L'd Up: Examining the Effects of a New York City Metro Shutdown on Public Discourse Using Twitter Data

Andreas Petutschnig*, Bernd Resch* **, Jen Nelles***, Laxmi Ramasubramanian****

andreas.petutschnig@sbg.ac.at, bernd.resch@sbg.ac.at, gjennife@hunter.cuny.edu, laxmi.ramasubramanian@sjsu.edu

* Department of Geoinformatics - Z_GIS, University of Salzburg, 5020 Salzburg, Austria
** Center for Geographic Analysis, Harvard University, Cambridge, MA 02138, USA
*** CUNY Institute for Sustainable Cities (CISC), Hunter College CUNY, New York City, NY 10065 USA
**** Department of Urban and Regional Planning, San José State University, San José, CA 95192 USA

Abstract. In 2016, announcements of a major renovation plan of the L train Metro line in New York City sparked intense discussion among commuters. In this study, we use Twitter data from 01/2016 – 04/2019, geolocated in New York City, to investigate the sentiment in the population towards topics related to different aspects of the shutdown. The results indicate the strongest sentiments towards alternative travel modes and the effects caused by the shutdown. We further show how the sentiments differ in their spatial clustering characteristics. Tweets conveying a negative sentiment toward the L train tend to cluster in lower Manhattan whereas positive and neutral hot spots are slightly less intense and spread out more evenly.

Keywords. transportation, sentiment analysis, twitter, urban planning, traffic disruption, planned events

Published in "Adjunct Proceedings of the 15th International Conference on Location Based Services (LBS 2019)", edited by Georg Gartner and Haosheng Huang, LBS 2019, 11–13 November 2019, Vienna, Austria.

This contribution underwent double-blind peer review based on the paper. https://doi.org/10.34726/lbs2019.47 |

1. Introduction

The New York City Subway serves as a means of transport for almost 5.48 million daily passengers, up to 420,000 of which use the L train on BMT Canarsie Line[1], which connects the boroughs of Manhattan and Brooklyn with a tunnel crossing the East River. After being flooded and damaged in 2012, the tunnel was only partially renovated. In 2016, the operating agency proposed two scenarios for a full renovation, starting in April 2019: either a three-year partial closure affecting one of the two tubes at a time, or an 18-month full shutdown. The plans sparked intense discussion among those affected by the shutdown, as it would interfere with their commuting routines by causing side effects like longer travel times or higher costs. For context, table 1 provides an overview over the timeline of events. In this paper, we use geolocated Twitter data from 01/2016 – 04/2019 (n = 29,556,272) from New York City and use natural language processing and spatial hot spot detection methods to assess how the shutdown reflects in the public discourse and sentiment of Twitter users and what the spatial characteristics of these effects are. Due to the Twitter data being obtained from different sources, the total number of tweets varies over time, which explains the overall increase of data counts from late 2018 onward. The insights gained from this study can help decision makers understand the impact of traffic disruptions on the affected population's subjective feelings on a highly granular level. This is not only important for the planning process, but also for information announcement strategies that acknowledge and respect sensitive topics.

Table 1: Timeline of selected events and public announcements regarding the L Train

25.07.2016	MTA decides on full shutdown for 18 months beginning in 2019
24.01.2018 - 14.02.2018	MTA holds open houses to raise awareness of the official L train shutdown mitigation plan
17.03.2017	MTA announces the closure will only last 15 months beginning in 2019
03.01.2019	Gov. Cuomo announces there will be no shutdown
13.02.2019	MTA releases a draft plan for evening and weekend repairs starting 27.04.2019

[1] http://web.mta.info/nyct/facts/ridership/ridership_sub.htm

2. Methods

Text preprocessing: To make the text matching used for the subsequent sentiment analysis consistent and performant, we developed a tailored preprocessing workflow using the built-in text search functionality of the PostgreSQL database management system. In our custom text search configuration, the preprocessed texts do not include stopwords or non-words like numbers or URLs. To further improve the text-matching quality, we applied stemming to eliminate ambiguous word endings of synonymous words. We also manually defined a number of n-grams, so groups of multiple words that are treated as one semantic term, e.g. *real estate* or *New York*.

