Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)

Incident Alert by an Anomaly Indicator of Probe Trajectories

Masanori Yoshida [a]*, Yosuke Kawasaki [a], Shogo Umeda [a], Masao Kuwahara [a]

[a] *Tohoku University, 6-6-06 Aramaki aza aoba, Aoba-ku, Sendai city, Miyagi pref. 980-8579, Japan*

Abstract

Herein, we propose methods that use three-dimensional trajectory data (one-dimensional time × two-dimensional space) acquired from probe vehicles to detect damaged spots on roads caused by natural disasters. Detecting road damages immediately after a disaster is essential to ensure quick and safe evacuations, rescue and relief operations, and efficient road repairs. However, road damages are currently monitored only by closed-circuit television, administrators patrolling the roads, and reports from users of the roads. Consequently, the number of locations that are being monitored is limited, and damage detection is delayed. Our proposed methods automatically estimate the locations of widespread damages on general roads. After analyzing and extracting the features of probe data obtained during a disaster, our proposed methods were verified using the data of past disasters. Our methods can identify anomalous vehicle behaviors and indicate various possibilities for detecting road damages.

Keywords: large-scale disaster; probe trajectory data; traffic-fault detection

1. Introduction

Recently, the frequency of large-scale disasters, such as the Great East Japan Earthquake (2011), has increased globally. Road damages caused by such disasters must be promptly detected to prevent secondary disasters such as road collapses and landslides. However, road damages are mainly detected through closed-circuit television (CCTV) road monitoring, patrolling by road administrators, and reporting by road users. These methods cause delays in the

* Corresponding author. Tel.: +81-22-795-7492; fax: +81-22-795-7494.
E-mail address: masanori.yoshida.s4@dc.tohoku.ac.jp

2352-1465

reporting of damages. For municipalities with long roads under their jurisdictions, these methods cannot immediately reveal the extent of road damages caused by a disaster. Thus, delayed reporting of road damages is a major concern that influences the traffic regulation timing of roads.

Furthermore, probe data are deemed a reliable resource to detect road damages. Probe data capture more detailed vehicle behaviors over a wide range than those captured by conventionally installed sensors such as vehicle detectors and CCTV. Additionally, the acquisition of probe data is not interrupted by disasters. The current study proposes methods to detect road damages using probe data.

Various studies have attempted to understand actual traffic situations and detect road damages during disasters. For example, Zhu et al. (2010) analyzed traffic pattern variations that occurred after the collapse of the I-35W bridge using vehicle sensors, bus user statistics, and questionnaire surveys. However, the limited coverage of vehicle detectors prevented a comprehensive overview of the traffic situations of the entire road network. Furthermore, as questionnaire surveys rely on the memory of the subjects, the survey answers could be inaccurate. However, probe data collected using mobile phones and car navigation systems are able to accurately capture the routes and speeds of vehicles traveling over an entire road network. Bengtsson et al. (2011), Lu et al. (2012), Hara et al. (2015), and Kawasaki et al. (2017) discussed the traffic situations captured by probe data in detail. Using the probe data and quantifying the differences between normal and disaster situations with regard to travel times and traffic congestion, Kawasaki et al. (2017) analyzed the actual traffic situations and evacuation efforts during the Kumamoto earthquake of April 2016. These examples have proved that probe data are useful to detect actual traffic situations occurring during a disaster.

Sekizuka et al. (2016) and Asakura et al. (2016) detected the time and location of the occurrence of road damage by analyzing the shock waves connecting the inflection points of vehicle trajectories in two dimensions (one-dimensional time × one-dimensional distance) when the damage occurred. However, these studies focused on expressways with vehicle detectors and access control. Therefore, these methods are not directly applicable to general road networks. Additionally, as the studies have evaluated only two-dimensional (2D) probe data, they did not capture vehicle behaviors such as speed reduction, stopping, and detouring around the damaged sites, on general roads during a disaster. Similar to our research, Cai et al. (2015) used the probe data to detect road damages at intersections on general roads. However, they focused on micro-vehicle behaviors, such as lane departures at intersections, excluding the extensive damages observed throughout the road network when the disaster occurred.

