International Cartographic Association
Association Cartographique Internationale

Where and why there: location analytics of routine occurrences (LARO) with a case study on traffic accidents

Yanan Wu [a,] *, May Yuan[a]

[a] *School of Economic, Political and Policy Sciences, The University of Texas at Dallas, yxw190021@utdallas.edu, myuan@utdallas.edu*

* Corresponding author

Keywords: where and when, routine occurrences, periodic pattern, traffic accidents, POIs,

Abstract:

Understanding where and when events happened provides a foundation to explore why and how the events took place and to predict their occurrences. Point events, such as traffic accidents, disease cases, and criminal activities, are common in spatial clustering or hotspot analysis. These spatiotemporal clusters or hotspots are subject to two disadvantages. First, the clusters can disappear and re-appear due to varying spatial and temporal units or be subject to the modifiable area unit problem (MAUP). Second, the clusters are fixed in space and time. So, analyses focus on clusters may overlook the correlation among clusters and the repeated point-events. Instead, this research focuses on track the timing and subsequent events, which means detecting where point events routinely occur and how site characteristics and situation dynamics at these locations may explain the routine occurrences. We develop the algorithm, location analytics of routine occurrences (LARO) to uncover locations where events occur routinely. We demonstrate the method with a case study of over 250,000 reported traffic accidents from 2010 to 2018 in Dallas, Texas, the United States of America. We assume that the locations of routine traffic accidents reflect the patterns of life in Dallas' human dynamics and use points of interest (POIs) to reckon the site characteristics and situation dynamics surrounding these locations.

LARO first identifies the routine occurrences by exploiting the periodicity of point events, which in turn determines the proper temporal units of analysis to detect locations of routine occurrences. The periodicity indicates that a value repeats itself in a regular time interval. An approach, discrete Fourier transformation (DFT), identifies the underlying patterns by transforming data items from the time domain to the frequency domain and using Fourier analysis to find the dominant periodicities. The periodic pattern, we coin as routine occurrences, takes place hourly, daily, weekly, or in another period. The general idea of DFT is to discretize time into a temporal bin and distributes point-events into respective temporal units. The result is a frequency distribution of point events over time. We use Fourier transformations to aggregate traffic crash counts at 1-h intervals during weekdays and separately during weekends. Our findings show that traffic accidents follow robust temporality with hour-of-the-day (HOD) and day-of-the-week (DOW) (Figure 1). The multiple R^2 of weekday transformation and weekend, transformation were 0.9598 and 0.823, while the adjusted R^2 of the weekday transformation and weekend transformation were 0.9557 and 0.792.

The second stage of LARO algorithms aims to detect the location of routine occurrences. Conventional approaches to detect clusters or hotspots are susceptible to MAUP. We use disaggregated data to overcome this problem and detect spatial-temporal routine occurrences based on each event and its proximity events. The consideration of proximity is common in criminology, which means that when a crime occurs at a specific location, the surrounding area may experience similar crimes occurring over a period. Based on the idea of proximity, we consider the proximity in our study as a pattern that subsequent events successively occur in proximity around the location of an existing event. The area of spatial proximity is determined by the road network in Dallas, which is one-half of the average size of a city block (~ 750m). The algorithm identifies the hour of a day (HOD) and the day of a week (DOW) of one event at a time and finds all events occurring at the same HOD and DOW within the set spatial proximity. The ability to measure spatiotemporal proximity over time could provide valuable insight into the location of existing events experiencing an increasing, a decreasing, or a constant number of subsequent events over time (Figure 2). Locations in northern Dallas experience increasing subsequent proximate events over years, locations in central Dallas have notable proximate events in 2018, locations in Eastern Dallas exhibit a constant number of proximate events over time, and locations in Southern Dallas show decreasing frequency over the years. The qualification of the location of routine occurrence varies with temporal units. If a location has an event occurring every Tuesday at 7 am, the maximal number of reoccurrences will be 52 events per year. A threshold of 10 percent will require five events on Tuesday at 7 am and at least two consecutive years for the location qualification of routine occurrences.

The third stage of LARO aims to examine the spatial features and their functions around the location of routine occurrences (ROL). We assume the human dynamics contribute to the traffic accident pattern due to the special temporal characteristics of human activities. The human dynamics support by various spatial features at locations at different times.

Abstracts of the International Cartographic Association, 3, 2021.
30th International Cartographic Conference (ICC 2021), 14–18 December 2021, Florence, Italy.
https://doi.org/10.5194/ica-abs-3-318-2021 |

Point of interest (POI) is one of the most popular ways to represents geographic entities (e.g., restaurant, park, and night-life entertainment) with semantic data (e.g., name and function). They work as origin or destination venues for human mobility and can reflect the density of populated areas. To find out the distinctive combinations of spatial features surrounding the detected locations, we use the A-priori algorithm to extract frequent spatial features and generate association rules between two geographic classes, such as traffic accidents and POIs. Except for ROLs, we also differentiate the POI associations to other two locations (normal traffic accident locations and random locations with no traffic accident). In the end, the rules can give summative characteristics of the sites and situations at these three different locations.

