VISUAL INTEGRATION OF SOCIAL MEDIA ANALYSIS AND TRAFFIC ANALYSIS [SURVEY]

Prof. N. R. Sonawane [1],Nalini Somwanshi [2],Priyanka Palange [3],Pooja Fuse [4],Shradha Patil [5]

[1] *Prof., Information Technology, PVG's COET, Maharashtra, India*
[2] *Student, Information Technology, PVG's COET, Maharashtra, India*
[3] *Student, Information Technology, PVG's COET, Maharashtra, India*
[4] *Student, Information Technology, PVG's COET, Maharashtra, India*

ABSTRACT

Transportation frameworks in urban areas are regularly influenced by different sorts of occasions, for example, common calamities, mishaps, furthermore, open social affairs. Exceedingly thick and convoluted systems in the transportation frameworks engender disarray in the system since they offer different conceivable exchange courses to travelers. Perception is a standout amongst the most vital strategies for looking at such falls of irregular circumstances in the immense systems. This paper proposes visual mix of movement investigation and online networking examination utilizing two types of huge information: brilliant card information on the Tokyo Metro and web-based social networking information on Twitter. Our framework gives numerous organized perspectives to outwardly, naturally, and at the same time investigate changes in travelers' conduct and anomalous circumstances separated from keen card information and situational clarifications from genuine voices of travelers, for example, grievances about administrations separated from web-based social networking information. We exhibit the potential outcomes and value of our novel representation environment utilizing a arrangement of genuine information contextual analyses and area specialists' criticisms about different sorts of occasions.

Keyword : - *Smart card, Big data*

1. INTRODUCTION

Open transportation frameworks, for example, railroads and metros, in uber urban communities are constantly required to build their strength to outrageous circumstances created by different occasions. For example, Tokyo, which is the greatest uber city in Japan, will have the 2020 Summer Olympics and Paralympics, which will bring about substantial scale developments of individuals over the wide region around Tokyo. Intense inland tremors are additionally assessed to conceivably happen in the Tokyo metropolitan region. Open transportation frameworks

are presently get ready reactions for these occasions .To expand the strength of the frameworks, lessons must be gained from past occasions to see how the frameworks are influenced by changes in travelers' practices. Incorporation of shrewd card information and web-based social networking information empowers us to replay past occasions and to find irregular circumstances of transportation frameworks, spreads of variations from the norm over transportation systems, and travelers' protests or disappointment about which even prepare framework administrators and station staff don't have a clue. While some investigation frameworks have used both versatility information and online networking information to comprehend human conduct or movement peculiarities, they can't bolster both finding strange circumstances from wide spatio-fleeting space and investigating spatio-transient spread of them, furthermore, intelligently investigating their reasons in detail. Creating a visual domain for investigating traveler practices in a perplexing transportation framework utilizing transportation logs and web-based social networking stream is still a testing assignment. For supporting powerful investigation, the earth needs to fulfill the accompanying prerequisites:

1] Finding uncommon marvels from the wide   scope of fleeting reviews that are gotten from contrasts amongst every day and occasion driven traveler practices. The methods for naturally confirming   impacts of known occasions and finding inconvenience obscure to try and prepare framework administrators are wanted.

2] Understanding changes in traveler streams and spatial spread of uncommon wonders in every time period on a wide region metro organize. A visual investigation environment is important to instinctively get it the course, speed, and scope of engendering of the surprising marvels, for example, anomalous crowdedness. These are troublesome for the prepare framework administrators to comprehend in light of the fact that the transportation framework arrange in Tokyo is to a great degree thick and convoluted.

3] Investigating explanations behind strange wonders or their impacts from genuine clients' voices. A framework is required for investigating data about travelers' dissensions, exercises, for example, utilization of cabs or transports, and confounding circumstances in stations, which frequently can't be acquired from client support or operation inconvenience databases.

This paper proposes a novel visual combination investigation framework that can bolster ex post assessments of inconvenience in a metro framework by utilizing two types of huge information: documented transportation logs from the savvy card arrangement of the Tokyo Metro and online networking information from Twitter. Information procured through the pictured outcomes generally reflects genuine circumstances, for example, fiascos, mischances, and open social affairs. To address the above prerequisites, we constructed the investigation framework by coordinating the accompanying representation segments:

1. HeatMap see gives a fleeting outline of uncommon marvels in traveler streams.
2. AnimatedRibbon see pictures fleeting changes in traveler streams with spatial settings and spread of uncommon wonders over the entire metro organize utilizing activity.
3. Tweet Bubble see gives an outline of patterns of catchphrases clarifying the circumstance amid the irregular wonders. outline, we have made the accompanying commitments.