2. Characteristics of the three-dimensional probe trajectory data

We analyzed and organized the characteristics of the probe data related to road damages that occurred owing to the Kumamoto earthquake (2016). In this study, we defined the time at which the main shock occurred (1:25 a.m. on April 16th) as the time of the occurrence of the disaster. After the earthquake, traffic regulations, such as road closures, were imposed at several locations on major/living roads that were damaged by steps, cracks, and collapses.

Figs. 1(a) and (b) depict the probe data obtained during the disaster and normal times, respectively. The probe data are for the period ranging from midnight (before the earthquake) to 3 a.m. (after the earthquake). The visualization targets were six tertiary meshes (49301587–494302508) that covered the National Highway No. 3 Matsuzaki Overriding Bridge in Kumamoto City, Kumamoto Prefecture, Japan (see Fig. 1(c)). The earthquake damaged the road around the bridge, and traffic restrictions were imposed at approximately 2:30 a.m. In this study, April 9 (one week before the disaster) was set as the normal time. Probe data characteristics were observed to differ between the disaster and normal times, as described below.

- Drastic increase in the number of probe vehicles: after the disaster (after 1:25 a.m.), the number of probe vehicles increased drastically. Immediately after the disaster, several vehicles stopped at identical locations, which were presumed to have been used as evacuation site after the disaster.
- Variations in the three-dimensional (3D) trajectories: Generally, most vehicles travel smoothly. During the disaster, the shape of the trajectories changed; furthermore, trajectories displayed abnormal behaviors, such as wandering around, implying that the traces of traffic disorders when the disaster struck could be denoted as the differences between the trajectories during disaster and normal times.

Based on the aforementioned analysis, we sketched Fig. 2 (a), which exhibited the expected vehicle trajectories both generally and post road damage. At the locations of the road damages, behaviors, such as unnatural deceleration, backtracking, detouring, and similar meanderings, were likely to be observed. These behaviors were assumed to induce notable changes in the trajectories through two spatial dimensions and one temporal dimension, i.e., using 3D trajectories.

3. Methods to detect anomalous car movements using three-dimensional probe trajectory data

To quantify the differences in the 3D trajectories of vehicles during normal and disaster situations and to detect vehicles behaving in an abnormal manner, we estimated the similarities between the probe data during normal and disaster times. In our proposed method, the vehicles are detected to be “abnormal” when only minor similarities exist. We compared the pairs of probe data acquired at times $t = t_0, t_1, t_2, \ldots$, separated by a small and constant time interval, Δt. Further, we consider that the two trajectories, I and J, of the probe data comprise discrete data points and present them as $I = \{x^i(t_0), x^i(t_1), \ldots, x^i(t_n)\}$ and $J = \{x^j(t_0), x^j(t_1), \ldots, x^j(t_m)\}$ ($n \leq m$). Here, $x^i(t_k)$ denotes the position of trajectory I at time t_k. When the two trajectories, I and J, begin and terminate at similar positions, for $x^i(t_k)$ and $x^j(t_k)$, they are expressed as $x^i(t_0) = x^j(t_0)$ and $x^i(t_n) = x^j(t_m)$.

The similarities between the two trajectories using time-series information are often evaluated using dynamic time warping (DTW). DTW is advantageous as it can evaluate the differences between trajectories of different lengths of time series by estimating a combination that minimizes the intervals between the points. However, two problems are observed under certain conditions. First, when the spatial differences between the trajectories to be compared are observed to be small, the aforementioned advantage is lost. The second problem is the high cost of calculation that is incurred both when determining the Euclidean distances between all the points of the two trajectories and when calculating the shortest path. To resolve these problems, we propose a method based on the time–space distance (TSD) and one based on the Levenshtein distance.

