Weighted Trilateration Using Domain Knowledge for Event Localization in Twitter Data

Usman Anjum
Dept. of Computer Science
University of Cincinnati, USA
anjumun@ucmail.uc.edu

Justin Zhan
Dept. of Computer Science
University of Cincinnati, USA
zhanjt@ucmail.uc.edu

ABSTRACT

In this paper, we develop a novel approach called *Domain Adapted Weighted Trilateration (DAWT)*. We use the idea of there are reference coordinates at known locations that scrape the microblog (tweet) counts in time and space (circular regions around the reference coordinate). The change in counts of tweets would be indicative of an event pattern. We propose the use of *DAWT* to combine the information from multiple reference points to find the exact location of the event. We propse the use of domain knowledge in the form of counts of tweet or spatial relationship between reference points to calculate the weights for trilateration. We use microblogging data collected from Twitter to evaluate our model and compare it with other baseline methods.

Keywords

social networking, event localization, weighted trilateration

INTRODUCTION

In today's digital age, social media platforms such as Twitter serve as real-time information hubs, offering crucial insights into unfolding events. The ability to accurately localize events based on geotagged tweets has significant applications in disaster response, crisis management, public safety, and market analysis. However, determining the precise geographic location of an event using social media data remains a challenging task due to the inherent sparsity, noise, and inconsistencies in the data.

Traditional event localization methods rely on direct spatial information, such as geotags, place names, or user locations. However, these methods face severe limitations due to the underdeveloped nature of Twitter data, where only a small fraction of tweets contain reliable geographic information. To address this challenge, previous research has proposed alternative approaches that analyze the spatial distribution of tweet activity instead of relying on explicit location data. One such approach is SPAtial REconstruction (SPARE) (Anjum, 2022; Anjum et al., 2022), which frames event localization as a problem of detecting changes in aggregated tweet counts across different regions.

Building upon this concept, we introduce Domain Adapted Weighted Trilateration (DAWT), a novel event localization method that enhances trilateration by incorporating domain-adapted weights. Traditional trilateration estimates an unknown location using distance measurements from multiple known reference points, a principle widely used in GPS and wireless positioning systems (Dargie & Poellabauer, 2010). Inspired by this, our approach treats social media users as "social sensors" whose tweet activity can signal the presence of an event. Unlike conventional trilateration, DAWT assigns weights to reference points based on variations in tweet counts, thereby improving the accuracy of event localization (Figure 1).

Our method consists of two key innovations: (1) A domain-adapted weighting scheme, which refines trilateration by prioritizing reference points based on tweet activity patterns; and (2) An enhancement using kriging, a geostatistical interpolation technique (Cressie, 1993; Wackernagel, 1995), which models the spatial relationship between reference points to improve weight assignments. We evaluate DAWT on mass shooting events, which serve as an ideal test

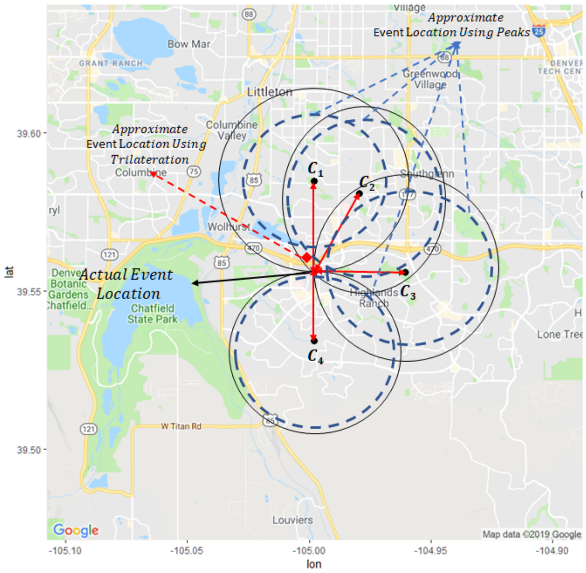


Figure 1. Basic SPARE framework

case due to their sudden and unpredictable nature. Our results demonstrate that DAWT can localize events more accurately than traditional trilateration techniques, although its effectiveness varies across datasets.

This work makes the following contributions:

- We propose DAWT, a novel adaptation of trilateration for event localization using social media data.
- We introduce a domain-adapted weighting strategy to improve the reliability of trilateration by leveraging tweet count variations.
- We incorporate kriging-based interpolation to refine weight calculations and assess its impact on localization accuracy.
- We conduct an empirical evaluation on real-world mass shooting events, demonstrating the strengths and limitations of DAWT in comparison to traditional methods.

While our study focuses on mass shooting events, the DAWT framework can be extended to other event types, such as natural disasters and traffic incidents. Future work will explore alternative weighting mechanisms, including temporal dynamics and user credibility, to further enhance localization accuracy.

