Research Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
1–10

DOI: 10.1177/03611981211037232
journals.sagepub.com/home/trr
SAGE

Twelve-Year Analysis of Transportation Research Board Annual Meeting's Official Hashtag

Subasish Das[1] and Anandi Dutta[2]

Abstract

Twitter, among other social media platforms, has become more popular over time. Social media platforms underpin the way scholars share ideas, propagate the latest emergence of evidence, and adopt new practices, by providing a virtual platform for interacting, socializing, and sharing information from academic conferences with the outside world beyond the physical location. The Transportation Research Board Annual Meeting (TRBAM) is the largest annual conference for transportation engineering and science, and the hashtag for the conference, #TRBAM, was used first in 2009. This paper aims to perform an observational study based on the interactions on Twitter surrounding this hashtag by collecting all original #TRBAM tweets for 12 years (2009–2020). A general trend in the data is that the quantitative measures (tweets, retweets, and favorites) are all much higher during the conference month compared with other months. Top trending topics included: *transit*, *safety*, *bike or non-motorized mode of transportation*, *data*, and *freight*. Overall, the communication pattern shows more dispersion than the central tendency. The findings of this study highlight the need to implement and improve strategies to help transportation research communities encourage continuous and active participation during and after conferences. More active engagement among attendees will help maintain the momentum of information sharing and expand the traffic safety platform globally.

In January 1922, the first annual Transportation Research Board (TRB) meeting was held directly after the National Board of Highway Research was created. The Federal-Aid Road Act of 1916 and the rapid growth of the motor vehicle population in the U.S. created a need to organize and distribute the newly provided national funding for roadway construction. The first annual TRB meeting was held with the purpose of facilitating dialogue among key stakeholders, including universities, highway departments, and highway industries. The meeting aimed to identify highway research needs, correlate research activities, collect and share research findings, and pool resources and knowledge among attendees (*1*). The TRB Annual Meeting (TRBAM) began in 1923 with only 30 attendees, and the number has grown significantly in the decades that followed. TRBAM has now become one of the largest gatherings of its kind in the world for transportation professionals. It consistently draws more than 14,000 attendees and features more than 5,000 presentations, 200 committee meetings, and 800 sessions each year (*1*, *2*).

Twitter is a microblogging social media platform where users can post tweets up to 280 characters (increased from 140 characters in November 2017). Academic conferences widely use conference hashtags to increase interactions, networking, and information sharing. It is generally anticipated that Twitter can facilitate knowledge transfer between conference attendees and conference followers. The official hashtag of TRBAM is #TRBAM. Since its first use in 2009, this hashtag has been used for 12 years. The current study collected all original tweets with #TRBAM through June 30, 2020. Because of the large base of TRBAM participants and the variety of interdisciplinary research amidst the ever-evolving transportation sector, it is important to shed light on the interaction, networking, and knowledge distribution pattern surrounding this hashtag. This is a research gap in understanding the insights of these interactions. This study aims to address this gap by collecting

[1]Texas A&M Transportation Institute, Texas A&M University System, Bryan, TX
[2]Department of Computer Science, University of Texas at San Antonio, San Antonio, TX

Corresponding Author:
Subasish Das, s-das@tti.tamu.edu

the most recent tweets by performing comprehensive exploratory text mining.

Literature Review

The literature review addresses three specific areas: (1) use of Twitter in knowledge sharing from conferences; (2) effectiveness of Twitter as a tool to understand knowledge sharing; and (3) assessment of research trends by performing text mining on conference information.

As one of the largest conferences of its kind, TRBAM provides attendees the opportunity to interact with other transportation professionals across the globe by providing unique educational and communication opportunities. These opportunities have become more available with the incorporation of microblogging via Twitter and communication spanning significantly beyond conference attendees into the broader mainstream. Studies examining the use of Twitter at conferences have found increased use of the social media tool over time, which can, in turn, provide vast amounts of information about knowledge and information sharing beyond the conference environment (*3*, *4*).Compared with other social media avenues like Facebook, Instagram, or Linkedin, Twitter is often used as a unique medium for information sharing. Many studies have specifically explored the use of Twitter as a learning and co-learning medium (*5–9*).

Ebner and Reinhardt conducted a hashtag analysis during the ED-MEDIA 2008 conference and found that micro-blogging not only enhanced participation of conference attendees but also allowed people to share resources and communicate with peers who were not in physical attendance (*10*). Aramo-Immonen et al. argued that Twitter provides participants opportunities to expand co-learning behavior by actively engaging in conversation in an informal environment and become involved in developing processes by sharing participants' expertise and knowledge (*11*). How these groups interact, and their level of influence, can also vary as some groups may have more influence than others (*12*).

Studies measuring Twitter use at different stages in a conference (before, during, and after) are limited. Reinhardt et al. conducted a survey on conference attendees concerning their levels of activity during different periods of time ($n = 41$) and found a majority of users tweeted between 11 and 20 messages per day during the conference to share resources, parallel discussions, take notes, and raise questions (*13*). Pre-conference activities typically related to event organization and promotion, while after-conference events typically involved wrapping up conference attendance, reflections, statistics, or obtaining conference feedback. Das conducted a study on the #TRBAM hashtag by collecting data for 10 years (*14*). The study focused on text mining and communication patterns.

There is a limited number of studies that investigate research trend patterns by performing text mining on conference and journal information. Four studies used various topic modeling techniques to determine research trends by using the TRBAM and Transportation Research Record (TRR) paper titles and abstracts (*15–18*). Another study applied the topic model to discover transportation research themes from 17,163 transportation-related journal articles (*19*). Another study performed similar analysis for maritime transport (*20*). Text mining and topic modeling have become topics of interest among transportation engineers. In recent years, many studies applied these techniques to identify hidden insights from unstructured textual contents (*21–29*).