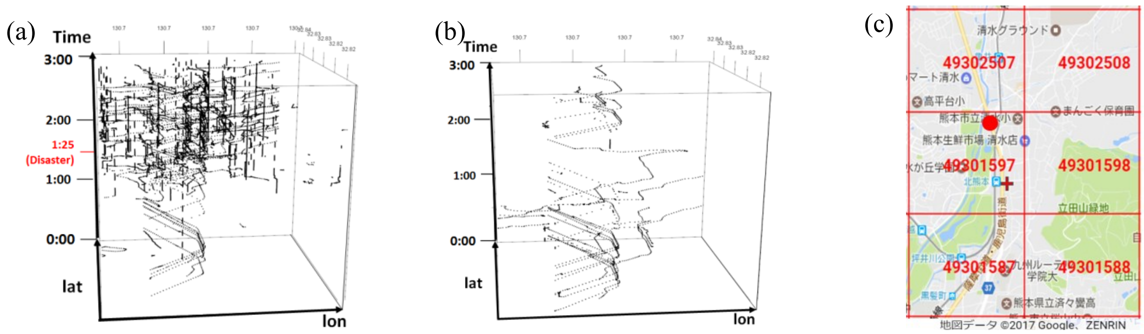

Fig. 1. (a) Probe data during the disaster; (b) Normal probe data; (c) Area to be visualized (In (a) and (b), the horizontal axis (Lon) is the longitude, the perpendicular axis (Lat) is the latitude, and the height axis (Time) is the time.)

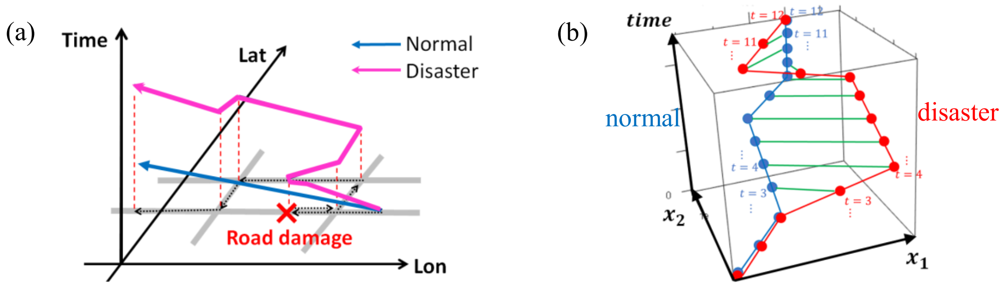

Fig. 2. (a) Changes in vehicle trajectory after the disaster; (b) schematic of TSD_{ij} and its component.

3.1. Abnormal vehicle detection method based on time–space distance (TSD)

The TSD index quantitatively evaluates the differences between two 3D trajectories. The TSD between two trajectories, I and J, with time-series lengths of n and m, respectively, can be expressed as TSD_{ij} (Eq. (1)), which evaluates the degree of similarity between I and J by adding the positional differences at each time point (from t_0 to $t_m > t_n$). To represent the distance between two points, the latitude and longitude are considered to be the coordinates on a 2D plane, and their difference is considered to be the Euclidean distance between the trajectories in this study.

The blue and red trajectories in Fig. 2(b) represent vehicle routes during normal and disaster times, respectively. At a given position, the trajectory during normal time occurs earlier than the one during the disaster. In particular, at the final positions of the two trajectories, $x^i(t_n) = x^j(t_m)$, the difference in arrival times, $t_m - t_n$, defines the delay of J caused by the incident. To evaluate the delay, the normal trajectory, I, is assumed to wait at the final position, $x^i(t_n) = x^j(t_m)$, from t_n to t_m.

$$TSD_{ij} = \int_{t_0}^{t_m} \left| x^i(t) - x^j(t) \right| dt \tag{1}$$

By discretizing the time variable, t, in Eq. (1) in units of Δt, we obtain Eq. (2).

$$TSD_{ij} = \sum_{t=0}^{t_m} \left| x^i(t) - x^j(t) \right| = \sum_{t=0}^{t_m} \sqrt{\left\{x_1^i(t) - x_1^j(t)\right\}^2 + \left\{x_2^i(t) - x_2^j(t)\right\}^2}\,, \tag{2}$$

where $x^i(t)$ is expressed in 2D coordinates as $x^i(t) = (x_1^i(t), x_2^i(t))$.

