at time t. The average of this quantity across all countries, m^t_{oecd}, is the average migration rate at time t for all considered OECD countries. The difference-in-differences estimator, Equation 1, allows for change in the Twitter users' population if it is similar to the population change in OECD countries:

$$\hat{\delta} = (m^t_c - m^t_{oecd}) - (m^{t-1}_c - m^{t-1}_{oecd}) \qquad (1)$$

Selection and changes in the Twitter population over time prevent making statistical inferences for a single point in time. However, if changes in the Twitter population match changes in the populations of countries, the comparison of relative changes for a country with relative changes for the Twitter users can be used to gather information on trends. For example, if the proportion of 25-year-old Twitter users is consistent with the proportion of 25 year olds in the general population, and an increase in the out-migration is observed for a given country, then the population-level migration rates increased, relative to other countries. [14]

3.2.4 Results

The average out-migration rates for OECD countries was found for each four-month period. However, the rates cannot be considered accurate as the sample is not representative of the whole population. The country specific out-migration rates from the general trend can be used to indicate changes in mobility patterns. Figure 1 shows the estimates of difference-in-differences $\hat{\delta}$ for out-migration rates for OECD countries. These countries consistently have a sample of at least 100 Twitter users in each four-month period. The results shown are the average $\hat{\delta}$ evaluated for the periods of May-Aug and Sept-Dec 2011 against the estimated rates for the same months in 2012. Positive values show a relative increase of out-migration rates, while negative values show a decrease in rates. [14]

The results show some interesting metrics. Using Mexico as an example, the country saw a decline in out-migration rates from 2011 to 2012 (as seen in Figure 1). Because a majority of Mexican migrants move to the US, the result can be interpreted as a decline in migration from Mexico to the US. This is consistent with estimates from the Pew Research Center for 2005-2010[3]. The Twitter data showed that the decrease in out-migration from Mexico is persisting. Normally, with census or other official statistics, it would take several years for this information to be observed. Unless recent trends are incorporated, the results of these projections might overestimate the migration rate from Mexico to the US. The data also showed that the decreased out-migration from the US, Italy, and Portugal had lowered mobility to other countries. Spain, Greece, and Ireland saw an increase in out-migration rates. [14]

To gauge the differences between internal and international migration and travel, Zagheni et al. estimated the radius of gyration of geolocated tweets for migrants and non-migrants in their home countries. They described the radius of gyration as, "a measure of the average distance of geolocated tweets from their baricenter [*sic*]" where the barycenter is the center of the bounding box of locations from which the user posted their geolocated Tweets [12]. Migrants were users who had moved to a different country for at least one four-month period. It was observed that for larger countries, the distance from the barycenter is larger. For most countries, international migrants traveled shorter distances in their home country than those who did not migrate internationally. A notable exception to this is the US, where the radius of gyration is larger. International migrants originating from the US are likely to be part of a group of people who travel more, both nationally and internationally. This may be related to the recent economic crisis that caused a general reduction of internal mobility in the US. [14]

3.3 Societal Happiness

Scholars and policymakers have, in recent times, been making efforts to measure the well-being of individuals and groups of regional, national, and global levels. Metrics such as Gross National Happiness have appeared and are inspiring governments and organizations worldwide to measure the happiness of people. In 2012, the United Nations General Assembly launched the annual World Happiness Report to rank countries by happiness. To calculate happiness for these reports, considerable amounts of time, money, and manpower are required. Because of this, reports use limited samples, or are administered annually or less frequently. Abdullah et al. released a report on using embedded photos in geolocated tweets as a data source for well-being reports. [1]

[3] `http://www.pewhispanic.org/files/2012/04/Mexican-migrants-report_final.pdf`

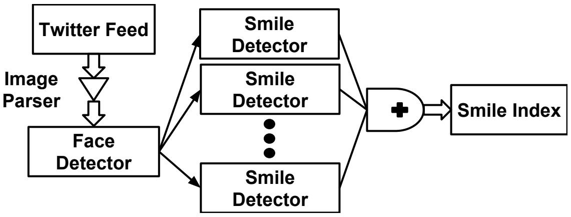

Figure 2: Framework for assessing happiness though the Smile Index of Abdullah et al. [1]

3.3.1 Data

A dataset of nine million geotagged tweets posted between January 1, 2012 and April 30, 2013 was used. This data was obtained via Twitter's "garden hose," an official sample of 10% of all tweets. Timezones from all tweets were normalized to Coordinated Universal Time (UTC). Prior work has shown that, generally, information is shared as it happens, in real time, and that tweets with location are posted from the user's current location. All of the images downloaded had been uploaded to Twitter's official photo-sharing service. After conducting a random sample, it was found that 72% of tweets with a photo used Twitter's official photo-sharing service. Metadata is removed from images as they are uploaded to this service, which means that uploaded photos do not have any location data embedded within them. The location of a tweet does not necessarily mean that the attached photo was also taken at that location. This introduces a discrepancy that is likely caused by photos posted by locals versus photos posted by tourists. Finally, tweets without photos were removed from the dataset. [1]