LITERATURE REVIEW

Our literature review consists of reviewing relevant work on event localization and trilateration.

Event Detection & Localization

Numerous surveys have summarized the research conducted on event detection within the microblogging domain (Akiladevi et al., 2023; Atefeh & Khreich, 2015; Cordeiro & Gama, 2016; Garg & Kumar, 2016; Hasan et al., 2018; Imran et al., 2015; Li et al., 2023; Ozdikis et al., 2017; Steiger et al., 2015; Zheng et al., 2018). Each survey focuses on a particular aspect of event detection. For instance, (Li et al., 2023) identified key datasets and potential research areas in event detection using social media, while (Akiladevi et al., 2023) provided an overview of primary issues addressed so far and suggested future research directions in the field of event detection using social media.

Similarly, (Atefeh & Khreich, 2015) extensively discussed the challenges and limitations in event detection and localization due to tweet content, such as ambiguous texts and lack of relevant data. A systematic survey on Location-Based Social Networks was presented in (Steiger et al., 2015). (Cordeiro & Gama, 2016) created a taxonomy of event detection in social media, classifying different methods by event type, detection method (supervised or unsupervised), and whether the detected event is new or recurring. (Garg & Kumar, 2016) focused on the various types of datasets (images, texts, audio, etc.) in social media used for event detection. (Imran et al., 2015) discussed the approaches, challenges, and benefits of using social media messages for detecting emergency events. Finally, (Hasan et al., 2018) conducted a survey on methods for real-time event detection.

The survey by Ozdikis et al. (Ozdikis et al., 2017) examined a list of techniques proposed for event localization on Twitter, classifying them based on spatial features used for location estimation and the granularity of location estimation. Similarly, Zheng et al. (Zheng et al., 2018) explored the geo-location problem and challenges associated with finding the home location, tweet location, and mentioned location. They highlighted that methods heavily rely on tweet content, and the noisy and short nature of tweets makes geo-location challenging.

In another work (Cheng & Wicks, 2014), clusters are created in space and time, and then classified based on the topics within each cluster. Unsupervised social event detection using hybrid graph contrastive learning and reinforced incremental clustering involves identifying social events without labeled data by leveraging graph-based representations and incrementally clustering data points to improve detection accuracy (Y. Guo et al., 2024).

In Abdelhaq et al., the region of interest is divided into cells, and keywords are extracted based on their temporal and geo-spatial properties and then clustered. The *Eyewitness* algorithm (Krumm & Horvitz, 2015), and its real-time version (Comito et al., 2017), analyze a corpus of geotagged tweets over localized regions for unusual spikes in tweet counts. They divide the area of interest into triangles and use time periods of different lengths. An event is defined as a peak above a baseline tweet count, obtained through regression. A geo-social event detection method focusing on the geographical regularities of local crowd behaviors to detect events has also been proposed (Lee & Sumiya, 2010). They implemented their method using a fixed time window, and their geographic grids are created based on a clustering-based space partition method.

Sakaki et al. (Sakaki et al., 2010) used tweets to find the epicenter of an earthquake and the trajectory of typhoons. Another work (Ozdikis et al., 2013) estimates an event's location by assigning probabilities using Dempster–Shafer (DS) theory based on geotags, texts in tweets, and user profiles. The location of the events is found by clustering. However, they only considered two levels of granularity and require coordinates and names for assigning probabilities. This work was extended to incorporate real-time tweets (Ozdikis et al., 2016). Dempster–Shafer (DS) was also used to find coarse-grained information (like city names) and fine-grained information (coordinates of the event) (Shahraki et al., 2019). They focused mostly on traffic accidents.

The most relevant work to our proposed research is based on the framework called *SPatial Aggregation REconstruction (SPARE)*. SPARE considered peaks to find the location of an event (Anjum, 2022; Anjum et al., 2022). The geographic region was divided into concentric circles. Peaks were obtained when the number of tweets measured at specific radii from the center of the circle was disaggregated. Peaks were obtained from multiple reference coordinates, and then the information was combined using trilateration to obtain a more exact geographic coordinate. Low pass filters were also used to remove any unwanted random peaks. A more advanced version of a filter to remove random peaks, called the Semantic Decay Filter (SDF), was proposed (Ben Lazreg et al., 2020). The SDF removes peaks that have low similarity between texts in tweets. This work was further expanded in (Anjum et al., 2022), where instead of peaks, bi-directional long short-term memory (Bi-LSTM) was used to find event signatures. This method was called *SPatial Event Localization (SPEL)*. Bi-LSTM's training was enhanced by using synthetic data generated using generative adversarial networks (GAN). Synthetic data generated like agent-based modeling (Anjum et al., 2021) and BERT (X. Guo et al., 2022) has also been proposed to improve event detection .