In Fig. 2(b), TSD_{ij} is the sum of the green lines, which represents the sum of the distances between I and J at each point in the time series. I, with a shorter time-series length, is assumed to spatially coincide with J from t_n to t_m.

3.2. Abnormal vehicle detection method based on Levenshtein distance

Alternatively, we propose a method that absorbs the observation error introduced using the global positioning system (GPS) by superposing the probe data on a mesh. Yamamoto et al. (2014) converted the trajectory pairs into character strings and calculated the similarities between the strings as the similarities between the trajectories. In a similar manner, our method can be executed in two steps as follows:

- converting the probe data trajectory into a string;
- using the Levenshtein distance to calculate the similarities between the character strings.

3.2.1 Method of converting the probe data into character strings

This subsection describes a method to convert a probe trajectory into a character string. First, the time is defined as $t = t_0, t_1, t_2, \ldots$ (discretized by a microscopic time interval Δt), and the space is defined in 2D with axes, x_1 and x_2, discretized by small intervals of Δx_1 and Δx_2, respectively. Further, a mesh delimited by Δx_1 and Δx_2 is set on the 2D space, and arbitrary characters (a, b, c…) are assigned to the mesh. The probe data are assumed to be obtained at a fixed time interval of Δt; further, at each Δt, the probe data are converted into letters corresponding to the mesh on a 2D plane. Figs. 3(a) and (b) depict the manner in which trajectories are converted into character strings under normal (traveling straight ahead) and detouring conditions. Spatial and temporal differences are represented by different types of characters and different lengths of character strings, respectively. Therefore, spatiotemporal differences, such as deceleration and detours, in the probe data are represented by different character strings. Hereafter, we refer to a character string converted from the probe data as a *locus character string* and the characters forming the string as the *locus characters*. By applying the probe data to a mesh of a certain size, the slight deviation of the probe data from its actual position due to GPS error can also be considered to exist at the same mesh (i.e., in the same position). Therefore, absorbing GPS errors in the probe data is possible. In contrast, in TSD, direct evaluation of the Euclidean distance between two probe data may affect calculation results due to the accumulation of errors contained in the GPS position data.

3.2.2 Method to evaluate the similarities between character strings

This subsection describes a method that uses the Levenshtein distance to evaluate the similarities between two character strings. The Levenshtein distance can be defined as the minimum number of times that one character string should be edited (insertion, deletion, or substitution) to be converted to another string. As the number of edits increases, the Levenshtein distance also increases. To calculate the Levenshtein distance, the algorithm proposed by Levenshtein (1966) was used. Fig. 3(c) depicts an example of Levenshtein distance calculation, wherein we define two character strings, $S_1 = \{\text{s a k a}\}$ and $S_2 = \{\text{a o b a}\}$, and proceed to edit S_1 to match S_2. However, S_1 must be edited at least thrice; thus, the Levenshtein distance between S_1 and S_2 is 3.

4. Verification of the proposed methods

We applied the previous section's proposed methods to the actual probe data and confirmed that the probe data could detect abnormal vehicle behavior. The methods were applied to two cases of road damages: one was the damage observed during the Kumamoto earthquake, whereas the other was a traffic accident site in the Sendai area. In this study, we used the statistically processed probe data that did not identify any individuals.

