

Identifying Eyewitness News-Worthy Events on Twitter

Erika Varis Doggett

Timeline Labs

Santa Monica, CA

erika.varis@timelinelabs.com*

Alejandro Cantarero

Timeline Labs

Santa Monica, CA

alejandro@timelinelabs.com†

Abstract

In this paper we present a filter for identifying posts from eyewitnesses to various event types on Twitter, including shootings, police activity, and protests. The filter combines sociolinguistic markers and targeted language content with straightforward keywords and regular expressions to yield good accuracy in the returned tweets. Once a set of eyewitness posts in a given semantic context has been produced by the filter, eyewitness events can subsequently be identified by enriching the data with additional geolocation information and then applying a spatio-temporal clustering. By applying these steps we can extract a complete picture of the event as it occurs in real-time, sourced entirely from social media.

1 Introduction

Current information has always been of paramount interest to a variety of professionals, notably reporters and journalists, but also to crime and disaster response teams (Diakopoulos et al., 2012; Vieweg et al., 2010). With the explosion of the internet and social media, more information is available—and at a faster rate—on current events than ever before. A large proportion of non-professional people are now in the position of news reporters, spreading information through their social networks in many cases faster than traditional news media (Beaumont, 2008; Beaumont, 2009; Ritholtz, 2013; Thielman,

* Current affiliation: The Walt Disney Studios
Erika.Varis.Doggett@disney.com

† Current affiliation: tronc, Inc.
acantarero@tronc.com

2013; Petrović et al., 2013), whether by sharing and re-sharing the first news message of an event, or through immediate personal eyewitness accounts. This average eyewitness represents a valuable and untapped information source for news professionals and others. But given the wide variability of texts and topics included in social media, the question remains: how to sift the wheat from the chaff?

When your information source is an average user, distinction of eyewitness posts from noise and off-topic data is difficult. Everyday social media users may be situated on the scene, but they do not always frame their posts in the most informative or consistent language. Their intended audience is a personal network of friends for whom a lot of contextual information is already available, allowing their utterances to be highly informative for that network, but still ambiguous for strangers, or automated programs.

1.1 Related work

Previous studies have attempted to programmatically identify eyewitnesses with limited success. Imran (2013) achieved only a .57 precision accuracy for their machine learning eyewitness classifier. Diakopoulos et al. (2012) reported a high accuracy for a static eyewitness classifier at .89, but it is unclear exactly how it was constructed, losing replicability and verifiability. In addition, their classifier only analyzed static datasets, whereas the speed of current events reporting on social media calls for a tool for online use.

In this paper we present a linguistic method for identifying eyewitness social media messages from

a microblogging service such as Twitter, in a real-time streaming environment. Our system identifies messages on different eyewitness topics, including shootings, police activity, and protests, and can easily be extended to further areas, such as celebrity sightings and weather disasters. We further identify events corresponding to groups of related messages by enriching the data with geographical location and then running a spatio-temporal clustering algorithm.

Our work provides the novel contributions of a system that functions in a real-time streaming environment, combines information such as semantic and spatio-temporal clustering, and utilizes simple and fast computational tools over classifiers that require large training data and long setup time. In §2 we outline our process for finding eyewitness posts and events. Section 3 presents results from this system, and we provide some concluding remarks in §4.

2 Method

An eyewitness post on a social network is a text document giving a first person account from a witness to the event. As such, we looked to build filtering rules based on language related to an event, excluding posts from official agencies (e.g. police, fire departments), news outlets, and after-the-fact or remote commentary. In this section, we describe how filters can be constructed that are capable of doing this in real-time on Twitter.

2.1 Datasets

We collected Twitter data from several real events to find a set of eyewitness tweets to inform the creation of linguistic filters. Such a dataset can be collected from the Twitter API (or any suitable 3rd party vendor) by doing broad searches in a narrow time window right after an event has happened. To build a rule set for shootings, for example, we pulled data from multiple mass shootings, including shootings in 2013-2014 at LAX in Los Angeles; Isla Vista, CA; and the Las Vegas Wal-mart and Cici’s Pizza. In these cases, searches around the local place-names at the time of the shootings produced a very broad set of documents, which were then manually checked for true eyewitness texts, resulting in an informative set of eyewitness tweets. By examining these eyewitness tweets, we discovered several

consistent language patterns particular to eyewitness language.