2. RELATED WORK

2.1 Shrewd Card Information Examination

Smart card data is one of the data sources to analyze operation of public transportation systems [3], [4]. Ceapa et al. focused on congestion patterns of some underground stations in London to reveal station crowding patterns to avoid traffic crowdedness [5]. They utilized data of oyster cards, the smartcards used on the London Underground. Their spatio-temporal analysis showed a highly regular crowding pattern during the weekdays with large spikes occurring in short time intervals. Sun et al. provided a model to predict spatio-temporal density of passengers and analyzed it for one MRT line in Singapore [6]. However, previous work only focused on a single selected line or some stations. One reason is that most smart card data does not include transfer station information. Our work speculates the most probable path of each trip from origin and destination in smart card data and succeeds in visualizing propagation of effects of trouble on the metro network.

Zeng et al. [7] gave a perception framework to investigate traveler versatility in a Singapore open transportation frameworks counting Metro framework and open transport arrange. They assessed different versatility related elements, for example, holding up time, riding time, exchange time, and travel effectiveness utilizing information including traveler travel information through RFID card, travel line plan information, and transportation organize information. They at that point imagined them to investigate land availability, time-productive courses and their worldly varieties along the starting point goal ventures. In spite of the fact that they concentrated on envisioning versatility related data along courses from a particular birthplace in a tree structure, our work concentrate on picturing spatio-transient spread of crowdedness or vacancy in a muddled system. To the extent we know, there has been no examination on the perception of engendering of impacts spreading over an extensive variety of open transportation frameworks, for example, metro systems.

2.2 Spatio-Temporal Information Visualization

There has been some examination on and frameworks created for the perception of geo-spatial and worldly values on a delineate. Andrienkos and Slings by et al. used numerous heatmaps in a guide isolated into general rectangles [8], [9]. Each heatmap in the lattice spoke to worldly review of movement volume in a predetermined territory. Wang et al. likewise utilized heatmaps to envision worldly changes in movement speeds in chosen street segments [10]. Bringing 3D symbols into a guide set up of heatmaps is one of the option approaches for speaking to worldly outline of quality values in determined spots [11], [12]. Their approach concentrated on portraying changes in qualities at free focuses or zones and did not give a strategy to speaking to fleeting changes in values between two focuses or streams.

There has been some examination on dissecting portability information what's more, separating and picturing critical occasions or portability designs. Doraiswamy et al. acquainted procedures with concentrate urban occasions from substantial spatio-transient information, for example, taxi trips in New York City1 [13]. Wang et al. removed and envisioned roads turned parking lots and their proliferation from information of taxi trips in Beijing [14]. Andrienko et al. removed and described imperative spots from versatility information, for example, GPS tracks of autos and flight directions and envisioned them in 3D spatio-transient space [15], [16]. Dissimilar to these cases, smart card information in our framework incorporates birthplace goal (OD) information without direction data. We along these lines need to estimate the most plausible course for each trek from OD information and envision collected traveler practices.

Ferreira et al. given a visual situation to looking at fleeting changes in qualities, for example, trip span or number of outings between chose OD locales utilizing taxi trip information in New York City [17]. Flows rates imagined a transient review of stream extents among different OD combines by utilizing heatmap and two separate maps [18]. Jiang et al. presented Roundabout pixel charts to speak to spat io-fleeting examples of OD conveyances from or to a chose area utilizing roundabout heatmaps [19]. In spite of the fact that their approach can picture transient changes in streams among two isolated districts, it can't speak to changes in streams in the particular directions or courses on the guide, for example, lines of trains or streets.
Representation of time-fluctuating changes in the quantity of travelers in a Metro system is a sort of element chart representation. Andrienkos proposed techniques for speaking to summed up developments as system of streams [22], yet they didn't give a technique to picturing fleeting changes in streams. Despite the fact that there has been some exploration on element diagram perception [23], [24], [25], to the extent we know, there has been no examination to have the capacity to at the same time give bits of knowledge into numerous properties of both hubs and edges in a diagram while inserting them in their spatial and transient setting as our Animated Ribbon view does.