Trilateration

Trilateration in the domain of sensors has been used extensively (Kumari et al., 2019). Localization is performed using received signal strength indicator (RSSI), time of arrival (ToA), angle of arrival (AoA), and time difference of arrival (TDoA). A node can be localized using trilateration, triangulation, quadrilateration, or multilateration. These method requires the knowledge of distances from three or more reference points. The node or the sensor with its unknown location computes its position as the intersection point of the circles using radius as the distance between the reference node and the sensor nodes.

The concept of weight based trilateration has been proposed by many papers for indoor positioning. (Kumari et al., 2019) applied the fingerprinting technique and then applied the K-nearest neighborhood (KNN) algorithm to find location of indoor occupants using RSSI measurements. Another fingerprinting technique used Kalman filters for weighted trilateration for indoor positioning in Bluetooth low energy devices (BLE) (Cantón Paterna et al., 2017). Similarly, (Garcia et al., 2024) proposed weighted trilateration with extended Kalman filters for indoor positioning with ultra wideband systems.

In terms of trilateration for event detection, trilateration was used to find the gunshot location with the use of AI agents as sensors. (Westhuis et al., 2011) used trilateration to identify location using RFID signal strength.

Trilateration for event detection in microblogging data has not been studied extensively and has only been proposed by (Anjum, 2022; Anjum et al., 2022, 2023) which has been described in the previous subsection.

METHODOLOGY

We begin this section by first describing the basics behind the SPARE methodology. Next we define what trilateration is and finally, we define *Domain Adapted Weighted Trilateration (DAWT)*.

Spatial Temporal Grid (STG) for Twitter Data

The use of trilateration for event localization was first proposed within the framework called SPatial Aggregation REconstruction (SPARE). In SPARE, multiple reference points measure aggregated counts of tweets in incremental concentric circles (layers). The layer where a possible event occurs is identified by the largest jump (or peak) between consecutive layers. By using this layer as the distance to the event and combining these distances from different reference points through trilateration, the exact geographic coordinates of the event can be determined.

Event localization using trilateration

In this section, we describe basic formulation of trilateration.

For each of the reference coordinates ($C_i = (x_i, y_i)$), we would identify a specific layer as a radius at which the event occurred. To get the exact latitude and longitude, we combine the different layers for each of the reference coordinate to get an exact latitude and longitude (instead of only a radius), as was illustrated in Figure 1. Trilateration has been used widely in sensor localization to find the location of an unknown sensor based on its distance from sensors at fixed known locations (Dargie & Poellabauer, 2010).

Let there be n reference coordinates, C_i where $i = 1....n$ (called anchors nodes in trilateration), whose coordinates are represented in the 2D Cartesian plane as (x_i, y_i) where $i = 1....n$. The unknown coordinate (which is the possible event location) is represented by coordinates $\mathbf{x} = (x, y)$. The distance between the approximate event location and the reference coordinates is the layer at which the significant peak lies. It is denoted r_i for reference coordinate C_i respectively. The relationship between sensor nodes, approximate events, and distances is represented as $A\mathbf{x} = b$, where:

$$A = \begin{bmatrix} 2(x_n - x_1) & 2(y_n - y_1) \\ 2(x_n - x_2) & 2(y_n - y_2) \\ \vdots & \vdots \\ 2(x_n - x_{n-1}) & 2(y_n - y_{n-1}) \end{bmatrix}$$

$$b = \begin{bmatrix} {r_1}^2 - {r_n}^2 - {x_1}^2 - {y_1}^2 + {x_n}^2 + {y_n}^2 \\ {r_2}^2 - {r_n}^2 - {x_2}^2 - {y_2}^2 + {x_n}^2 + {y_n}^2 \\ \vdots \\ {r_{n-1}}^2 - {r_n}^2 - {x_{n-1}}^2 - {y_{n-1}}^2 + {x_n}^2 + {y_n}^2 \end{bmatrix}$$

Using least square estimation, $\mathbf{x}$ can then be found using:

$$\mathbf{x} = (A^T A)^{-1} A^T b \tag{1}$$

Weighted Trilateration

Our proposed approach *Domain Adapted Weighted Trilateration (DAWT)* is based on weighted trilateration. Weighted trilateration is an advanced technique used for determining the precise location of a point by measuring distances from multiple known reference points, similar to standard trilateration. However, it incorporates weights to account for the varying accuracy and reliability of these distance measurements. This method is particularly useful in scenarios where some measurements are more accurate than others due to factors like signal strength, environmental conditions, or sensor quality.

Weighted trilateration can offer several advantages over traditional trilateration, particularly in scenarios where data quality and reliability vary. It can improve accuracy by assigning weights to different data points based on their reliability or some other metric like proximity to event, user credibility, timestamp relevance, number of twitter users, etc. Hence, weighted trilateration can reduce the impact of noisy or inaccurate data. This leads to more precise localization. Weighted trilateration can also more robustness to noisy data by reducing relevance of unreliable data.