This study can be considered as a follow-up to Das's study (*14*). The focus of this study is to perform in-depth exploratory data analysis and conduct topic modeling to determine trends and patterns of communication via social media.

Methodology

Data Collection

Twitter has an average of 328 million monthly active users as of the first quarter of 2017 (*30*). A tweet is a post on Twitter. It is a publicly available short statement or note, limited to 280 characters, that is published from a Twitter user's profile. Twitter users choose a username, or "handle," for the platform that is preceded by the @ symbol (e.g., @NASEMTRB is the Twitter handle for the official Twitter account of TRB). Hashtags, prefixed with a # symbol, are used to denote user-specified strings and have been widely used to identify topics of discussion on Twitter. In this study, the authors employed five different open source R software packages as well as several natural language processing (NLP) tools to perform NLP tasks (*31*). Any direct messages exchanged between two users or in a group were not considered for this purpose because that information is not publicly available.

Exploratory Twitter Mining

On Twitter, a person or entity's original post from their handle is known as the "original" Tweet. If someone shares the same content by using the Twitter sharing option, the tweet is called a "retweet." The number of "likes" for a tweet post is the count of favorites for that tweet. Figure 1 shows the frequencies of original tweets, retweets, and favorites associated with #TRBAM and related hashtags, in order from the highest frequency to the lowest frequency. It indicates that the number of favorites has increased significantly in recent years. Here,

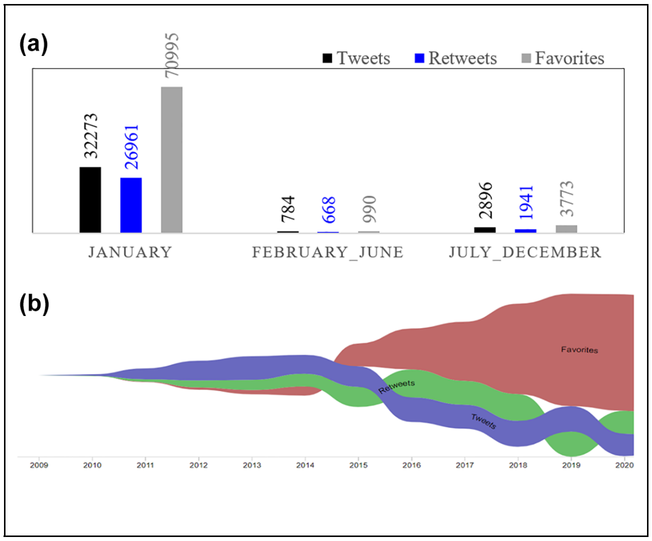

Figure 1. Frequencies of Twitter posts: (*a*) distribution of tweets, retweets, and favorites, and (*b*) bump chart for tweets, retweets, and favorites.

it is worth noting that Twitter adoption has risen significantly and steadily since 2011, and many participants who had never explored Twitter as a learning or networking tool have slowly embraced the platform. This effect, though hard to measure, is undeniably impactful for the conference attendees, with many new participants adding their views and opinions to the topics discussed.

Figure 1*a* splits up the frequencies by three different time frames (during 2009-2020): January, February–June, and July–December. For the periods of February–June and July–December, there was a very minimal number of tweets, retweets, and favorites. This trend is obvious, as these two periods are outside of the month of January when TRBAM takes place in Washington, D.C. From February–June, there were 784 tweets, 668 retweets, and 990 favorites. From July–December, the numbers were slightly higher: there were 2,896 tweets, 1,941 retweets, and 3,773 favorites. As January approached, people typically began to share their session information in advance via social media. In the month of January alone, there were significantly higher frequencies for tweets, retweets, and favorites. In January, there were 32,273 tweets, 26,961 retweets, and 70,995 favorites.

Figure 1*b* shows the bump chart of the distribution of tweets, retweets, and favorites. The quantity of favorites has increased significantly since 2014 compared with tweets and retweets. The number of tweets increased from two in 2009 to 3,256 in 2012 and then remained fairly consistent until 2020, ranging from 3,129 to 4,281. The number of retweets followed a similar pattern. Retweets increased from zero in 2009 to 3,314 in 2015 and then remained fairly consistent until 2020, ranging from 3,129 to 4,616. The number of favorites, however, has significantly increased over the past 12 years. In 2009, there were zero favorites, and in 2020 there was a total of 19,250. Some of the most significant increases in the numbers of favorites were from 2017 to 2018 (difference of 5,181) and from 2018 to 2019 (difference of 3,626).

To understand the temporal distribution of the tweets precisely, it is important to examine the day level tweet analysis. Figures 2 and 3 show the day level number of tweets from 2009 to 2020. The grey color indicates that there were no tweets in a day. The color palette green to orange indicates low to high frequency. The before-conference months (July–December) show more per day tweets compared with after-conference months (February–June). A quick glance at these plots can also show when the physical conference took place each year, based on the color palette measures. It is interesting to see that recent years (2015–2020) show more after-conference (February–June) tweets compared with the after-conference months in the earlier years (2009–2014).