3.3.2 Smile Index Framework

Figure 2 shows the phases of the happiness assessment. Images were converted to a standard JPEG format with the colors converted to grayscale. Faces were detected using a cascade of boosted classifiers with Haar-like features as described previously in Section 2.3. Through this system, an efficient removal of false positives was achieved, and ultimately attained a 100% accuracy rate for detecting faces in the test dataset that was used. After this phase, only images that contain faces are used. [1]

After detecting faces in the images, smiles are found using another Haar feature based classifier trained by Hormada et al. [4], who manually labeled images from several sources. The classifier has a strong prediction rate. If the image contains more than one face, each face is run through the smile detection algorithm followed by performing a logical OR on the detected faces. This means that if any detected face in an image is smiling, the image will be considered as a positive instance of smiling. The system was tested with the GENKI[4] database of diverse images of people (gender, ethnicity, age, geographical location, facial features (e.g., glasses, facial hair), and photo setting (e.g., indoor, outdoor)) from public personal web pages. The system achieves an accuracy of 0.85 on the GENKI-4K database, a subset of 4,000 images manually labeled for smile presence. [1]

Due to the popularized nature of rapidly growing social media networks such as Twitter, most captured faces tend to be smiling. As a result, the raw output of smile-detection needs to be adjusted to more accurately quantify happiness. The number of smile containing images is compared to the number of images posted at a given location during a given time period, t:

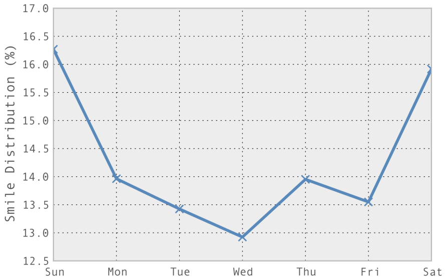

Figure 3: Distribution of smiles in U.S. tweets over days of the week [1]

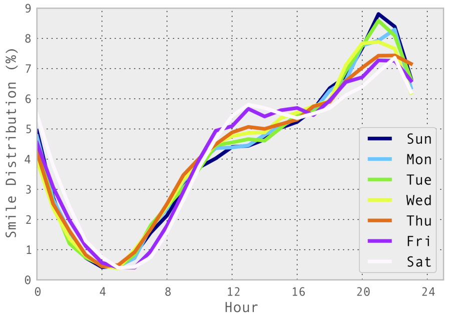

Figure 4: Daily breakdown of hourly smile distribution [1]

$$R_t = \frac{S_t}{I_t} \qquad (2)$$

This ratio, R_t, is S_t, the raw smile count from time period t from the given location, over I_t, the total number of images from time period t from the given location. Because faces are more likely to be smiling, the number of images without faces can be used to gauge happiness. For example, during Hurricane Sandy, a rise in the number of images without faces was observed. This lower ratio of smiles led to the detection of negative sentiment to the event. [1]

3.3.3 Results

In line with prior research, a significantly higher proportion of smile-containing images were posted during the weekend (Figure 3) [1]. Within a single day, happiness rises gradually in the morning, levels off and dips in the mid-afternoon, before increasing in the evening (Figure 4). The happiness drops steadily until the middle of the night, when depressed individuals are more active on Twitter. The hourly happiness trend found is similar to others with the exception of the evening, when it is much higher than other studies. Four case studies of emotionally significant U.S. events verified that the Smile Index increases in response for holidays and celebratory events, while it decreases during tragedies and disasters. Through connecting emotional health with demographic and emotional indicators, there are positive correlations between income and predominately white-majority geographical areas and the reverse in black-majority areas. While testing the prediction capabilities of the Smile Index, it was demonstrated that the metric can forecast happiness 7 days into the future with just 13 days of historical data. [1]

[4] MPLab GENKI database `http://mplab.ucsd.edu/wordpress/?page_id=398`

4. DISCUSSION AND CONCLUSIONS

The modified CrisisTracker by Ikawa et al. clustered tweets into stories for managing disasters. It was found that the certainty of a given event and its associated location decreases as the volume of tweets increases. Four approaches of gathering location from Twitter and a location inferencing method were proposed. The prototype performed well for major places, and reasonably well for smaller places. It is thought that adding additional data sources would improve performance of the system. [5]