2.3 Spatial Tweet Representation

LeadLine [26] recognized occasions from online networking information, removed data around 4 Ws (who, what, when, and where) identified with the occasions, and afterward envisioned the data in composed perspectives. SensePlace2 [27] gave an coordinated condition for separating and picturing space-time topic data from twitter streams. Thom et al. given visual investigation framework to distinguishing spatio-fleeting irregularities from geo-found tweets and envisioning them as word mists portrayal on a guide [28]. Their approaches concentrate on investigating occasions from online networking information without utilizing other information assets.
Skillet et al. given a framework to movement irregularity identification from human versatility information and peculiarity examination utilizing web-based social networking information [1]. They utilized term mists to envision terms identified with the recognized irregularities. In spite of the fact that in their approach perception is utilized for

demonstrating distinguished comes about, our work concentrate on giving intuitive conditions for discovering peculiarities and investigating them in detail by utilizing two types of information from shrewd card framework and online networking.

3.EXISTING SYSTEM

The smart card is being utilized increasingly by open travel offices, this innovation is not new. Since 1990, that the use of the sharp card has ended up being basic, with the exponential development of the web and the expanded modernity of portable correspondence advancements. Indeed, the shrewd card has many utilizations, including access control, and to store data like biometrics, photographs, fingerprints, medicinal information, DNA comes about, religious association, and saving money information.

4.DATASETS

4.1 Smart Card Data

We use a large scale data set of travel records from March 2011 on the Tokyo Metro extracted from the smart card system. Tokyo has the complicated train route map. It consists occasions also, some different days in get-away seasons are dealt with as ends of the week.

Travelers are relied upon to act with some occasional designs, particularly every day ones, along these lines we attempt to do the factual examination of this information. demonstrates the normal and standard deviation of the quantity of travelers at each day and age of the day through one year, from April 2012 to March 2013. The blunder bar of each point demonstrates standard deviation. we first gauge the excursion time length of each trip log and afterward gather the number of travelers who were going at a specific time period. The eras are isolated at regular intervals. Weekdays and ends of the week have plainly unique request designs. The deviations of weekdays are extensively littler than those of ends of the week. This implies generally travelers really act in an intermittent way, so we might be ready to identify some unpredictable mishaps or occasions by looking at the distinctions with the normal number of travelers at each segment..

4.2 Social Media Data

Online networking promptly reflects genuine occasions, for example, mischances. In this paper, we use Twitter as an online networking information asset. We have been creeping through more than four years of Twitter information from Twitter API from Walk 11, 2011. Our creeping began from 30 well known Japanese clients by getting their past courses of events. At that point we over and over extended the arrangement of clients by taking after retweets what's more, notices showed up in their courses of events. We have persistently played out the client extension and following of their courses of events. We then got information of more than 2 million dynamic clients and 25 billion tweets on 2015.

5.WORKING MODEL

Visual Exploration is a process to known events happening in city. In this application we are going to show the events in particular office complex. If complex consist of 10 to 15 stories building in this complex there are many number of office, here we are going to maintain the data of all employees at particular floor for example how many employees are working in single office.

Employee need to login at particular entry gate and in office so admin can calculate the number of employee and visitor present at particular office and on floor. Suppose if fire catches at particular floor e.g. at 10th floor then admin will fire an event which is shown on user application, so they can take another way to dispatch from floor. So our application is useful to find out the how many people are missing from office or complex.

This paper proposes a novel visual combination investigation framework that can bolster ex post assessments of inconvenience in a metro framework by utilizing two types of huge information: chronicled