2.2 Language patterns

One of the challenges of social media language is that users exhibit a wide range of phrasing to indicate an event has occurred. It is because of our world knowledge that we are able, as human language speakers, to understand that the person is discussing a newsworthy event (Doyle and Frank, 2015).

With that in mind, we propose building filters that consist of three parts. The first is a semantic context. Examples here might be criminal shootings, unusual police activity, or civil unrest. This semantic context may be built using heuristic rules, or it may also be derived from a machine learning approach. The second part is the existence of salient linguistic features that indicate an eyewitness speaker. Finally, we look at similar linguistic features that indicate the user is not an eyewitness, useful for blocking non-eyewitness posts.

2.2.1 Eyewitness features

First person. First person pronouns are often dropped on social media, but when present this is a strong indicator that the event being described was witnessed first-hand.

Immediate temporal markers. Words such as “just”, “now”, “rn”¹ indicate the event happened immediately prior to the tweet or is ongoing.

Locative markers. Language may be used to define a place in relation to the speaker, such as “home”, “work”, “school”, or “here”.

Exclamative or emotive punctuation. Eyewitnesses to an exciting or emotional event express their level of excitement in their messages. Common ways this may be achieved are through punctuation (exclamation and question marks), emoticons, emoji, or typing in all capital letters. These are relatively common features used in social media NLP (Thelwall et al., 2010; Agarwal et al., 2011; Neviarouskaya et al., 2007).

Lexical exclamations and expletives. A normal person is likely to use colorful language when witnessing an event. Phrases such as “smh”², “wtf”, and expletives are often part of their posts.

¹Twitter short-hand for “right now”.

²“Shake my head”

2.2.2 Non-eyewitness features

Non-eyewitness features are crucial as a post may match the semantic context and have at least one of the linguistic eyewitness markers above, but still not be an eyewitness account of an event. The main markers we found for non-eyewitness language fall into a handful of categories, described below.

Jokes, memes, and incongruous emotion or sentiment. The expected reaction to a violent crime or disaster may include shock, sadness, anxiety, confusion, and fear, among others (Shalev, 2002; Armsworth and Holaday, 1993; North et al., 1994; Norris, 2007)³. As such, it is reasonable to remove posts with positive sentiment and emotion from eyewitness documents related to a traumatic incident (e.g. shootings).

Wrong part of speech, mood, or tense. The verb “shoot” in the first person is unlikely to be used in a criminal shooting context on a social network. The conditional mood, for example in phrases such as “what if, would’ve, wouldn’t”, indicates a hypothetical situation rather than a real event. Similarly, future tense does not indicate someone is witnessing or has witnessed an event.

Popular culture references. Flagging and removing posts with song lyrics or references to music, bands, video games, movies, or television shows can greatly improve results as it is not uncommon for eyewitness features to be referencing a fictional event the user saw in one of these mediums.

Temporal markers. Language such as “last night, last week, weeks ago, months ago” and similar phrases suggest an event happened an extended period of time in the past, and is not a current eyewitness.

2.3 Finding eyewitness events

Identifying an event from a set of eyewitness posts can be done using a simple clustering approach. Most common clustering methods on text data fo-

³While it is possible to have a psycho- or sociopathic witness who would not react with typical trauma emotions (Fowles, 1980; Ekman, 1985; Herpetz et al., 2001; Patrick, 1994), we judged this situation to be sufficiently rare to be discounted. In addition, while “gallows humor” occurs among first responders, it is not generally widely spread outside their cohort (Moran, 2002), or is indicative of after-the-fact third party commentary (Phillips, 2015; Davies, 2003).

cus on semantic similarity. However, the eyewitness filters we created already enforced a level of semantic similarity for their resulting documents, so such clustering would not be effective for our use case.