4.1. Outlines of the case studies

In case of the Kumamoto earthquake, the target vehicles were probe vehicles that had passed through two intersections located on both sides of the Matsuzaki Overpass Bridge on National Highway No. 3 (Fig. 4 (a)). Probe data of the vehicle passing through the Murozono intersection at 0:25 a.m. on April 9 were set as the reference data during normal times. The vehicles passing through the Murozono intersection between 12 a.m. and 3 a.m. on April 16 (the day of the earthquake), between 12 a.m. and 3 a.m. on April 9 (on a normal day) and passing through the Kaketsuken-mae intersection later were targeted for detection.

In case of the Sendai area, the target was a traffic accident that occurred at 4:04 p.m. on September 28, 2014, on National Highway No. 4 in Sendai City. The vehicles passing through Intersection A, as depicted in Fig. 4(b), and passing through Intersection B later were targeted for detection. The probe data of a vehicle passing through A at 3:33 p.m. on September 21, 2014, were set as the reference data at normal times. The targets were vehicles traveling from 3:00 p.m. to 6:00 p.m. on September 28, 2014 (the day of the accident) and vehicles traveling from 3:00 p.m. to 6:00 p.m. on September 21, 2014 (a normal day). The mesh ranges were determined to include the origin-destination to be detected (Fig. 4.), and divided into 25. Further, the distance that a vehicle could traverse during each interval of probe data acquisition and the average road distance should be considered to determine the mesh setting; however, this is a topic for future studies.

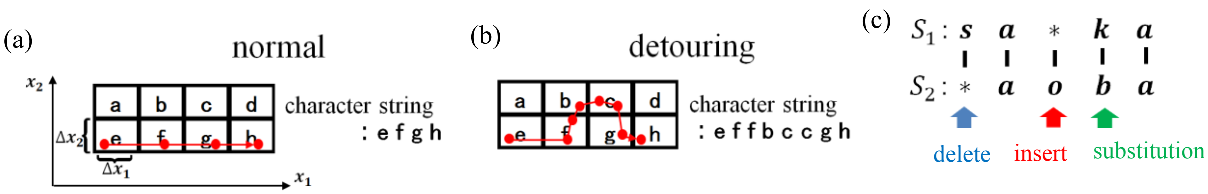

Fig. 3. Examples of trajectory conversions to character strings in (a) normal time and (b) detouring time. (c) Example of Levenshtein distance calculation.

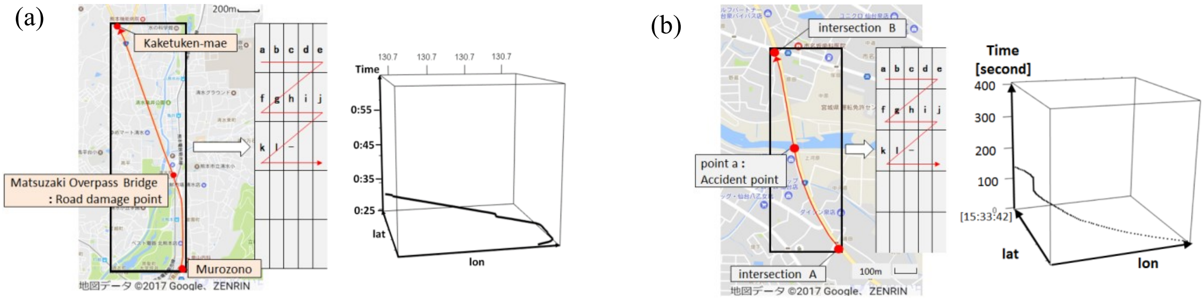

Fig. 4. Setting the verification target area, normal probe trajectory, and mesh setting during character string conversion in case of the (a) the Kumamoto earthquake and (b) Sendai area.