transportation logs from the keen card arrangement of the Tokyo Metro and online networking information from Twitter.
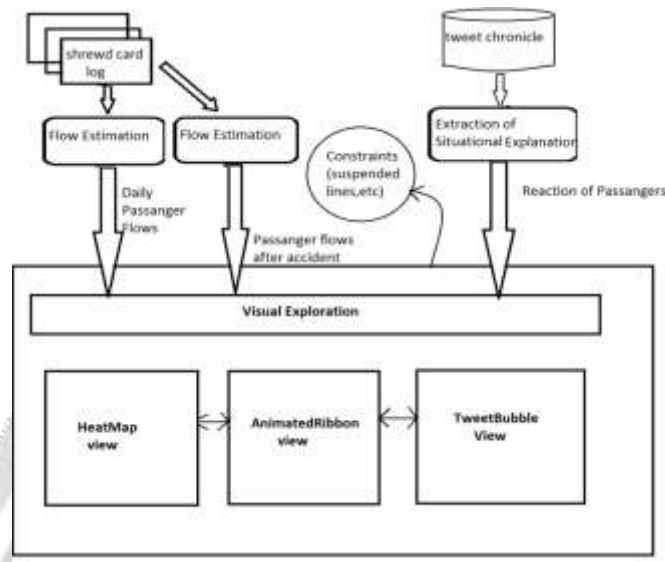

Fig1. Working Model

6. EXTRACTION OF TRAVELER STREAMS

With the recorded brilliant card information, we can see how numerous travelers utilized a specific station. Be that as it may, that information does exclude the passage time, so we couldn't perceive how numerous travelers are there inside a specific day and age. Also, on the off chance that we attempt to appraise the crowdedness of each prepare, or impacts of a mischance at a specific area, the beginning goal combine is lacking. We should make sense of the travel way of every traveler for such demands.

6.1 Assessing Day by day Traveler Streams

There are a few conceivable ways to take from an inception station to a goal station. A keen card log contains data about where a traveler touched in and where and when he/she touched out. It does exclude the passageway time or exchange stations' data. We consequently guess the most plausible way for every outing (starting point and goal combine) by expecting that they take the briefest time way. We accept that aggregate travel time (t) of every trek is characterized as

$$t \frac{1}{4} T \, \flat \, C \, \flat W$$

where:

• T is the time while travelers are riding trains. It characterized by utilizing the timetable.
• C is the strolling time while travelers are exchanging trains. It identifies with the structure of the station, so it varies at each station. We generally characterize these circumstances by utilizing the data from the prepare organization.
• W is the time sitting tight for an exchange. We characterize this as (normal prepare interim/2) separated from the timetable. It contrasts on each line. With this model, we can ascertain the assessed travel time of any travel way. We then scan for the briefest time way of each root goal stations match by utilizing the Dijkstra calculation.

We need to discover surprising wonders that vary from the regular repeating examples of the travelers. For this reason, we first gauge in which segment of a line a traveler passes in a specific day and age from the guessed

briefest time way and leave time. We then gather the quantity of travelers who voyaged a specific segment in a specific time period (like clockwork or 60 minutes). Information for weekdays what's more, information for a considerable length of time are independently broke down on the grounds that weekdays and ends of the week demonstrate obviously extraordinary examples as appeared from that point forward, we compute the basic moving normal (SMA) of the past one year for every month what's more, ascertain standard deviation utilizing a similar time window. SMA reflects every day recurrent examples, and surprising examples can be distinguished by contrasting it and log information. All traveler streams from one-day brilliant card records take a few minutes to gauge utilizing one CPU center. This process can be effectively parallelized in light of the fact that way estimation of each outing are absolutely autonomous. We can proficiently utilize 20 CPU centers on one server, in this manner we accomplish to remove one-month worth of traveler streams inside a few minutes. The measure of calculation time is satisfactory for current use. Many parts of our present framework are tentatively executed and have huge opportunity to get better as far as the execution.

6.2 Evaluating Traveler Streams after Mishaps

Mishaps some of the time cause benefit suspensions at a few segments, making travelers take temporary routes. In such circumstances, the course estimation technique proposed in Area 5.1 can't ascertain a fitting course on the grounds that the most limited way would be changed by administration suspensions. We can recomputed the briefest ways considering the suspension data, for example, suspended areas and time. This refines traveler stream estimation to make it more fitting to depict what occurred around then. We give interfaces to information imperatives of suspended lines as well as segments and begin and end times of suspension our visual investigation environment appeared.

The visual investigation environment pictures the recomputed result. We can then outwardly check how travelers take bypasses and focus on specific lines contrasts traveler streams.

We can utilize suspension data acquired from an outer data asset, for example, the metro working organization or the vehicle data website page as contributing imperatives for our framework. The metro working organization holds data about occasions that upset their tram framework. We likewise gather the prepare working condition data from the vehicle data website page of a third party company.