Multiple separate events of a newsworthy nature are unlikely to be occurring simultaneously in the same location at the same time, or such instances will be considered part of a single large event. Therefore, we propose using a spatio-temporal clustering algorithm to identify potential events. By forcing the spatial proximity to be small (limited to approximately a neighborhood in size) and the temporal locality to be similarly tight, say less than 30 minutes, we can effectively group documents related to events. A good summary of such methods is provided in Kisilevich et al. (2010).

2.4 Method summary

In this section, we describe a complete process for finding eyewitness posts and events.

We start by enriching each document in the feed with geolocation information of the Twitter user, for use in the clustering step to identify events. Geolocation information is central to our approach to finding events, but less than 5% of tweets have location data available. There are many approaches that can be used to enrich social media posts with a prediction of a user’s location. Good summaries are available in Ajao et al. (2015) and Jurgens et al. (2015). We implemented the method described in Apreleva and Cantarero (2015) and were able to add user location information to about 85% of users in our datasets with an 8 km median error. This is accurate enough to place users in their city or neighborhood and enables us to find more posts related to the same event.

After enriching the data, we apply the semantic context topic filters, then the eyewitness linguistic features, and then remove documents matching the non-eyewitness features. This produces a set of eyewitness documents. Specific examples of how to construct these filter rules for criminal shootings, police activity, and protests are available on github⁴. Further event types could be easily constructed by combining a relevant semantic context with the eyewitness linguistic features presented here.

⁴<https://github.com/erikavaris/topics>

This set of eyewitness documents can then be run through a spatio-temporal clustering approach to find events. In our examples, the set of eyewitness documents never had more than around 100-200 documents in a 24-hour period. Since this set is so small, we were able to use a simple approach to clustering. We start by computing the complete distance matrix for all points in the dataset using the greater circle distance measure. The greater circle distance is the shortest distance between two points on a sphere, a good approximation to distances on the surface of the Earth. We can then cluster points using an algorithm such as DBSCAN (Ester et al., 1996). DBSCAN clusters together points based on density and will mark as outliers points in low-density regions. It is commonly available in many scientific computing and machine learning packages in multiple programming languages, and hence a good choice for our work.

For each cluster we then look at the max distance between points in the cluster and check that it is less than a distance τ_d . In our experiments we set the threshold to be about 10 km. If the cluster is within this threshold, we then sort the posts in the cluster by time, and apply a windowing function over the sorted list. If there are more than τ_s documents in the windowed set, we label this set as an event. We used $\tau_s = 1$ for our experiments and time windowing functions, t_w , in sizes between 20 and 40 minutes.

3 Experiments

Using the method described in the previous section, we built filter rules for criminal shootings, unusual police activity, and protests. We ran each filter rule over the Twitter Firehose (unsampled data) on a 24/7 basis. We then sampled random days from each filter, pulling data back for 24 hours, and applied the spatio-temporal algorithm to identify potential events.

Since the resulting sets of documents are relatively small, we measured the accuracy of our method by hand. Generally a method of this type might report on the precision and recall of the solution, but it is not possible to truly measure the recall without reading all messages on Twitter in the time period to make sure posts were not missed. In

our case, we simply conducted a search of news articles after the fact to see if any major events were not picked up by the filter rules on the days that were sampled. For the days that we sampled, there were no major news events that failed to show up in our filters.

We optimized for precision than recall in this study as the goal is to surface potential events occurring on Twitter that may be newsworthy. It is more useful to correctly identify a set of documents as interesting with high accuracy than it is to have found every newsworthy event on the network but with many false positives.

Labeling the accuracy (precision) of a method surrounding semantic topic goals is subjective, so we had multiple people classify the resulting sets as eyewitness, non-eyewitness, and off-topic, and then averaged the results. We used the label “non-eye” on posts that were referencing real events, but were clearly second- or third-hand commentary and not clearly embedded in the local community. Most often these posts represented a later stage of news dissemination where the author heard about the event from afar and decided to comment on it.

While the authors of these tweets were not true eyewitnesses to these events, they are potentially interesting from a news commentary perspective, and were accurate to the event topic. Thus, we may consider the general event semantic accuracy as the combined values of “eyewitness” and “non-eye” tweets.

3.1 Results

3.2 Eyewitness posts

Accuracy results for different sets of eyewitness posts on different dates are shown in Table 1.