The output of this algorithm represents a location of point-event as $lp(x, y, t, f)$ with location (x, y), time of occurrences (t), and a set of features (f) characterizing the location at fine units to differentiate from other locations. Among them, the spatial features characterize that most frequently (at the fourth quantile, > 75% of all the same POIs in Dallas) POIs appear around the ROLs are of parks, parking lots, and entertainment places. Specifically, rules between ROLs and high frequencies of these POIs with the support of almost 21% and confidence of almost 85%. These indicate that 21% of locations with high frequencies of these POIs are within 750m of ROLs, and around 85% of all ROLs are within 750m of locations with high frequencies of these POIs. While for normal and random locations, they associated with low frequencies (at the first quantile, < 25% of all the same POIs in Dallas) of POIs. This study enriches the geographical meaning of the location of events based on these attributes; a location can be not only a place to support event occurrences but also serves an interrelating role that interacts with time, proximity geographical dynamics, and nearby spatial features. The better quantification of the dynamic evolution of the locations of events provides greater insights as to the relevant department. For the location of routine occurrences, appropriate mitigation measures can be applied according to the time of occurrences and spatial features around the location.

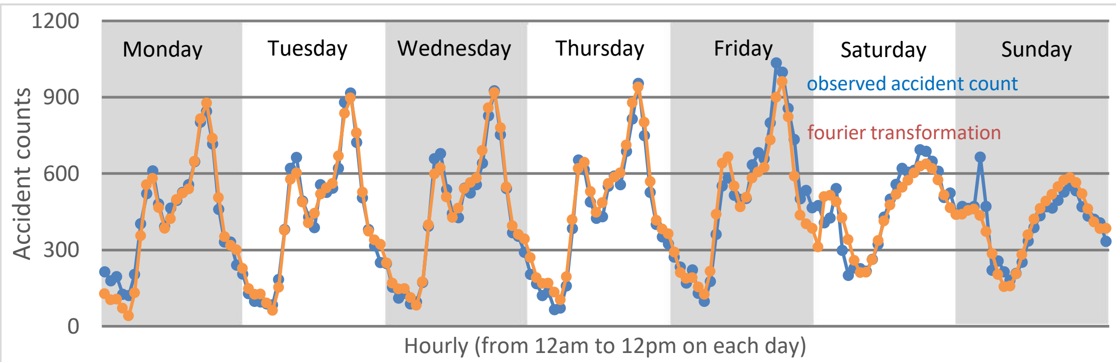

Figure 1. Observed traffic accident counts and fitted value of Fourier transformation

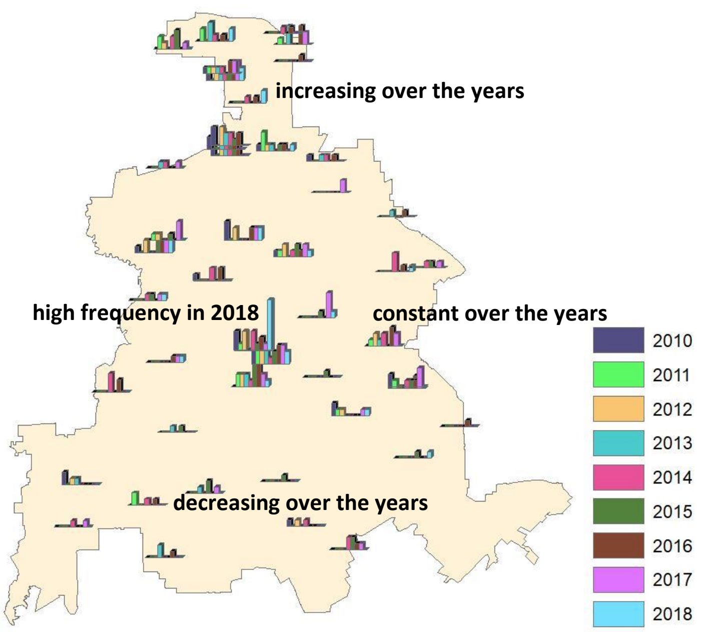

Figure 2. Locations where traffic accidents occurred year after year at $\mathbf{7}_{th}$ hour Mondays

Abstracts of the International Cartographic Association, 3, 2021.
30th International Cartographic Conference (ICC 2021), 14–18 December 2021, Florence, Italy.
https://doi.org/10.5194/ica-abs-3-318-2021 | © Author(s) 2021. CC BY 4.0 License.