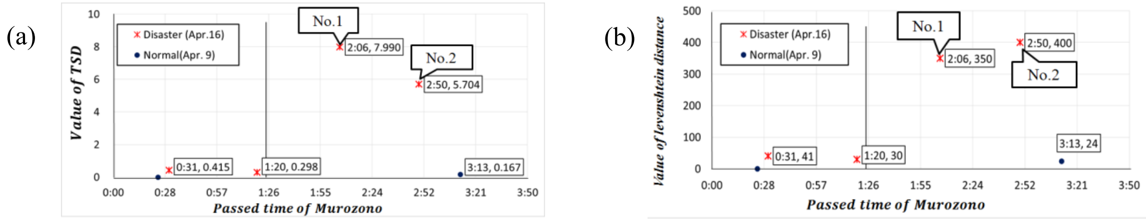

Fig. 5. Calculations of (a) TSD and (b) Levenshtein distances vs transit times over the Murozono intersection.

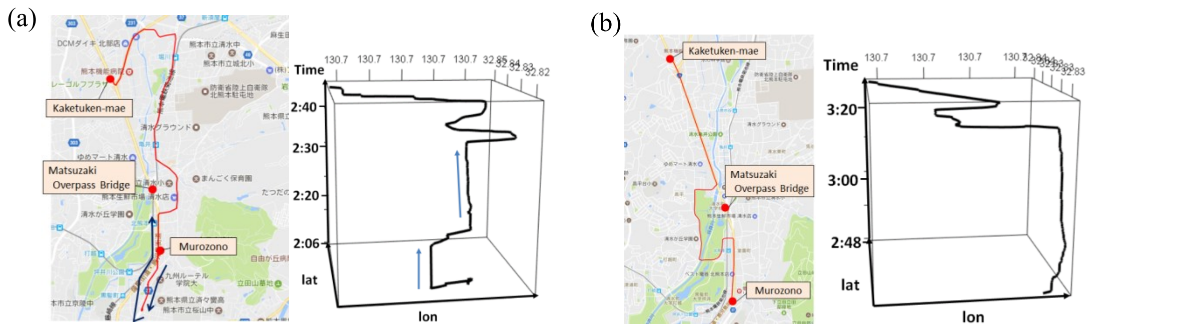

Fig. 6. Post-disaster trajectories of vehicles (a) No. 1 and (b) No. 2. In each panel, the left and right images represent the trajectories in 2D and 3D, respectively.

4.2 Calculations of similarities

We verified whether the proposed method could detect abnormally behaving vehicles. First, we calculated the similarities in case of the Kumamoto earthquake. Fig. 5(a) plots the result of the TSD whereas Fig. 5 (b) plots the Levenshtein distance. Both the TSD and Levenshtein distances of the two vehicles were several times to several tens of times higher after the disaster as compared to that at normal times. The two vehicles were detected as abnormal because the results (of both methods) for both increased dramatically after the disaster. Fig. 6 presents their probe data. By observing the two trajectories, we can infer that we were able to extract the characteristics of the vehicles, which were traveling at low speeds and wandering around.

Next, we applied the proposed method to the traffic accident site in the Sendai area. From Fig. 7, similar to the Kumamoto earthquake scenario, both the TSD and Levenshtein distances were several times to several tens of times higher after the accident as compared to that observed at normal times. Furthermore, the TSD and Levenshtein distances of the probe vehicles returned to near-normal values at a certain time (4:58 p.m.) after the accident. As the results of both the methods dramatically increased for the vehicles No.1 and No.2, the two vehicles were detected as abnormal. Fig. 8 (a) and (b) presents the probe data for these vehicles, which passed after the accident. Comparing the features of this case with those of the Kumamoto earthquake, the spatial differences between the trajectories are observed to be very small and only the temporal difference affects the similarities between the trajectories. Even in such situations, the proposed method is able to detect the abnormal behaviors of the vehicles.