What day data was pulled had an impact on the accuracy measurement. Table 1 illustrates this difference particularly in the shooting results. For 02/02/2015, there was more Twitter traffic pertaining to shootings than there was on 06/15/2015, which likely influenced the higher eyewitness accuracy of 72% vs. 46%. We have generally observed on Twitter that when major events are occurring the conversation becomes more focused and on topic, and when nothing major is happening results are lower volume and more noisy.

Table 1: Accuracy results for different eyewitness filters. Data were pulled for a 24 hour period on the random dates shown. Count is the total number of documents in each set.

Topic	Date	Count	Eyewitness	Non-Eye	Off-Topic	Semantic Accuracy
Shooting	2/4/15	126	72%	21%	7%	93%
Shooting	6/15/15	41	46%	20%	34%	66%
Police	3/24/15	100	73%	14%	13%	87%
Police	6/17/15	293	71%	11%	18%	82%
Protests	5/23/15	196	56%	31%	13%	87%
Protests	6/23/15	89	52%	25%	24%	77%
Averages			62%	20%	18%	82%

In these data pulls, the combined eyewitness and non-eyewitness general semantic accuracy was 93% and 66%, respectively. We note that on average the accuracy of our filters is 82% across the days and filters measured.

3.2.1 Events

The approach outlined in 2.3 successfully surfaced on-topic events from the sets of eyewitness tweets. We found its effectiveness to be low on individual filter rules due to the low volume of tweets. We were able to find more relevant clusters by combining the criminal eyewitness topic streams – shootings, police activity, and protests – that corresponded to events we could later find in the news media.

In running these experiments, we found that it was important to add an additional parameter to the cluster that ensured there were tweets from different users. It was common to find multiple tweets that would cluster from the same user that was sharing updates on a developing situation. Both of these behaviors are of potential interest and the algorithm may be adjusted to weight the importance of multiple updates versus different user accounts.

4 Conclusion

This paper presents a novel combinatory method of identifying eyewitness accounts of breaking news events on Twitter, including simple yet extensive linguistic filters together with grouping bursts of information localized in time and space. Using primarily linguistic filters based on sociolinguistic behavior of users on Twitter, a variety of event types are explored, with easily implemented extensions to further event types.

The filters are particularly appealing in a business application; with minimal training we were able to teach users of our platform to construct new rules to find eyewitness events in different topical areas. These users had no knowledge of programming, linguistics, statistics, or machine learning. We found this to be a compelling way to build real-time streams of relevant data when resources would not allow placing a computational linguist, data scientist, or similarly highly trained individual on these tasks.

The system offers a straightforward technique for eyewitness filtering compared with Diakopoulos et al. (2012), easily implemented in a streaming environment, requiring no large training datasets such as with machine learning, and achieving higher accuracies than comparable machine learning approaches (Imran et al., 2013). Together with spatio-temporal clustering to identify eyewitness tweets that are spatially and temporally proximate, our eyewitness filter presents a valuable tool for surfacing breaking news on social media.

For future research, a machine learning layer could be added with broader linguistic filters, and may help achieve higher recall while maintaining the high accuracy achieved with our narrow linguistic keywords.

References

- Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment analysis of twitter data. In *Proceedings of the Workshop on Language in Social Media*, LSM 2011, Portland, Oregon, June. Association for Computational Linguistics.
- Oluwaseun Ajao, Jun Hong, and Weiru Liu. 2015. A sur-