Weighted trilateration can be presented by modifying Equation 1:

$$\mathbf{x} = (A^T W A)^{-1} A^T W b \tag{2}$$

where $W = [w_1, ...w_n]$ are the weights for each reference coordinate C_i.

Assigning Weights for Weighted Trilateration

The major task in DAWT is how to find the weights. When using microblogging data (like tweets) for event localization, the aggregated tweet counts from different reference points can provide valuable information about the proximity of an event. We hypothesize that locations closer to the event will generally have higher jumps in tweet counts, manifested by peaks, while those farther away will have lower jumps in tweet counts. A simple and straightforward way of assigning weights is by normalizing the differences in tweet counts between the layers of a reference point. In this way we can assign weights that reflect the relative reliability of each reference point's data.

However, there are some issues with assigning weights in this way. It may not be necessary that magnitude of the peaks represent close proximity to events. Peaks may be due to noise or geographic and temporal distribution of people which may show peaks of high magnitude regardless of proximity to events.

Hence, we propose to use Ordinary Kriging (OK) to find weights for the trilateration. Kriging is a geostatistical method that provides an optimal way to assign weights to data points when estimating values at unsampled locations (Wackernagel, 1995). It takes into account both the distance between data points and the spatial correlation structure of the data, making it a powerful tool for spatial interpolation.

Ordinary Kriging is a specific type of kriging that assumes a constant but unknown mean across the study area (Wackernagel, 1995). It calculates weights based on the spatial correlation (semivariogram) and minimizes the estimation variance, leading to more accurate and reliable estimates. Hence, It can provide more reliable localization results by leveraging the spatial structure of the data, making it a better choice for applications requiring high accuracy.

To find the weights w_i, solve the kriging system:

$$\mathbf{W} = \begin{bmatrix} \gamma(C_1, C_1) & \gamma(C_1, C_2) & \cdots & \gamma(C_1, C_n) & 1 \\ \gamma(C_2, C_1) & \gamma(C_2, C_2) & \cdots & \gamma(C_2, C_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(C_n, C_1) & \gamma(C_n, C_2) & \cdots & \gamma(C_n, C_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(C_1, C_0) \\ \gamma(C_2, C_0) \\ \vdots \\ \gamma(C_n, C_0) \\ 1 \end{bmatrix}$$

where $C_1 \dots C_n$ are the n reference points, $\gamma(C_1, C_2)$ is the semivariogram of reference points C_1 and C_2 respectively. The semivariogram is calculated using the formula:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(C_i) - Z(C_i + h)]^2 \tag{3}$$

where $Z(C_i)$ is the value at reference point C_i, h is the lag distance, i.e. the distance between two reference points in a spatial dataset. It is used to group the square differences between the values of the reference points: $Z(C_i) - Z(C_j)]^2$. For each group of distances, we calculate the average of the squared differences. This average is the semivariogram value $\gamma(h)$ for that distance.

Finally, we fit a model to understand the spatial correlation and determine the weights. The typical models are spherical, exponential, linear, and gaussian. Each model has it own advantages and we pick the model that presents the best localization accuracy.

Practical Implementation

There are multiple ways to implement the inverse function for trilateration. Two common ways of implementation are using the `minimize` function (Developers, 2023) or through `np.linalg.inv` (Koubaa, 2024). Both these methods have their advantages and disadvantages and give different results.

The `minimize` function from the `scipy.optimize` module is used for optimization problems. It minimizes the error between the measured distances and the distances calculated from the estimated coordinates. This method is flexible and can handle non-linearity and constraints, but it is computationally expensive and requires a good initial guess for convergence.

$$\text{minimize} \quad f(\mathbf{x}) = \sum_{i=1}^{n} w_i \left(\sqrt{(x - x_i)^2 + (y - y_i)^2} - d_i \right)^2 \tag{4}$$

The `np.linalg.inv` function from the `numpy` module computes the inverse of a matrix. It can be used to solve a system of linear equations derived from the distance measurements. This function is computationally efficient, simple to implement for linear problems. However, it assumes a linear relationship, sensitive to noise and measurement errors, and requires the matrix to be invertible (non-singular).

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \Rightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{5}$$

In our experiments, we use both methods for implementation and show the difference in the results.

EXPERIMENTS AND RESULT

In this section, we use the real-world dataset to look at the performance of our model.