Table 1 shows the other most frequent hashtags used each year from 2010 to 2020. The year 2020 saw the most spread-out range of frequent hashtags including #mobility (n = 56), #equity (n = 32), #research (n = 31), #safety (n = 29), and #transit (n = 28). In 2019, #tsmo (n = 92), #thisisits (n = 79), #mobility (n = 57), and #roadsafety (n = 54) were the most frequently used hashtags. The year 2018 had the greatest overall number of hashtags, and there were several hashtags with relatively high frequencies, including #iot (n = 91), #moveequity (n = 82), #roadsafety (n = 77), #6minpitch (n = 76), and #visionzero (n = 61). The hashtag #transit was the most consistently used hashtag, from 2012 to 2020, and the frequency of #transit ranged from 16 to 46. The hashtag #freight was used in 2011 (n = 21), 2012 (n = 25), 2013 (n = 50), and 2014 (n = 39), and the frequency then decreased in future years before peaking again in 2019 (n = 17).

Clusters of Keywords, Twitter Handles

Figure 4 illustrates the occurrence of Twitter-related keywords associated with #TRBAM. The heat map for tweets illustrates words that were commonly used in tweets. A total of 73,730 unique keywords have been used associated with #TRBAM tweets. After fixing the minimum counts at 10, a total of 1,248 keywords were used to generate the heatmap shown in Figure 4. These words include *booth* (n = 828), *car* (n = 457), *pedestrian* (n = 380), *transportation research* (n = 209), and *congratulations* (n = 186). The use of the keyword *booth* indicates that conference participants frequently post about the location of booths in a way to meet in person with a

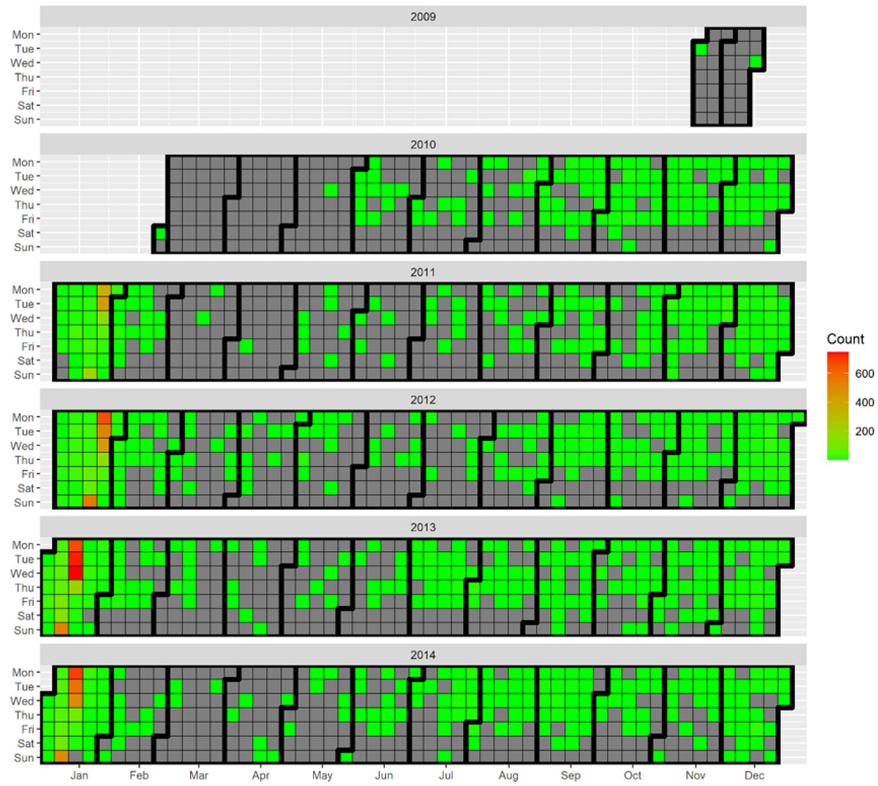

Figure 2. Calendar plots of tweet frequencies between 2009 and 2014.

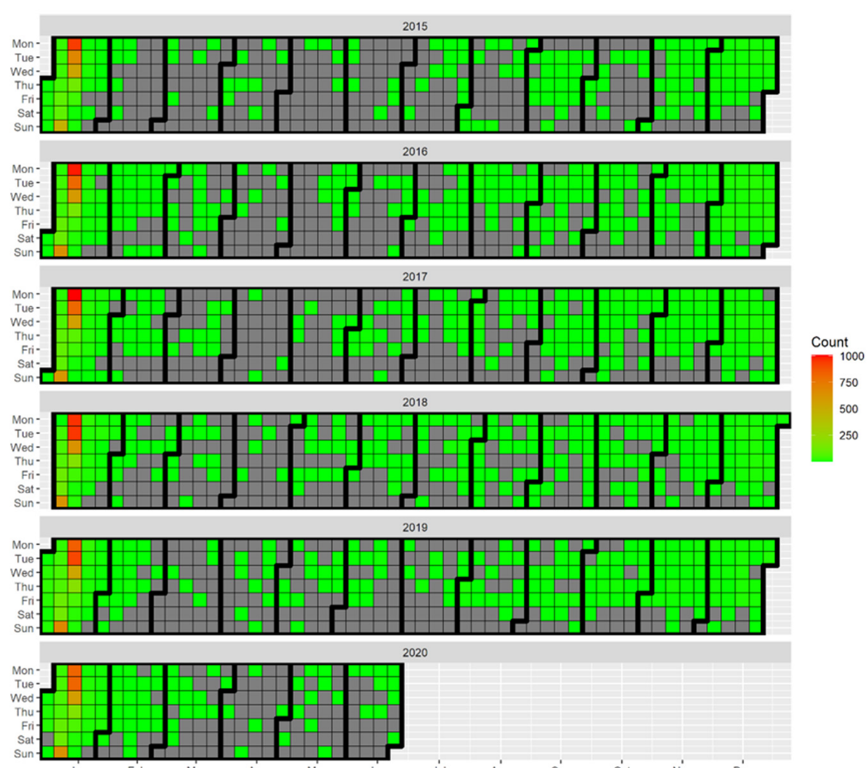

Figure 3. Calendar plots of tweet frequencies between 2015 and 2020.