Message Categorization: Because we aim to understand public discourse, we categorized all message texts based on manually defined semantic groups, each related to an aspect of the shutdown. Each group is made up of several keywords, which we initially chose manually and then refined based on the most frequently used words in each group. The groups are policy, effects, alternatives, actors, destination/purpose of travel, location, and L train. If a message text contains a keyword from a given group, a link between the message and the group is established. This check is performed for all groups, thus allowing $n{:}m$ relationships. Categorizing the data allows us to observe whether public discourse of the shutdown changes over time and what topics are most prevalent. It further allows us to focus parts of the analysis specifically on messages including the L train. All results shown below are derived from tweets belonging to the L train topic (n = 3,348).

Sentiment Analysis: We performed sentiment analysis of the tweets' texts to determine whether a tweet contained a positive, neutral or negative emotion. We matched every word or n-gram in the message texts to the corresponding sentiment value in a sentiment lexicon (Hu and Liu, 2004) and summed up the values of the whole message text. The resulting sentiment score s was interpreted as negative, if $s < -1$, neutral if $|s| \leq 1$ and positive if $s > 1$ (Kovacs-Gyori *et al.*, 2018).

Hot Spot Analysis: We performed a hot spot analysis based on the Getis-Ord G_i^* statistic (Ord and Getis, 1995) for the point coordinates of tweets associated with the L train topic, grouped by sentiment. We chose a p value of 0.05, so only hot spots with $|G_i^*| > 1.96$ were considered in the results. The grid size of the statistical units is based on a heuristic used for square cells (Wong and Lee, 2005) and adapted to yield hexagonal grid cells of equal area with an in-circle radius $r_i = \sqrt{A/(n \cdot \sqrt{3})}$ with A being the area of the study region and n the total number of tweets. This setup allowed us to detect whether there are statistically significant spatial clusters of the sentiment groups in the area of interest and if so, compare them.

3. Results

Different discussion topics vary not only in size, but also in composition of sentiments over time. The groups of highest relative sentiment are *Alternatives* and *Effects* (see figure 1). This means that messages concerning these topics tend to be composed of a more emotion-laden vocabulary than others and gives an indication of a strong opinion towards the topic. The value of this knowledge lies in understanding the topic-sentiment links and acting accordingly. In this case, the responsible agency could respond by showing to the public their efforts to minimize negative effects on commuters and creating or promoting viable alternatives. The result maps of the spatial analysis shown in figure 2 indicate that, as expected, L train related tweets tend to cluster around the L train Metro line. The strongest clustering effect for all three sentiments is in the vicinity of the 14th Street station and negative hot spots are more concentrated in Manhattan. Like the results shown above, the spatial distribution of sentiments is also essential for understanding and acting upon the public opinion towards a topic. The strong clustering of tweets with negative sentiment in lower Manhattan may indicate that a disproportionate amount of commuters are experiencing difficulties in that area, although this conclusion may be skewed by the high tweet frequency in the area.

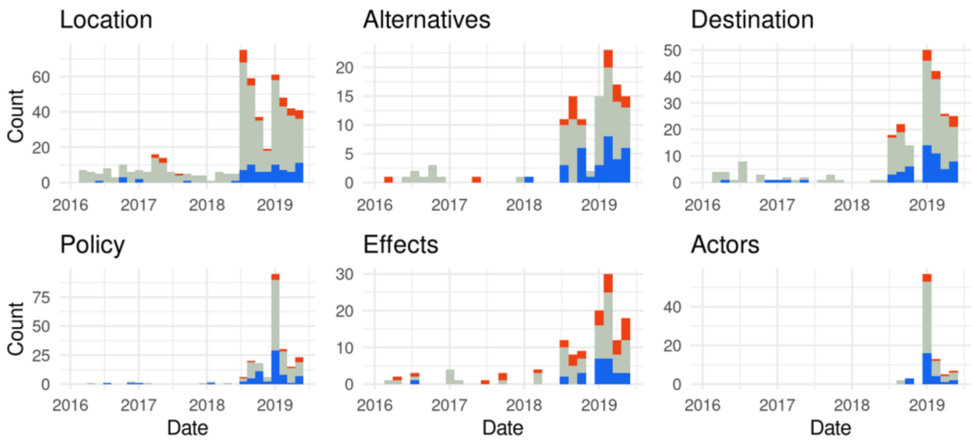

Figure 1: Sentiments in different groups over time within the L train topic (red=negative, gray=neutral, blue=positive). For readability, the graphs use different scales on the y-axis.