6.3 Exchange

Our travelers stream model is developed on the suspicion that each traveler will take the settled briefest way. Their genuine conduct is more differing; they will likewise consider travel charges, crowdedness, simplicity of prepare exchange, and so forth. Utilizing a probabilistic conduct model might be more proper to going on time. Trains were halted at specific areas on account of genuine mishaps, however more unobtrusive activities for example, deferring trains or somewhat killing administrations would have occurred much of the time. We have effectively attempted to build a preparatory conduct model of travelers after mishaps have happened [29] and plan to enhance the demonstrate by presenting such subtle elements. Since the present smart card framework does not have the data of passageway time or exchange purposes of every excursion, we can't assess the accuracy of the evaluated traveler streams straightforwardly. We talked with a portion of the staff of the prepare administrator and found that the separated stream appears to relate as far as anyone is concerned of the everyday operation. We plan to assess the accuracy through correlation with other factual overview results, for example, movement censuses.

7. CONCLUSION

We proposed a novel visual combination condition to investigate changes in streams of travelers on the Tokyo Metro and their circumstances and end results by utilizing over four years' worth of information removed from the smart card framework and Twitter.

Our novel approach empowers us to remove and imagine (1) traveler streams on an entangled metro organize from huge scale information from the smart card framework and (2) bizarre marvels what's more, their engendering on a spatio-worldly space. In addition, (3) we coordinated two types of huge (information from the shrewd card framework and Twitter) into a visual investigation framework to investigate causes as well as impacts of uncommon wonders. The contextual investigations and surveys via prepare working framework specialists demonstrated the conceivable outcomes and helpfulness of our framework to watch genuine circumstances amid the

occasions. We plan to give components to programmed identification also, expectation of occasions, and forecast and control of traveler streams on wide and complex transportation systems. through intertwining different sorts of enormous information streams including prepare trips data.

8. REFERENCES

[1] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst. 2013, pp. 334–343.

[2] R. Kr€uger, D. Thom, and T. Ertl, "Visual analysis of movement behavior using web data for context enrichment," in Proc. IEEE Pacific Vis. Symp., 2014, pp. 193–200.

[3] M.-P. Pelletier, M. Trepanier, and C. Morency, "Smart card data use in public transit: A literature review," Transp. Res.
Part C: Emerging Technol., vol. 19, no. 4, pp. 557– 568, 2011.

[4] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: Concepts, methodologies, and applications," ACM Trans. Intell. Syst. Technol., vol. 5, no. 3, pp. 1–55, 2014, Art. no. 38.

[5] I. Ceapa, C. Smith, and L. Capra, "Avoiding the crowds: Understanding tube station congestion patterns from trip data," in Proc. ACM SIGKDD Int. Workshop Urban Compute., 2012, pp. 134–141.

[6] L. Sun, D.-H. Lee, A. Erath, and X. Huang, "Using smart card data to extract passenger's spatio-temporal density and train's trajectory of MRT system," in Proc. ACM SIGKDD Int. Workshop Urban Compute., 2012, pp. 142–148.

[7]W. Zeng, C. Fu, S. M. Arisona, A. Erath, and H. Qu, "Visualizing mobility of public transportation system," IEEE Trans. Vis. Compute. Graph., vol. 20, no. 12, pp. 1833–1842, Dec. 2014.

[8] G. L. Andrienko and N. V. Andrienko, "Spatio-temporal aggregation for visual analysis of movements," in Proc. IEEE Symp. Visual. Anal. Sci. Technol., 2008, pp. 51–58.

[9]A. Slingsby, J. Wood, and J. Dykes, "Treemap cartography for showing spatial and temporal traffic patterns," J. Maps, vol. 6, no. 1, pp. 135–146, 2010.

[10] Z. Wang, T. Ye, M. Lu, X. Yuan, H. Qu, J. Yuan, and Q. Wu,"Visual exploration of sparse traffic trajectory data," IEEE Trans.Vis. Comput. Graph., vol. 20, no. 12, pp. 1813–1822, Dec. 2014.

[11] C. Tominski, P. Schulze-Wollgast, and H. Schumann, "3D information visualization for time dependent data on maps," in Proc. 9th Int. Conf. Inf. Vis., 2005, pp. 175–181.