Table 1. Other Popular Hashtags by Year

Hashtags	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
#transit	3	6	39	27	46	30	17	16	41	39	28
#roadsafety	0	0	0	5	17	17	22	14	77	54	16
#visionzero	0	0	0	0	2	10	25	47	61	37	23
#freight	0	21	25	50	39	12	6	2	6	17	9
#mobility	0	0	0	1	0	5	8	17	40	57	56
#tsmo	0	0	0	0	0	0	1	9	55	92	7
#safety	0	4	4	12	7	12	20	8	23	26	29
#thisisits	0	0	0	0	0	0	6	6	28	79	20
#research	0	0	1	10	12	16	5	12	17	32	31
#6minpitch	0	0	0	2	0	1	13	11	76	3	22
#av040	0	0	0	5	42	62	2	0	0	0	0
#moveequity	0	0	0	0	0	0	0	0	82	21	0
#data	0	0	1	6	2	2	10	12	17	29	22
#iot	0	0	0	0	0	0	1	6	91	1	1
#infrastructure	0	2	5	4	3	5	3	4	20	34	18
#autonomousvehicles	0	0	0	0	1	2	0	17	33	28	15
#equity	0	0	0	4	1	1	10	18	17	11	32
#bigdata	0	0	0	8	4	9	8	12	30	8	5
#bikeshare	0	2	3	8	3	0	10	12	28	12	4
#aviation	0	0	4	2	8	4	12	3	8	13	18

broader group. Some of the key clustered keyword groups are "*booth*, *networking*, *fair*, *organization*," "*convention center*, *noon*, *connect*, *mariott marquis*," "*my1sttrb*, *yptvoice*," and "*car*, *fair*, *pedestrian*, *cycling*, *cyclist*." These clusters display the conventional communication topics during the meeting. However, a more in-depth analysis of the keyword associations is needed, which is done in the next section. All of these combinations can be explained and justified. However, the current scope of the paper is limited to develop a generalized analysis of the overall insights. An explanation of each combination group is outside the scope of the current analysis.

"Twitter handles" represent the people or entity, and the connections between Twitter users are "followers." The current communication pattern defined in this paper is based on the standard criteria that involved the use of #TRBAM and mentioned someone in the same tweet. The purpose was to identify the influential handles in the Twitter network and their individual influence zones; doing so allowed for quantifying an account's influence. The final dataset revealed 8,000 unique interactions in the form of comments or mentions. Single interactions are around 85% of all interactions. Figure 5 shows the clusters developed based on the patterns of mentions of the Twitter handles. A total of 12,817 mentions were found in the collected dataset. After determining the threshold of five mentions, 853 handles were visible in Figure 5. Some frequently mentioned handles are *mamakoid* (n = 474), *transportgooru* (n = 442), *phxdowntowner* (n = 360), *trecpdx* (n = 343), *kittelson* (n = 337), and *kklevine* (n = 316). Figure 5 clearly shows the networking is very closed and within groups.

Latent Dirichlet Allocation

This study used Latent Dirichlet allocation (LDA), a widely popular method to extract hidden insights from unstructured text, to perform topic modeling (*32*). The basic idea is that a generative process derives each document in the corpus (i.e., a collection of documents or texts) in which each document contains a finite distribution of topics. Each topic is considered as a multinomial distribution of keywords in the corpus. The latent elements in the LDA algorithm are the distributions of topics per document and the distribution of words per topic.

Consider, there are M documents in which w indicates a vector of N words in document i. Again, consider topic z is allocated to each word w_j of a document i. It makes each document a combination of topics (represented by some topic distribution θ over the document i). There are two performance measures:

- **α values:** Large value indicates that each document has a comparatively even distribution of the topics. If not, it indicates a sparse distribution of all topics.
- **β values:** High value represents if topics are a comparatively even distribution of the lexicon of words. If not, it indicates a sparse distribution of words per topic.

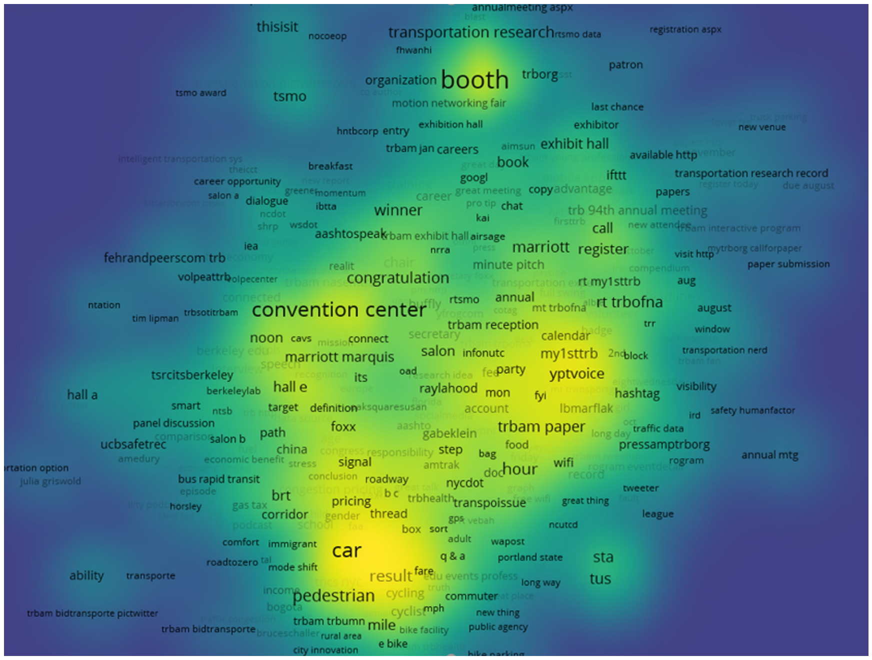

Figure 4. Clusters of keywords.
Note: The larger words and words in brighter yellow indicate a greater cooccurrence in tweets.