Further, we considered the automatic detection of abnormally behaving vehicles using the values of the TSD or Levenshtein distances. The automated detection of abnormal vehicle behavior requires a threshold determination of the TSD or Levenshtein distances. As few of the probe data were sampled during the disaster, the index values during the disaster were difficult to understand and should be compared with their values during normal travel periods. Using the accident site's normal probe data, we calculated the TSD of the data (3:00 p.m.–6:00 p.m.; September 28, 2014; n = 5) on the day of the accident and on a normal day (12:00 a.m.–11:59 p.m.; September 21, 2014; n = 40). Fig. 8 (c) presents the corresponding distributions. Each vehicle at the time of the accident occupies a unique and separate position on the distribution. The probability distribution of the TSD or Levenshtein distances under normal conditions could be statistically evaluated if a sufficient amount of data was analyzed. Further, using the probability distribution, the probability of the abnormality will be able to be quantitatively determined (e.g., the Hotelling T^2 method).

5. Conclusions

This research clarified the spatial and temporal differences between the features of 3D probe vehicle trajectory data during normal and disaster periods and further exploited these differences in traffic-fault detection methods. Specifically, we proposed two detection methods: TSD, which directly evaluated the Euclidean distances between trajectories and a method based on the Levenshtein distance, which evaluated a trajectory on a mesh and converted the trajectory into a character string. These methods successfully detected abnormally behaving vehicles, indicating a possibility to detect locations of road damages. The proposed traffic-fault detection method focuses on both the temporal and spatial differences of the probe data. Thus, the model can achieve the following two goals:

- behavior evaluation of 3D probe vehicle trajectories on general roads;
- inspection of an entire road network (including narrow streets) without requiring the preparation of a high-speed network data. This can be achieved by applying the probe data as point data independent of the network data.

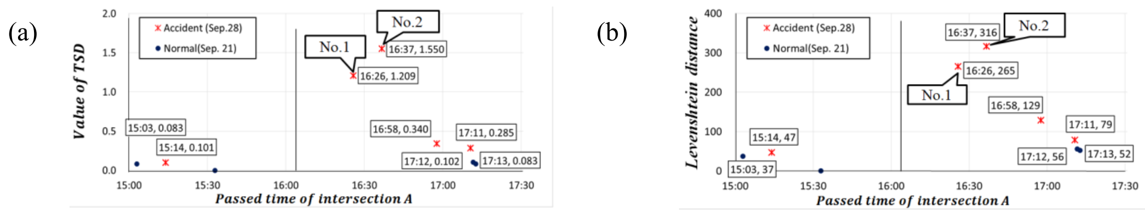

Fig. 7. Calculation results of the (a) TSD and (b) Levenshtein distances after different transit times at Intersection A

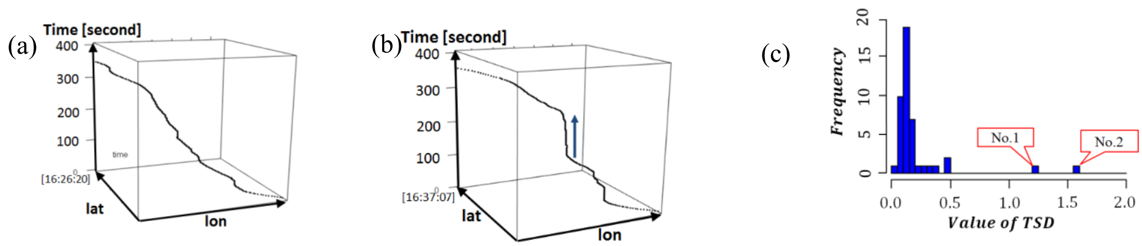

Fig. 8. Trajectories of vehicles (a) No.1 and (b) No. 2 after the accident on the bridge section of National Highway No. 4, Sendai City; (c) TSD distributions (n = 45) of both the vehicles at normal times and on the day of the accident.