Dataset

The data for analysis was obtained from TBAM (Anjum et al., 2021) and SPARE (Anjum, 2022; Anjum et al., 2022)[1]. The description of the datasets is summarized in Table 1. The table shows the event name, the reference coordinates at which the number of tweets was collected, the actual coordinate of the event, and the time at which the event occurred. The *Label* refers to the name used for the data sets in the experiments. Each of the events is an example of an unknown event. For each reference coordinate, the number of tweets was collected from a radius of 1.0mile to 2.8miles with 0.1 mile increments. The number of tweets aggregated for the time 2 days before the event to 2 days after the event occurrence data with a 1-hour time step. Figures 2(a) and 2(b) shows the changing number of tweets at the outermost layer (which is at 2.8miles) collected from the 4 different reference points for the STEM and VIRG datasets respectively.

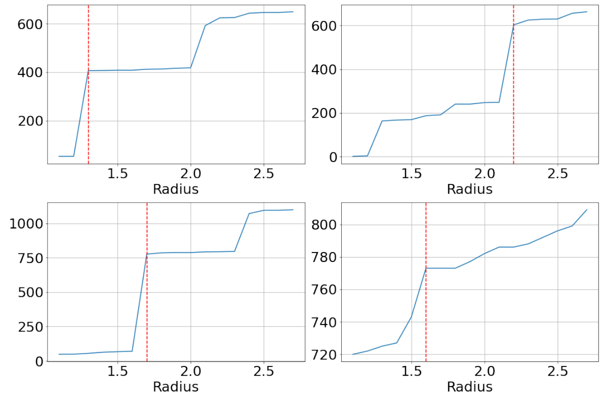
Results

The results of our analysis can be found in Table 2. In the dataset, **SPARE** is the column that represents the *SPARE* algorithm (Anjum et al., 2022). It is equivalent to trilateration algorithm with equal weights and using `np.linalg.inv` function. The **SPARE-DAWT** is our proposed algorithm column is the weighted trilateration with the weights equal to the highest difference between the number of tweets in consecutive radii. The **SPARE-DAWT-kriging** is the column weighted using the ordinary kriging method. The **nl** term in the columns denotes the use of `minimize` function to implement trilateration. Similar to before, the **DAWT-nl** and **DAWT-nl-kriging** are trilateration implemented using `minimize` function and weighted generated using difference in number of tweets or through kriging respectively.

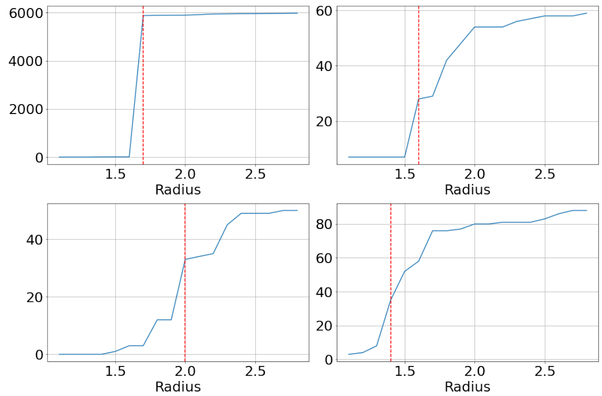
In the table, we report the distance error in miles. The distance error found between the actual event coordinates from Table 1 and the coordinates from trilateration using the Haversine Distance. The Haversine formula calculates the great-circle distance between two points on a sphere given their longitudes and latitudes:

$$d = 2r \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(x_i)\cos(x_j)\sin^2\left(\frac{\Delta\lambda}{2}\right)} \right)$$

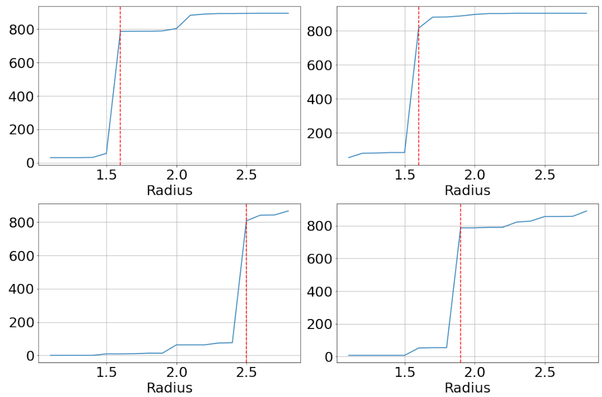
[1] https://github.com/usmananjum/SPARE_Data.git



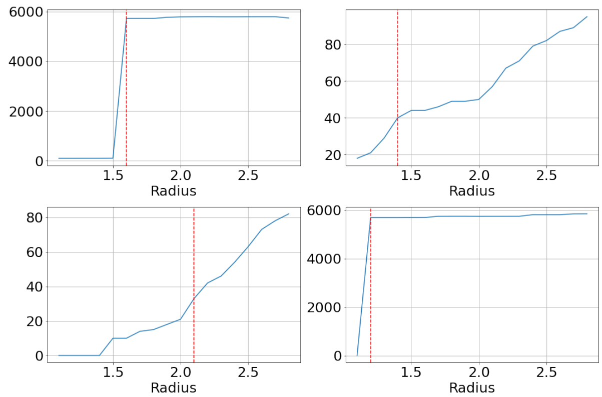
(a) STEM Dataset Collected from 4 Reference Points