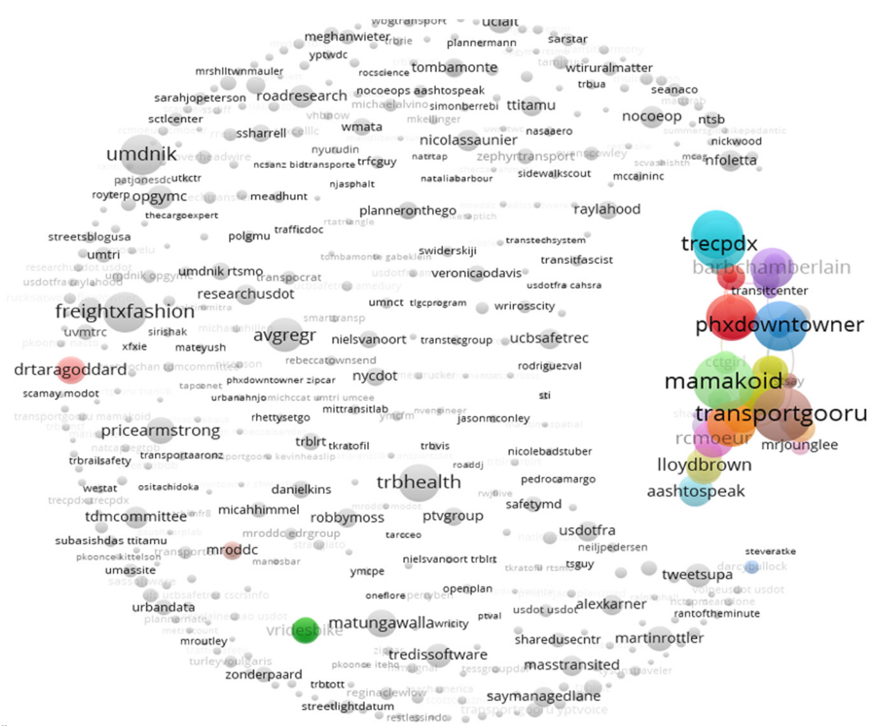

Figure 5. Clusters of Twitter mentions.
Note: The larger the size of the word indicates a higher frequency of use. The networking patterns of the handles with larger circle size also indicates that the conversation and discussion are limited in between them. The colors of the circles indicate different clusters.

One potential drawback of LDA is the difficulty of interpreting some of the topics (*33*). Sievert and Shirley developed a relevancy metric to resolve this issue (*34*). This metric can be shown as:

$$rel(term\ w|topic\ t) = \lambda * p(w|t) + (1 - \lambda) * p(w|t)/p(w) \tag{1}$$

The weighting parameter λ can influence the word ranking per topic based on their relevance. Before starting the topic modeling technique, several basic steps of data cleaning were performed. Stop words, redundant words, numbers, and punctuations were removed initially. Lemmatization uses vocabularies and morphological analyses to remove the inflectional endings of a word and to convert it into its dictionary form. For the final dataset for topic modeling, lemmatization was conducted.

Figure 6 shows topic models in word cloud format based on different temporal durations. For each of the temporal clusters, top keywords (range: 30–50) have been used to generate four topic clusters from different temporal durations. A comprehensive list of stop words (for example: *research*, *paper*, *papers*, *session*) was excluded from the analysis because of the non-information nature of these words.

Topic Models Based on Year-Based Corpora. The topic models (shown in word cloud format) based on 2009–2012 tweets show topics such as *involvement in committees* (Topic 1), *bike and transit* (Topic 3), *receptions in the TRB meetings* (Topic 2), and *parking and bikeshare* (Topic 4). Some of the intuitive keywords in several topics are *health* (Topic 1) and *safety* (Topic 4). The top four topics in 2013–2016 tweets are *freight/rail* (Topic 1), *bike/data* (Topic 2),