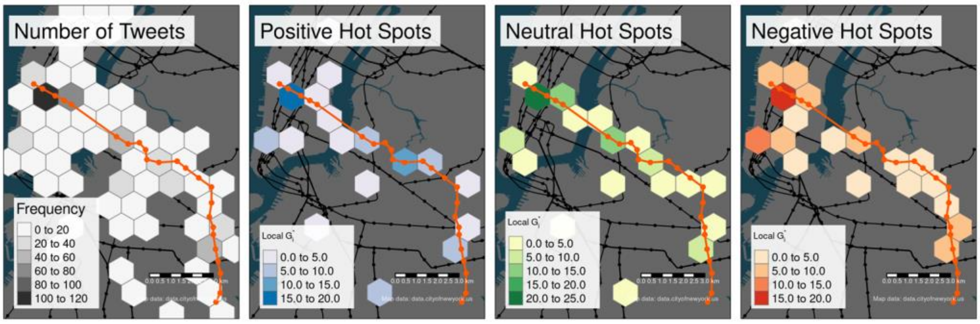

Figure 2: Number of L Train related tweets and distribution of L train topic tweets with positive, neutral, and negative sentiment.

4. Discussion and Conclusion

There are several promising strands of analysis that could be completed in the future. Because, beginning in April 2019, L train service was curtailed on evenings and weekends analysing the content of tweets by time of day will reveal the degree to which the slowdown affects the sentiments of those tweeting. Pairing these data with a spatial analysis would identify particular service pinch points. These could be contrasted with similar data collected on evenings and weekends in July and August 2019, where service is slated to be shut down completely during certain hours. Again, this will help us understand what areas of the city were most affected by the service changes and when allowing us to better model who is using the subway and for what purposes. Contrasting both with construction shutdown dates prior to the slowdown will enable us to gauge how sentiment reacts to (relatively) planned and publicized service changes (such as the current L train slowdown) versus unanticipated closures. Finally, analyzing the content of the tweets may enable researchers to pinpoint what types of adaptation people may have taken in response to the service changes.

However, some caution should be taken in interpreting this data. The segments of the population that take to Twitter to vent their transit-related frustrations (or successes) publicly may not be an accurate sample of L train ridership. Similarly, limiting the analysis to English-language tweets excludes the large Spanish- and other foreign-language speaking population resident along this transit line. Further, the sentiment analysis results still leave room for interpretation. For example, all topics show a bias towards positive emotions, which might seem surprising given the context of traffic disruptions. This might be attributable to the universal language positivity bias shown by (Dodds *et al.*, 2015). Other effects like the relative decrease of responses in the *Alternatives* category can be observed but not causally explained, which is a limitation from working with only one data source. Future work would therefore benefit from integrating additional data sources like questionaries or news articles in the analysis.

References

Dodds, P. S. *et al.* (2015) 'Human language reveals a universal positivity bias', *Proceedings of the National Academy of Sciences of the United States of America*, 112(8), pp. 2389–2394. doi: 10.1073/pnas.1411678112.

Hu, M. and Liu, B. (2004) 'Mining and summarizing customer reviews', in *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*. doi: 10.1145/1014052.1014073.

Kovacs-Gyori, A. *et al.* (2018) '#London2012: Towards Citizen-Contributed Urban Planning Through Sentiment Analysis of Twitter Data', *Urban Planning*, 3(1), p. 75. doi: 10.17645/up.v3i1.1287.

Ord, J. K. and Getis, A. (1995) 'Local Spatial Autocorrelation Statistics: Distributional Issues and an Application', *Geographical Analysis*. doi: 10.1111/j.1538-4632.1995.tb00912.x.

Wong, D. W.-S. and Lee, J. (2005) *Statistical analysis of geographic information with ArcView GIS and ArcGIS*. John Wiley & Sons Hoboken, NJ.