The research done by Zagheni et al. provided a way to collect recent data from Twitter that allows for more accurate migration trend projections. The observations made from the data were for short-term trends and the results should be used with care, as there may be some stochasticity in the results. The data can be used for estimating recent trends before more official data is available. The data could be made more accurate if it were trained with official data from the same time, however that data was unavailable. The difference-in-differences approach was created to reduce bias for statistical inference. Zagheni et al. asks the Web Science community "how can we make statistical inference from online data when there is not any 'ground truth' data that can be used as a training reference?" [14]

The Smile Index as developed by Abdullah et al. uses the amount of smile detected images versus the total number of images posted to determine a real time, unobtrusive, and continuous look at the moods of a population. By using images for the index versus text, it allows for this method to easily scale around the world across many languages. Variances in the culture around photography, such as East Asians' facial expressions tending to show lower intensity in photographs, do not cause a significant change in the effectiveness of the Smile Index as the changes in ratios of smiling and non-smiling images is the same across cultures. Additional data sources would help reduce population bias, and further investigation of using images for sentiment analysis would help to better understand how smiles and images reflect overall happiness. [1]

In general, using Twitter as a data source allows for high volume, immediate data. This can be a significant improvement over traditional data sources such as surveys which often have much lower response rates, are administered infrequently, difficult to scale, and expensive. Using Twitter as a data source raises several questions with bias and accuracy. The population that uses Twitter does not accurately represent the general population. In the United States, around 15% of online adults use Twitter, which overrepresents minorities and younger people when compared to the overall population as Twitter has a higher proportion of these users [1]. In the studies conducted by Zagheni et al. and Abdullah et al., facial recognition was used. Profile photos were used for demographic information such as age and gender. This is likely to be quite inaccurate at an individual basis, and likely to be more accurate at a holistic level, but profile photos can be very outdated, of a different person, or have several people in the photo. Sample bias is an inherent part of using data from Twitter, but large-scale patterns and trends can still be useful for predictions and estimates before official or survey based statistics are available.

Acknowledgments

Thanks to Kristin Lamberty, Elena Machkasova, and Jacob Thebault-Spieker for their guidance and feedback.

5. REFERENCES

[1] S. Abdullah, E. L. Murnane, J. M. Costa, and T. Choudhury. Collective smile: Measuring societal happiness from geolocated images. In *Proceedings of CSCW '15*, pages 361–374, New York, NY, USA, 2015. ACM.

[2] GeoNames. Geonames. Online, `http://www.geonames.org`. Accessed: 2016-03-01.

[3] B. Hecht and M. Stephens. A tale of cities: Urban biases in volunteered geographic information. In *Proceedings of the 8th International Conference on Weblogs and Social Media*, pages 197–205, 2014.

[4] D. D. Hromada, C. Tijus, S. Poitrenaud, and J. Nadel. Zygomatic smile detection: The semi-supervised haar training of a fast and frugal system: A gift to opencv community. In *International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2010 IEEE*, pages 1–5, Nov 2010.

[5] Y. Ikawa, M. Vukovic, J. Rogstadius, and A. Murakami. Location-based insights from the social web. In *Proceedings of WWW '13*, pages 1013–1016, New York, NY, USA, 2013. ACM.

[6] I. Johnson, S. Sengupta, J. Schoening, and B. Hecht. The geography and importance of localness in geotagged social media. In *International Conference on Human Factors in Computing Systems (CHI 2016)*, California, USA, 2016. ACM.

[7] L. Li, M. F. Goodchild, and B. Xu. Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*, 40(2):61–77, 2013.

[8] S. Petrović, M. Osborne, and V. Lavrenko. Streaming first story detection with application to twitter. In *HLT '10*, pages 181–189, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[9] J. Rogstadius, M. Vukovic, C. A. Teixeira, V. Kostakos, E. Karapanos, and J. A. Laredo. Crisistracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research and Development*, 57(5):4:1–4:13, Sept 2013.

[10] Twitter. Twitter developers. Online, `https://dev.twitter.com`. Accessed: 2016-03-19.

[11] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.*, volume 1, pages I–511–I–518 vol.1, 2001.

[12] Wikipedia. Barycenter — Wikipedia, The Free Encyclopedia. Online, `https://en.wikipedia.org/wiki/Barycenter`. Accessed: 2016-03-17.

[13] Wikipedia. Haar-like features — Wikipedia, The Free Encyclopedia. Online, `https://en.wikipedia.org/wiki/Haar-like_features`. Accessed: 2016-04-04.

[14] E. Zagheni, V. R. K. Garimella, I. Weber, and B. State. Inferring international and internal migration patterns from twitter data. In *Proceedings of WWW '14*, pages 439–444, New York, NY, USA, 2014. ACM.