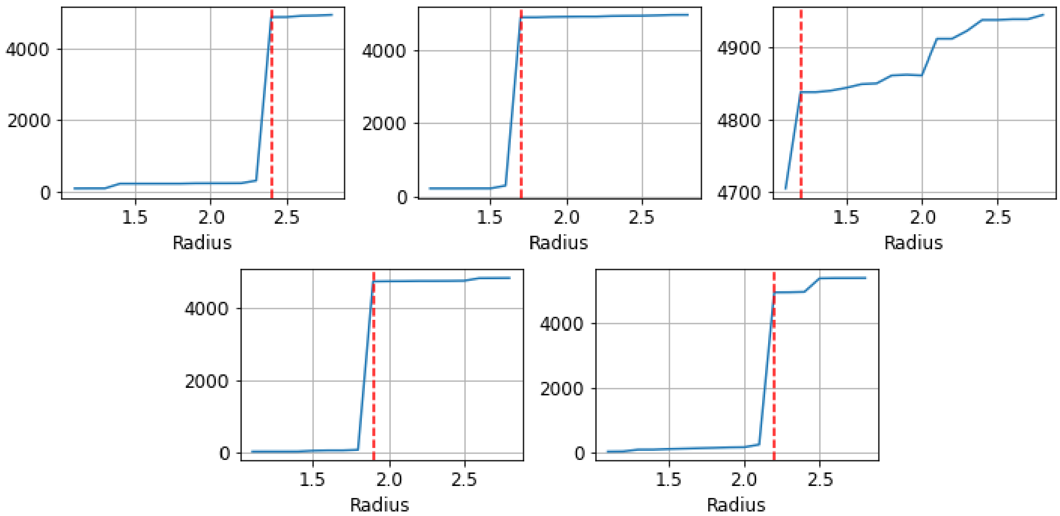
(b) VIRG Dataset Collected from 4 Reference Points



(c) GAR Dataset Collected from 4 Reference Points



(d) ELP Dataset Collected from 4 Reference Points



(e) SANTA Dataset Collected from 5 Reference Points

Figure 2. Changing Aggregated Number of Tweets with distance from C_i (red-line shows the possible event location) - y-axis is counts of tweets and x-axis is the distance from C_i in miles

Table 1. Detailed information about the events - the location, time and coordinates for the location of reference coordinates (location of areas from where tweet counts are measured)

Data Set	Reference Coordinate (Latitude, Longitude)	Label	Event Date & Time	Event Coordinate (Latitude, Longitude)
STEM School Shootings	39.58482, −104.99790	STEM	05-07-2019 1:53pm	39.556, −104.9979
	39.58096, −104.97928			
	39.55599, −104.96067			
	39.53438, −104.99790			
Virginia Beach Shootings	36.75089, −76.02167	VIRG	05-31-2019 4:44pm	36.7509, −76.0575
	36.75089, −76.02167			
	36.72206, −76.05750			
	36.75089, −76.09333			
Garlic Festival Shootings	37.02661, −121.58528	GAR	07-28-2019 5:40pm	36.997778, −121.585278
	36.99777, −121.54933			
	36.96894, −121.58528			
	36.99777, −121.62123			
El-Paso Shootings	31.80596, −106.38430	ELP	09-03-2019 10:45am	31.7771, −106.3843
	31.7771, −106.3505,			
	31.74824, −106.38430			
	31.7771, −106.4096			
Santa Clarita Shootings	34.44885, −118.52000	SANTA	11-14-2019 7:38am	34.4419, −118.5177
	34.4200, −118.4852,			
	34.4450, −118.5024,			
	34.39115, −118.52000			
	34.4200, −118.5548			

where: d is the distance between the two points (in miles), r is the Earth's radius (mean radius = 3,959 miles), x_i and x_j are the latitudes of the two points (in radians), Δx is the difference in latitudes, and $\Delta\lambda$ is the difference in longitudes.

Table 2. Summary of Results - Error in distance between actual event coordinates and event coordinates found from different methods (in miles)

Dataset	SPARE	SPARE-DAWT	SPARE-DAWT-kriging	nl	DAWT-nl	DAWT-nl-kriging
STEM	0.4079	0.4792	0.8721	**0.4077**	0.5490	0.8507
VIRG	0.2683	0.4230	0.2182	0.1329	0.3330	**0.0647**
SANTA0	**0.2884**	0.3696	0.6421	0.4753	0.3006	0.4173
SANTA1	0.2562	0.1197	0.1591	0.4729	**0.0887**	0.1287
GAR	0.4756	0.4756	1.3514	0.3699	0.3665	**0.3496**
ELP	0.2427	**0.0383**	0.7435	0.2382	1.5813	0.2865

Discussion

The results from Table 2 reveal some interesting conclusions and each dataset provides a unique usecase for the which method of trilateration to use.