[12] S. Thakur and A. J. Hanson, "A 3D visualization of multiple time series on maps," in Proc. 14th Int. Conf. Inf. Vis., 2010, pp. 336–343.

[13] H. Doraiswamy, N. Ferreira, T. Damoulas, J. Freire, and C. T. Silva, "Using topological analysis to support eventguided exploration in urban data," IEEE Trans. Vis. Comput. Graph., vol. 20, no. 12, pp. 2634–2643, Dec. 2014.

[14] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. van de Wetering,"Visual traffic jam analysis based on trajectory data," IEEE Trans. Vis. Comput. Graph., vol. 19, no. 12, pp. 2159–2168, Dec. 2013.

[15] G. Andrienko, N. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel, "From movement tracks through events to places: Extracting and characterizing significant places from mobility data," in Proc.IEEE Conf. Visual Anal. Sci. Technol., 2011, pp. 161–170.

[16] G. L. Andrienko, N. V. Andrienko, P. Bak, D. A. Keim, and S. Wrobel, Visual Analytics of Movement. New York, NY, USA:Springer, 2013.

[17] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, "Visual exploration of big spatio-temporal urban data: A study of New York city taxi trips," IEEE Trans. Vis. Comput. Graph., vol. 19, no. 12, pp. 2149–2158, Dec. 2013.

[18] I. Boyandin, E. Bertini, P. Bak, and D. Lalanne, "Flowstrates: An approach for visual exploration of temporal origin-destination data," Comput. Graph. Forum, vol. 30, no. 3, pp. 971–980, 2011.

[19] X. Jiang, C. Zheng, Y. Tian, and R. Liang, "Large-scale taxi O/D visual analytics for understanding metropolitan human movement patterns," J. Vis., vol. 18, no. 2, pp. 185– 200, 2015.

[20] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko "stacking-based visualization of trajectory attribute data," IEEE Trans. Vis. Comput. Graph., vol. 18, no. 12, pp. 2565–2574, Dec. 2012.

[21] T. Cheng, G. Tanaksaranond, C. Brunsdon, and J. Haworth "Exploratory visualization of congestion evolutions on urban transport networks," Transp. Res. Part C: Emerging Technol., vol. 36, no. 0, pp. 296–306, 2013.

[22] N. V. Andrienko and G. L. Andrienko, "Spatial generalization and aggregation of massive movement data," IEEE Trans. Vis. Comput. Graph., vol. 17, no. 2, pp. 205– 219, Feb. 2011.

[23] T. von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. van Wijk, J. Fekete, and D. W. Fellner, "Visual analysis of large graphs: State-of-the-art and future research challenges," Comput. Graph. Forum, vol. 30, no. 6, pp. 1719–1749, 2011.

[24] F. Beck, M. Burch, S. Diehl, and D. Weiskopf, "The state of the art in visualizing dynamic graphs," in Proc. EuroVis STAR, 2014, pp. 83–103.

[25] S. Hadlak, H. Schulz, and H. Schumann, "In situ exploration of large dynamic networks," IEEE Trans. Vis. Comput. Graph., vol. 17, no. 12, pp. 2334–2343, Dec. 2011.

[26] W. Dou, X. Wang, D. Skau, W. Ribarsky, and M. X. Zhou, "LeadLine: Interactive visual analysis of text data through event identification and exploration," in Proc. IEEE Conf. Visual Anal. Sci. Technol., 2012, pp. 93–102.

[27] A. M. MacEachren, A. R. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford, "SensePlace2: GeoTwitter analytics support for situational awareness," in Proc. IEEE Conf. Visual Anal. Sci. Technol., 2011, pp. 181–190.

[28] D. Thom, H. Bosch, S. Koch, M. W€orner, and T. Ertl, "Spatiotemporal anomaly detection through visual analysis of geolocated Twitter messages," in Proc. PacificVis, 2012, pp. 41–48.

[29] D. Yokoyama, M. Itoh, M. Toyoda, Y. Tomita, S. Kawamura, and M. Kitsuregawa, "A framework for largescale train trip record analysis and its application to passengers' flow prediction after train accidents," in Proc. 18th Pacific-Asia Conf. Adv. Knowl. Discovery Data Mining, 2014, pp. 533–544.