- vey of location inference techniques on twitter. *Journal of Information Science*, 41(6):855–864.
- Sofia Apreleva and Alejandro Cantarero. 2015. Predicting the location of users on twitter from low density graphs. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 976–983, Oct.
- Mary W. Armsworth and Margot Holaday. 1993. The effects of psychological trauma on children and adolescents. *Journal of counseling and Development : JCD*, 72(1), September.
- Claudine Beaumont. 2008. Mumbai attacks: Twitter and flickr used to break news. *The Daily Telegraph*, November.
- Claudine Beaumont. 2009. New york plane crash: Twitter breaks the news, again. *The Daily Telegraph*, January.
- Christie Davies. 2003. Jokes that follow mass-mediated disasters in a global electronic age. In Peter Narváez, editor, *Of Corpse: Death and Humor in Folklore and Popular Culture*. Utah State University Press, Logan, Utah.
- Nicholas Diakopoulos, Munmun De Choudhury, and Mor Naaman. 2012. Finding and assessing social media information sources in the context of journalism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pages 2451–2460, New York, NY, USA. ACM.
- Gabriel Doyle and Michael C. Frank. 2015. Shared common ground influences information density in microblog texts. In *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*, Denver, Colorado, June. Association for Computational Linguistics.
- Paul Ekman. 1985. *Telling lies: Clues to deceit in the marketplace, marriage, and politics*. New York: Norton.
- Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, pages 226–231. AAAI Press.
- Don C. Fowles. 1980. The three arousal model: Implications of gray’s two-factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology*, 17(2):87–104.
- Sabine C. Herpetz, Ulrike Werth, Gerald Lukas, Mutaz Qunaibi, Annette Schuerkens, Hanns-Juergen Kunert, Roland Freese, Martin Fleisch, Ruediger Mueller-Isberner, Michael Osterheider, and Henning Sass. 2001. Emotion in criminal offenders with psychopathy and borderline personality disorder. *Arch Gen Psychiatry*, 58(8).
- Muhammad Imran, Shady Elbassuoni, Carlos Castillo, Fernando Diaz, and Patrick Meier. 2013. Extracting information nuggets from disaster-related messages in social media. In T. Comes, F. Fiedrich, S. Fortier, J. Geldermann, and L. Yang, editors, *Proceedings from the 10th International ISCRAM Conference*, Baden-Baden, Germany, May.
- David Jurgens, Tyler Finethy, James McCorriston, Yi Tian Xu, and Derek Ruths. 2015. Geolocation prediction in twitter using social networks: a critical analysis and review of current practice. In *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media (ICWSM)*.
- Slava Kisilevich, Florian Mansmann, Mirco Nanni, and Salvatore Rinzivillo, 2010. *Data Mining and Knowledge Discovery Handbook*, chapter Spatio-temporal clustering, pages 855–874. Springer US, Boston, MA.
- Carmen C. Moran. 2002. Humor as a moderator of compassion fatigue. In Charles R. Figley, editor, *Treating Compassion Fatigue*, Psychological Stress Series. Routledge, June.
- Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. 2007. Textual affect sensing for sociable and expressive online communication. *ASCII*, 4738.
- Fran H. Norris. 2007. Impact of mass shootings on survivors, families, and communities. *PTSD Research Quarterly*, 18(1).
- Carol S. North, Elizabeth M. Smith, and Edward L. Spitznagel. 1994. Posttraumatic stress disorder in survivors of a mass shooting. *American Journal of Psychiatry*, 151(1), January.
- Christopher J. Patrick. 1994. Emotion and psychopathy: Startling new insights. *Psychophysiology*, 31.
- Saša Petrović, Miles Osborne, Richard McCreadie, Craig Macdonald, Iadh Ounis, and Luke Shrimpton. 2013. Can twitter replace newswire for breaking news? In *International AAAI Conference on Web and Social Media, ICWSM*, Boston, MA, July. Association for the Advancement of Artificial Intelligence (AAAI).
- Whitney Phillips. 2015. *This is Why We Can’t Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture*. MIT Press, February.
- Barry Ritholtz. 2013. Twitter is becoming the first and quickest source of investment news. *The Washington Post*, April.
- Arieh Y. Shalev. 2002. Treating survivors in the acute aftermath of traumatic events. In Rachel Yehuda, editor, *Treating Trauma Survivors with PTSD*. American Psychiatric Publishing, Inc., Washington, D.C.
- Mike Thelwall, Kevan Buckley, Georgios Paltoglou, and Di Cai. 2010. Sentiment strength detection in short informal text. *Journal for the American Society for Information Science and Technology*, 61(12).
- Sam Thielman. 2013. Twitter breaks news of the boston marathon explosions. *AdWeek*, April.

Sarah Vieweg, Amanda L. Hughes, Kate Starbird, and Leysia Palen. 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*, pages 1079–1088, New York, NY, USA. ACM.