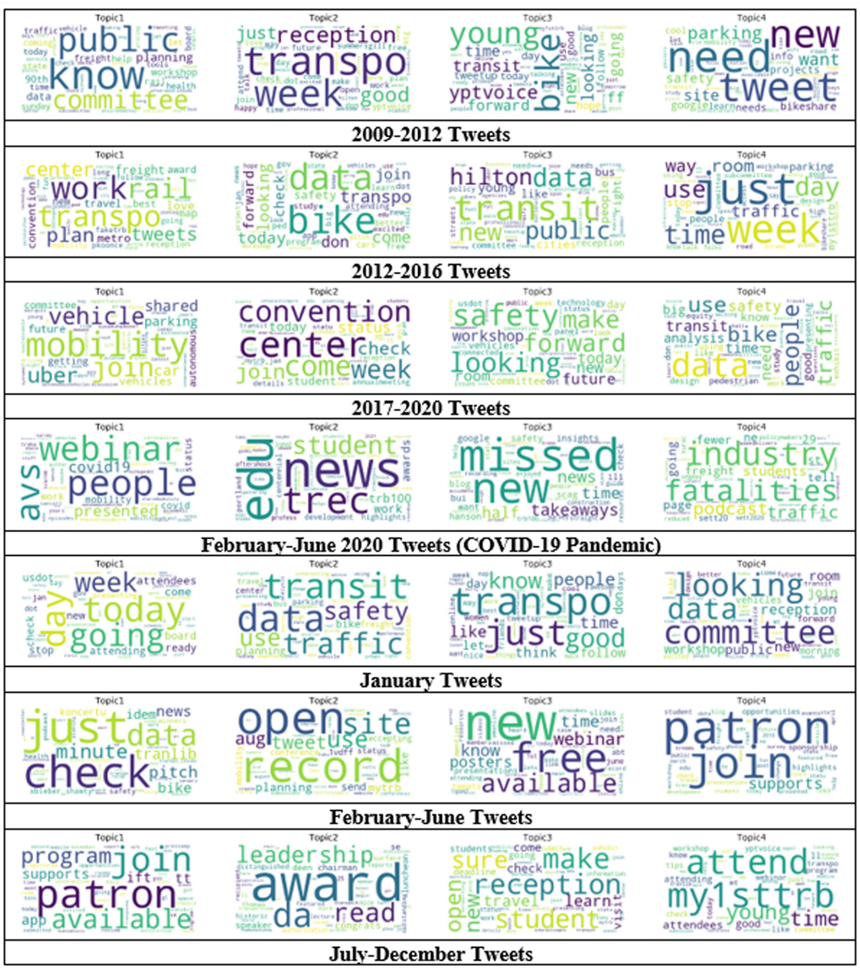

Figure 6. Topic models based on different temporal clusters.
Note: The size of the word indicates the higher frequency of the words. A specific color palette (viridis) is used for ease of visual interpretation.

public transit (Topic 3), and *parking* (Topic 4). Topic 2 (*bike/data*) also has several relevant words, such as *vehicles*, *safety*, and *app*. Topic 3 is also associated with city science topics such as *cities*, *buses*, *needs*, and *policies*. *Mobility* (Topic 1), *convention center* (Topic 2), *safety* (Topic 3), and *bike/transit* (Topic 4) are the four major topics in 2017–2020 tweets. Some of the topics included in Topic 1 are *mobility*, *ridesharing services* (for example, *Uber*), *future transportation*, and *autonomous vehicles*. Topic 2 broadly discusses the new annual meeting venue *Convention Center*. Topic 3 (*safety*) also contains relevant keywords such as *technology*, *vehicles*, *future*, and *connected*. Topic 4 (*bike/transit*) broadly covers relevant issues such as *pedestrians*, *safety*, *traffic*, and *need*. One general finding is that *bike/transit* is the common topic present in all three clusters of topics. The dominant topics during 2017–2020 are *safety* and *mobility*. New and emerging transportation technologies such as *connected and autonomous vehicles* and *ridesharing services* are present in the recent year (2017–2020) tweet corpora (collection of tweets or documents).

Topic Models Based on Month-Based Corpora. Three month-based corpora (January, February–June, and July–December) have been developed to perform the topic modeling separately for each of the month-based corpora. Additionally, the COVID-19 (February 2020–June 2020) corpus has been developed to see the most recent trends of transportation-related discussion in the era of a pandemic. It is interesting that COVID-19-related discussions are not present in the topics. It is mostly because of the number of #TRBAM-related tweets during the April–June period. From Figure 3, it is found that most of the tweets occurred during February–March, and COVID-19 discussion was limited. Additional investigation on this issue has not been done here as the broader aspects of "social media and transportation during COVID-19" could be a separate research topic.

January tweets have four major topics: *attendance in TRB sessions* (Topic 1), *transit and safety* (Topic 2), *networking* (Topic 3), and *committees and receptions* (Topic 4). February–June tweets have topics such as *paper submission/site opening* (Topic 2), *posters/presentations/webinars* (Topic 3), and *patron/opportunities* (Topic 4). July–December tweets have four major topics: *patron/availability* (Topic 1), *leadership award* (Topic 2), *reception at TB meeting* (Topic 3), and *first TRBAM attendance* (Topic 4). As mentioned earlier, during-TRBAM (January) and after-TRBAM (other months) have significantly different tweeting patterns in relation to quantity and topics. Monthly clustering shows that January tweets are mostly associated with physical-meeting-related keywords. For other months, the general trend is to discuss the papers/presentations and potential opportunities at the meeting.

Conclusion

With a flexible and ever-increasing network of users, Twitter offers transportation researchers promising possibilities concerning general information updates and exchanges while also fulfilling the demand for fast and immediate modes of communication. Twitter mining was used in several aspects of transportation engineering studies (*35–48*). Past studies of the use of conference hashtags indicate that this is a dynamic space where attendees share their thoughts, experiences, and ideas through the use of social media hashtags. This study reflects on the use of the hashtag #TRBAM over a period of 12 years and examines the associated communication patterns and trends of topics. This study started the analysis by performing exploratory Twitter mining. The topic models for different temporal durations indicate how topics evolve over time. The study found that January tweets (especially during the conference dates) are significantly higher than the rest of the months because of the notion of live tweeting during the conference. There is need for added promotion for the continuation of this hashtag throughout the year. It is usually seen that discussions on TRB- or TRB-related committees usually mention or use hashtag #NASEMTRB instead of #TRBAM, if the discussions or interactions do not occur in January. Future researchers should extend the range of hashtags to capture the discussions throughout the year.