For the $STEM$ dataset, the non-weighted basic trilateration algorithm performs the best. Adding weights to trilateration actually significantly reduces performance. We believe this may be due to the placement of the reference points for $STEM$. The reference points are at bearings 0°, 30°, 90°, and 180°, which means that the weights generated from the number of tweets and kriging may not have captured the exact spatial relationships, thereby reducing overall performance. In this case, we may need to explore other external factors such as temporal analysis, past tweeting behavior, and the geographical distribution of people and their Twitter usage in a specific area.

For the $VIRG$ and GAR dataset, we can clearly see that **DAWT-nl-kriging** gives the best accuracy and kriging is able to find the spatial relationship between the reference points that can be used for weights for trilateration. For these datasets, the bearings are 0°, 90°, 180°, and 270° and all at actual $2 - miles$ from the event.

The $SANTA0$ and $SANT1$ datasets are the $SANTA$ dataset with 3- or 4-reference points. The results show that adding a reference point improves the accuracy. However, we still need further research on how many reference points are the optimal number of reference points. For this data, the best method was when the data was weighted using the difference in number of tweets with the **DAWT-nl** providing the smallest error.

For the ELP dataset, the best performance is obtained using the **SPARE-DAWT** algorithm. In this case, we see that the difference in counts of tweets is able to provide very good weights for weighted trilateration.

Conclusion

In this paper, we proposed a novel algorithm called *Domain Adapted Weighted Trilateration (DAWT)*. Our approach introduces the idea of using weights for trilateration to localize events. We employed the *SPARE* methodology, which uses peaks to identify possible event locations. We determined weights for trilateration using two approaches: one based on the maximum value of the difference in tweet counts across consecutive layers, and the other using kriging. Our experiments on a real-world dataset demonstrated that $DAWT$ significantly improves the accuracy of localizing unknown events in most cases.

A major advantage of our method is that it helps us understand the main areas that are more affected by real-world events.

Our method opens up several interesting directions for future research. One potential avenue is to incorporate the temporal information of tweets to calculate weights using kriging. Additionally, we could explore other methods for weight calculation, such as other forms of kriging, regression, clustering, or more advanced neural network techniques like graph-based neural networks, Bayesian methods, Tsetlin Machines (Anjum & Zhan, 2024) or contrastive self-supervised learning.

It would also be interesting to explore other factors for determining weights for trilateration like past tweeting behavior, time of day, and the geographical distribution of people and their Twitter usage in a specific area. Our results show that spatial proximity or difference in counts of tweets may not be the only factors that affect the tweeting behavior of users, which may be why $DAWT$ performs not so well on some datasets. We also believe that data sparsity might also play a role in determining use of $DAWT$ and non-linear trileration method.

We have only used peaks as the distance to event. Consequently, we can also be explore other algorithm (Anjum et al., 2023) for finding event signatures and then using those signatures for event localization.

We have focused solely on shooting data in this study. Therefore, another promising direction for future research is to apply *DAWT* to other data sources and using domain knowledge to find weights for trilateration, such as localizing weather anomalies or pinpointing areas affected by climate change.

REFERENCES

Abdelhaq, H., Sengstock, C., & Gertz, M. (2013). Eventweet: Online localized event detection from twitter. *Proceedings of the VLDB Endowment*, *6*(12), 1326–1329.

Akiladevi, G., Arun, M., & Pradeepkandhasamy, J. (2023). Event detection in social media analysis: A survey. *Proceedings of the 20th International Conference on Inventive Communication and Computational Technologies (ICICCT)*, *757*, 39–53.

Anjum, U. (2022). *Localization of events using underdeveloped microblogging data* [Doctoral Dissertation]. University of Pittsburgh.

Anjum, U., Zadorozhny, V., & Krishnamurthy, P. (2021). Tbam: Towards an agent-based model to enrich twitter data. *18th ISCRAM Conference Proceedings. Blacksburg, VA (USA): Virginia Tech.*

Anjum, U., Zadorozhny, V., & Krishnamurthy, P. (2022). Localization of unidentified events with raw microblogging data. *Online Social Networks and Media*, *29*, 100209.

Anjum, U., Zadorozhny, V., & Krishnamurthy, P. (2023). Localization of events using neural networks in twitter data. In J. Radianti, I. Dokas, N. Lalone, & D. Khazanchi (Eds.), *Proceedings of the 20th international iscram conference* (pp. 909–919). University of Nebraska at Omaha.

Anjum, U., & Zhan, J. (2024). A novel tsetlin machine with enhanced generalization. *Electronics*, *13*(19), 3825.