Furthermore, we compared the differences between the two proposed methods, focusing on the complexities of the required calculations, necessity of presetting, and GPS errors. First, while calculating the similarities between two trajectories of lengths (number of dots), m and n ($m \geq n$), the computational complexity of the Levenshtein distance is O(mn) (similar to the result obtained by DTW) while that of TSD is O(m). Hence, TSD's computational complexity is comparatively less. Second, for presetting, the Levenshtein distance requires the mesh to be set in advance according to the probe data acquisition interval and the road network. However, TSD does not require presetting because it directly evaluates the Euclidean distance between two probe data points. Finally, because TSD directly evaluates the Euclidean distance between two probe data points, any errors contained in the GPS position data can accumulate and affect the calculations. While using the Levenshtein distance, the probe data is projected onto a mesh with a certain size. Therefore, GPS errors are absorbed, and their influence is diminished. Furthermore, comparing the accuracies of the false detections of general vehicles is necessary to compare the two indices. The results of the two methods may differ due to the different means of evaluating the distance of a trajectory according to the target event and shape of the road network. Therefore, applying the Levenshtein distance in various cases is necessary. Finally, topics for future studies can be summarized as follows.

- *Developing a method that automatically detects an anomaly in an entire road network*. Currently, for verification, we preset the OD of the location of the traffic problem and compare the similarities of the loci before and after the occurrence of the problem. However, to promptly detect traffic injuries during a disaster, the anomalous sections should be determined using the probe data of the entire network.
- *Setting the threshold to detect abnormalities*. Currently, vehicles with seemingly abnormal behavior were determined by calculating the extent of abnormal behavior. For automated detection, further data analysis under normal conditions and determination of the threshold value that indicates abnormal behavior is necessary.
- *Applying the method to other types of disasters*. The versatility and accuracy of this method must be verified by applying it to other types of disasters.

Acknowledgments

We express our appreciation for Mitsubishi Electric Corporation's collaborative research.

References

Zhu, S., Levinson, D., Liu, H.X., 2010. Kathleen Harder: The traffic and behavioral effects of the I-35W Mississippi River bridge collapse. Transportation Research Part A: Policy and Practice, Vol.44, Issue 10, 771–784.

Bengtsson, L., Lu, X., Thorson, A., Garfield, R., von Schreeb, J., 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. PLoS Med., Vol.8, No.8, p. e1001083.

Lu, X., Bengtsson, L., Holme, P., 2012. Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. USA, Vol. 109, No. 29, 11576–11581.

Hara, Y., Kuwahara, M., 2015. Traffic Monitoring immediately after a major natural disaster as revealed by probe data-A case in Ishinomaki after the Great East Japan Earthquake. Transportation Research Part A: Policy and Practice, Vol. 75, 1–15.

Kawasaki, Y., Kuwahara, M., Hara, Y., Mitani, T., Takenouchi, A., Iryo, T. and Urata, J., 2017. Investigation of Traffic and Evacuation Aspects at Kumamoto Earthquake and the Future Issues. Journal of Disaster Research, Vol. 12, No. 2.

Sekizuka, K., Mitani, T., Kawasaki, Y., Masuda, T., Nagai, S., Kuwahara, M., 2016. Travel time prediction under incident using only vehicle trajectories. JSCE, Vol. 53, CD-ROM, in Japanese.

Asakura, Y., Kusakabe, T., Nguyen, L., X., and Ushiki, T., 2016. Incident Detection Methods using Probe Vehicles with on-board GPS Equipment. Transportation Research Part C: Emerging Technologies, In press.

Cai, Y., Wang, H., Chen1, X., Jiang, H., 2015. Trajectory-based anomalous behavior detection for intelligent traffic surveillance. IET Intelligent Transport Systems, Vol.9, 810-816.

Ohashi, H., Shimizu, T., Yoshikawa, M., 2016. Similarity search of trajectory data with dynamic attributes. DEIM2016, G3-1, in Japanese.

Yamamoto, M., Kurita, K., Sasawaga, H., Arimura, H., 2014. Faster pattern matching algorithm for 2-dimensional trajectory data. FIT2014, Vol.13, No.2, 153-154, in Japanese.

Levenshtein, V., 1966. Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady, Vol.10, No.8, 707-710.