The traditional means of distributing knowledge through TRBAM is not enough; a larger social media presence can reach a much broader audience base. The current study examines the content and co-learning environment on social media. This study has several limitations. The cluster analysis requires additional examination of the patterns of interactions associated with lower frequencies. Impression analysis on the posted tweets provides additional information on the outreach of these tweets. This study has limited resources to reach out to the handles of the popular tweets to acquire that information. With the mass presence of mobile devices and the wide adoption of social networking platforms among users of all ages in the transportation profession, such platforms will certainly play a critical role in adapting and growing the conference learning experience in the years to come. Future research should aim to extend the study to encompass a larger set of transportation engineering conferences and compare the findings to the current study.

Author Contributions

The authors confirm the contribution to the paper as follows: study conception and design: S. Das; data collection: S. Das, A. Dutta; analysis and interpretation of results: S. Das; draft manuscript preparation: S. Das, A. Dutta. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Subasish Das https://orcid.org/0000-0002-1671-2753

References

1. Byrd, L. Transportation Research Board 1920-1995: The First Seventy-Five Years. The Highway Research Board. *TR News-Transportation Research Board*, Vol. 182, 1996, pp. 5–20.
2. Transportation Research Board, 2020. TRB 2020 Annual Meeting Highlights: A Century of Progress. TR News, Jan-Feb 2020 Issue, pp. 3-14. http://onlinepubs.trb.org/onlinepubs/trnews/trnews326CenturyofProgress.pdf. Accessed August 1, 2020.
3. Wen, X., Y. R. Lin, C. Trattner, and D. Parra. Twitter in Academic Conferences: Usage, Networking and Participation Over Time. Proc., 25th ACM Conference on Hypertext and Social Media, Santiago, Chile, 2014, pp. 285–290.
4. Chaoquin, N., and J. Jiang. Visualizing Computer Science Communities Using Conference Hashtags. *Proc., of the 2016 iConference*, (poster), Champaign, IL, 2016
5. Erickson, I. The Translucence of Twitter. In *Ethnographic Praxis in Industry Conference Proceedings*, Vol. 2008, Blackwell Publishing Ltd, Oxford, 2008, pp. 64–78.
6. Huberman, B. A., D. M. Romero, and F. Wu. Social Networks That Matter: Twitter under the Microscope. *arXiv:0812.1045*, 2008.
7. Gruzd, A., B. Wellman, and Y. Takhteyev. Imagining Twitter as an Imagined Community. *American Behavioral Scientist*, Vol. 55, No. 10, 2011, pp. 1294–1318.
8. Zappavigna, M. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. Continuum International Publishing Group, Great Britain, 2012.
9. Loureiro-Koechlin, C., and T. Butcher. The Emergence of Converging Communities via Twitter. *The Journal of Community Informatics*, Vol. 9, No. 3, 2013.
10. Ebner, M., and W. Reinhardt. Social Networking in Scientific Conferences–Twitter as Tool for Strengthen a Scientific Community. *Proc., 1st International Workshop on Science*, Vol. 2, Nice, France, 2009, pp. 1–8.
11. Aramo-Immonen, H., J. Jussila, and J. Huhtamäki. Exploring Co-Learning Behavior of Conference Participants With Visual Network Analysis of Twitter Data. *Computers in Human Behavior*, Vol. 51, 2015, pp. 1154–1162.
12. Wen, X., D. Parra, and C. Trattner. How Groups of People Interact with Each Other on Twitter During Academic Conferences. *Proc., of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Baltimore, MD, 2014, pp. 253–256.
13. Reinhardt, W., M. Ebner, G. Beham, and C. Costa. How People Are Using Twitter During Conferences. Creativity and Innovation Competencies on the Web. *Proc., 5th EduMedia*, Salzburg, Austria, 2009, pp. 145–156.
14. Das, S. #TRBAM: Social Media Interactions From Transportation's Largest Conference. TR News, Vol. 324, 2019, pp. 18–23.
15. Das, S., X. Sun, and A. Dutta. Text Mining and Topic Modeling of Compendiums of Papers From Transportation Research Board Annual Meetings. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2552: 48–56.
16. Das, S., K. Dixon, X. Sun, A. Dutta, and M. Zupancich. Trends in Transportation Research. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2614: 27–38.
17. Das, S. Data Dive Into Transportation Research Record Articles: Authors, Coauthorships, and Research Trends. TR News, Vol. 331, 2021, pp. 25–31.
18. Boyer, R. C., W. T. Scherer, and M. C. Smith. Trends Over Two Decades of Transportation Research. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2614: 1–9.
19. Bai, X., X. Zhang, K. X. Li, Y. Zhou, and K. F. Yuen. Research Topics and Trends in the Maritime Transport: A Structural Topic Model. *Transport Policy*, Vol. 102, 2021, pp. 11–24.
20. Sun, L., and Y. Yin. Discovering Themes and Trends in Transportation Research Using Topic Modeling. *Transportation Research Part C Emerging Technologies*, Vol. 77, 2017, pp. 49–66.
21. Das, S., and M. Theel. Pandemic and Transportation Research: Bibliometric Analysis and Topic Modeling. Presented at the 100th Transportation Research Board Annual Meeting, virtual, 2021.
22. Das, S. Exploratory Analysis of Unmanned Aircraft Sightings Using Text Mining. *Transportation Research Record: Journal of the Transportation Research Board*, Vol 2675, 2021, pp. 291–300.
23. Syed, A., X. Qin, R. Kate, D. Anisuzzaman, and Z. Yu. Identification and Analysis of Misclassified Work-Zone Crashes Using Text Mining Techniques. *Accident Analysis & Prevention*, Vol. 159, 2021, p. 106211.
24. Das, S. Fatal Crash Reporting in Media: A Case Study on Bangladesh. *Transportation Research Record: Journal of Transportation Research Board*, 2021.
25. Arteaga, C., A. Paz, and J. Park. Injury Severity on Traffic Crashes: A Text Mining With an Interpretable Machine-