Atefeh, F., & Khreich, W. (2015). A survey of techniques for event detection in twitter. *Computational Intelligence*, *31*(1), 132–164.

Ben Lazreg, M., Anjum, U., Zadorozhny, V., & Goodwin, M. (2020). Semantic decay filter for event detection. *17th ISCRAM Conference Proceedings. Blacksburg, VA (USA): Virginia Tech.*, 14–26.

Cantón Paterna, V., Calveras Auge, A., Paradells Aspas, J., & Perez Bullones, M. A. (2017). A bluetooth low energy indoor positioning system with channel diversity, weighted trilateration and kalman filtering. *Sensors*, *17*(12), 2927.

Cheng, T., & Wicks, T. (2014). Event detection using twitter: A spatio-temporal approach. *PloS one*, *9*(6), e97807.

Comito, C., Falcone, D., & Talia, D. (2017). A peak detection method to uncover events from social media. *Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on*, 459–467.

Cordeiro, M., & Gama, J. (2016). Online social networks event detection: A survey. In *Solving large scale learning tasks. challenges and algorithms* (pp. 1–41). Springer.

Cressie, N. A. (1993). *Statistics for spatial data*. John Wiley & Sons.

Dargie, W., & Poellabauer, C. (2010). *Fundamentals of wireless sensor networks: Theory and practice*. John Wiley & Sons.

Developers, S. (2023). *Minimize — scipy v1.15.1 manual*.

Garcia, J., Martinez, L., & Perez, A. (2024). Opportunistic fusion of ranges from different sources for indoor positioning. *Proceedings of the 2024 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 123–130.

Garg, M., & Kumar, M. (2016). Review on event detection techniques in social multimedia. *Online Information Review*, *40*(3), 347–361.

Guo, X., Anjum, U., & Zhan, J. (2022). Cyberbully detection using bert with augmented texts. *2022 IEEE International Conference on Big Data (Big Data)*, 1246–1253.

Guo, Y., Zang, Z., Gao, H., Xu, X., Wang, R., Liu, L., & Li, J. (2024). Unsupervised social event detection via hybrid graph contrastive learning and reinforced incremental clustering. *Knowledge-Based Systems*, *284*, 111225.

Hasan, M., Orgun, M. A., & Schwitter, R. (2018). A survey on real-time event detection from the twitter data stream. *Journal of Information Science*, *44*(4), 443–463.

Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2015). Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, *47*(4), 67.

Koubaa, A. (2024). *Numpy for data science cheat sheet*.

Krumm, J., & Horvitz, E. (2015). Eyewitness: Identifying local events via space-time signals in twitter feeds. *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 20.

Kumari, J., Kumar, P., & Singh, S. K. (2019). Localization in three-dimensional wireless sensor networks: A survey. *The Journal of Supercomputing*, *75*, 5040–5083.

Lee, R., & Sumiya, K. (2010). Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection. *Proceedings of the 2nd ACM SIGSPATIAL international workshop on location based social networks*, 1–10.

Li, Q., Chao, Y., Li, D., Lu, Y., & Zhang, C. (2023). Event detection from social media stream: Methods, datasets and opportunities. *arXiv preprint arXiv:2306.16495*.

Ozdikis, O., Oguztuzun, H., & Karagoz, P. (2013). Evidential location estimation for events detected in twitter. *Proceedings of the 7th Workshop on Geographic Information Retrieval*, 9–16.

Ozdikis, O., Oğuztüzün, H., & Karagoz, P. (2016). Evidential estimation of event locations in microblogs using the dempster–shafer theory. *Information Processing & Management*, *52*(6), 1227–1246.

Ozdikis, O., Oğuztüzün, H., & Karagoz, P. (2017). A survey on location estimation techniques for events detected in twitter. *Knowledge and Information Systems*, *52*(2), 291–339.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes twitter users: Real-time event detection by social sensors. *Proceedings of the 19th international conference on World wide web*, 851–860.

Shahraki, Z. K., Fatemi, A., & Malazi, H. T. (2019). Evidential fine-grained event localization using twitter. *Information Processing & Management*, *56*(6), 102045.

Steiger, E., De Albuquerque, J. P., & Zipf, A. (2015). An advanced systematic literature review on spatiotemporal analyses of t witter data. *Transactions in GIS*, *19*(6), 809–834.

Wackernagel, H. (1995). *Multivariate geostatistics: An introduction with applications*. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03098-1_11

Westhuis, J., Dunkel, J., & Bruns, R. (2011). Event-based location detection. *SoftCOM 2011, 19th International Conference on Software, Telecommunications and Computer Networks*, 1–5.

Zheng, X., Han, J., & Sun, A. (2018). A survey of location prediction on twitter. *IEEE Transactions on Knowledge and Data Engineering*, *30*(9), 1652–1671.