Learning Approach. *Safety Science*, Vol. 132, 2020, p. 104988.
26. Hosseini, P., P. Jalayer, M. Das, and S. Zhou, H. Identifying Wrong-Way Driving (WWD) Crashes in Police Reports Using Text Mining Techniques. Presented at the 100th Transportation Research Board Annual Meeting, virtual, 2021.
27. Das, S., A. Dutta, and I. Tsapakis. Topic Models From Crash Narrative Reports of Motorcycle Crash Causation Study. *Transportation Research Record: Journal of Transportation Research Board*, 2021, p. 036119812110025. https://doi.org/10.1177/03611981211002523.
28. Rahman, M. T., K. Dey, S. Das, and M. Sherfinski. Sharing the Road With Autonomous Vehicles: A Qualitative Analysis of the Perceptions of Pedestrians and Bicyclists. *Transportation Research Part F Traffic Psychology and Behaviour*, Vol. 78, 2021, pp. 433–445.
29. Das, S., and H. Zubaidi. Last Forty Years of ITE Journal Articles: A Scientometric Overview. ITE Journal Online Exclusives. https://www.ite.org/publications/ite-journal/last-forty-years-of-ite-journal-articles-a-scientometric-overview/. Accessed July 11, 2021.
30. Statista. Number of Monthly Active Twitter Users Worldwide From 1st Quarter 2010 to 1st Quarter 2019. https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/. Accessed August 1, 2020.
31. The R Project for Statistical Computing. https://www.r-project.org/. Accessed July 11, 2021.
32. Blei, D. M., A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, Vol. 3, 2003, pp. 993–1022.
33. Chang, J., S. Gerrish, C. Wang, J. L. Boyd-Graber, and D. M. Blei. Reading Tea Leaves: How Humans Interpret Topic Models. Presented at the Advances in Neural Information Processing Systems, Vancouver, British Columbia, Canada, 2009.
34. Sievert, C., and K. Shirley. LDAvis: A Method for Visualizing and Interpreting Topics. *Proc., of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, Baltimore, Maryland, 2014, pp. 63-70.
35. Das, S., X. Sun, and A. Dutta. Investigating User Ridership Sentiments for Bike Sharing Programs. *Journal of Transportation Technologies*, Vol. 5, 2015, pp. 69–75.
36. Sezgen, E., K. J. Mason, and R. Mayer. Voice of Airline Passenger: A Text Mining Approach to Understand Customer Satisfaction. *Journal of Air Transport Management*, Vol. 77, 2019, pp. 65–74.
37. Das, S., X. Sun, A. Dutta, and M. Zupancich. Twitter in Circulating Transportation Information: A Case Study on Two Cities. The 96th Transportation Research Board Annual Meeting, Washington D.C., January 8–12, 2017.
38. Sun, X., and S. Das. *User Sentiment Analysis With Louisiana Social Media Data for Better and Effective Crash Countermeasures*. Report No. 14-4TIRE. Louisiana Transportation Research Center, Baton Rouge, LA, 2015.
39. Gu, Y., Z. S. Qian, and F. Chen. From Twitter to Detector: Real-Time Traffic Incident Detection Using Social Media Data. *Transportation Research Part C Emerging Technologies*, Vol. 67, 2016, pp. 321–342.
40. Das, S. *Evaluating Safety Improvement From Edge Lines on Rural Two-Lane Highways*. Doctoral dissertation, Department of Civil Engineering, University of Louisiana at Lafyette, 2015.
41. Das, S., and H. A. Zubaidi. City Transit Rider Tweets: Understanding Sentiments and Politeness. *Journal of Urban Technology*, 2021, pp. 1–16. https://doi.org/10.1080/10630732.2021.1903288
42. Chen, F., and R. Krishnan. *Transportation Sentiment Analysis for Safety Enhancement*. *Technologies for Safe and Efficient Transportation*. U.S. DOT University Transportation Center, Washington D.C., 2013.
43. Dutta, A., and S. Das. Tweets About Self Driving Cars: Deep Sentiment Analysis Using Long Short-Term Memory Network (LSTM). *Proc., International Conference on Innovative Computing and Communications, Part of the Advances in Intelligent Systems and Computing Book Series (AISC)*, Vol. 1165, 2020, pp. 515–523 (virtual).
44. Das, S., and A. Dutta. Characterizing Public Emotions and Sentiments in COVID-19 Environment: A Case Study of India. *Journal of Human Behavior in the Social Environment*, Vol. 31, No. 1–4, 2021, pp. 154–167.
45. Das, S., A. Dutta, G. Medina, L. Minjares-Kyle, and Z. Elgart. Extracting Patterns From Twitter to Promote Biking. *IATSS Research*, Vol. 43, 2019, pp. 51–59.
46. Yazici, M. A., S. Mudigonda, and C. Kamga. Incident Detection Through Twitter. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2643: 121–128.
47. Das, S., A. Dutta, A. Mudgal, and S. Datta. Non-fear-Based Road Safety Campaign as a Community Service: Contexts From Social Media. In *Innovations for Community Services. I4CS 2020. Communications in Computer and Information Science, Vol. 1139* (S. Rautaray, G. Eichler, C. Erfurth, and G. Fahrnberger, eds.), Springer, New Mexico, 2020.
48. Zhang, Z., M. Ni, Q. He, J. Gao, J. Gou, and X. Li. Exploratory Study on Correlation Between Twitter Concentration and Traffic Surges. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2553: 90–98.