ANOMALOUS TRAJECTORY DETECTION METHOD BASED ON FRACTAL DIMENSION ANALYSIS OF GPS

[1]MINJEONG AN, [2]HONGCHUL LEE, [3]GARLAND CHOW

[1,2]School of Industrial Management Engineering, Korea University, Republic of Korea
[3]Sauder School of Business, The University of British Columbia, Canada
E-mail: [1]ansebby@korea.ac.kr, [2]hclee@korea.ac.kr, [3]Garland.Chow@sauder.ubc.ca

Abstract— As GPS devices are generally being supplied in transportation method such as taxi, delivery truck, a vast amount of GPS data is being accumulated. Through these GPS data, it's possible to monitor vehicle location in real-time and to analyze of the efficiency of the transportation. In this study, we had made an indicator for measuring tortuosity of vehicle movement based on fractal dimension analysis of the GPS trajectory. Using this indicator and Google direction API, we had extracted anomalous operation of the Vancouver taxi. We used the daily operations of 265 taxi drivers in over 81,000 of trips. This method can contribute to improve transportation service as being utilizing indicator for measuring transportation quality.

Index Terms—Anomalous Trajectory Detection, Trajectory Outlier Detection, Driving Evaluation, GPS Trace, Fractal Analysis.

I. INTRODUCTION

This study introduces a method for detecting anomalous driving trajectory of a GPS device attached vehicle. In recent years, as the attachment of GPS device in vehicles such as taxi and delivery truck became generalized, we can collect the GPS data and monitor the driving route of vehicles in real time. In using the collected GPS data, some researches about analysis of driver's behavior and detection of abnormal driving route have been performed around USA and China. The latter refers to as "anomalous trajectory detection", "driving fraud detection" and "trajectory outlier detection". In most research about this, they detected abnormal trajectory based on pattern recognition and stochastic process, utilizing accumulated past GPS data having same origin and destination. These methods can't be applied to the variety route having different origin and destination. And there are some overhead about accumulating GPS data and operating that.

In this study, we defined tortuosity as the anomalous degree of a single driving route based on fractal dimension, and suggested a method to utilize this as an index of driving service quality. To reflect the shape of real road, we measured the straightness index and compensated the tortuosity.

The remainder of this paper is organized as follows. After providing as overview the relevant existing literature on the anomalous trajectory detection and fractal dimension analysis in Section 2, we describe GPS data extraction and preprocessing in Section 3. In Section 4.1, we introduce method to compute the tortuosity of a route. In Section 4.2, we detail the tortuosity, including straightness index. The computational results are reported in Section 5, followed by the conclusions with some mention about future works in Section 6.

II. LITERATURE REVIEW

In this section, we briefly review the related literature which can be grouped into three categories. The first category includes the work on analyzing GPS traces. Liu et al. presented a good model to analyze taxi driver's operation pattern using GPS data of taxi [1], [2]. After classifying top driver and ordinary driver using K-means clustering method, they quantitatively analyzed a causal relationship between taxi driver's operation pattern and taxi driver's income. The operational scale of a top taxi driver was much larger than that of the ordinary driver. Within 3.5 km which is basic fare distance, the ordinary driver's driving was more tortuous than the top driver's driving. On the contrary to this, the top driver's driving was more tortuous than the ordinary driver's driving beyond 3.5 km. The top driver adopted the shortest path to save time within 3.5 km, while he adopted a more tortuous and longer path to increase the travel distance and maximize his income beyond 3.5 km. For measuring tortuosity of driving, they used fractal dimension formula. We thought that fractal value's gap of the top driver and the ordinary driver should be reduced to remain consistent taxi service quality for customer satisfaction. Some portion of this difference is derived from tortuous route of taxi driver. This gap could be influenced by taxi fare policy and taxi driver's income policy. If the taxi driver's mobility is measured and reported periodically by system, we can induce that taxi driver's service quality is remained consistently. Tang et al. analyzed travel demand distributions using Taxi GPS trajectories data [3]. They usedDensity-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster pick-up and drop-off locations. Finally, they constructed the traffic distribution patterns in Harbin city based on entropy-maximizing method based on actual OD matrix. This study presents a good method to forecast traffic distribution using accumulated past GPS data.

The second category focuses on anomalous trajectory detection. Taxi driving frauds are often committed by greedy taxi drivers who overcharge passengers by deliberately taking unnecessary detours [4]. This

problem occurs customer's complaint about taxi service. For solving this problem, many studies were performed in terms of anomalous trajectory detection. Zhang et al.proposed an Isolation-Based Anomalous Trajectory (iBAT) detection method verifying with large scaled taxi data [5]. After extracting taxi trajectories from GPS data, they split the city map into grid-cells of equal size. Then they grouped all taxi trajectories crossing the same source-destination cell-pair, augmented and represented each trajectory as a sequence of traveling cells. They exploited the few and different properties of anomalous trajectories, and applied the isolation mechanism to detect anomalous trajectories. Chen et al. proposed isolation-based online anomalous trajectory detection method (iBOAT) which is improved from iBAT in terms of travel area and real time processing [6]. These are methods of detecting anomalous routes by comparing the latter against time-dependent historically normal routes in common. Ge et al. developed a taxi driving fraud detection system based on Dempster-Shafer theory, which is combined travel route evidence and driving distance evidence [4]. Liu et al. proposed Speed-based Fraud Detection System (SFDS) in a trajectory-free and map-free scenario [7]. They discovered that 90.7% of the total fraud behaviors happened in transportation intersections and 47.2% of the total fraud behaviors happened during rush hours. Lee et al. developed a trajectory outlier detection algorithm TRAOD, which is consisted of partitioning and detection [8]. They detected the outlying sub-trajectories from a trajectory database in this study. Almost of study about fraud detection was based on accumulated historic trajectory data. We approached intuitively fraud detection from adopting fractal dimension which was used by Liu et al. [1].

The third category is the study on analyzing fractal dimension. As we have said above, Liu et al. adopted fractal dimension to analyze taxi driver's driving tortuosity [1]. Benhamou used fractal dimension to estimate the tortuosity of an animal's path [9]. Nams proposed the VFractal which is a new estimator for fractal dimension of animal movement paths [10]. Hadjileontiadisa et al. developed a fractal dimension-based crack detector (FDCD) in beam structures [11]. Sarkar and Chaudhuri proposed a method to estimate fractal dimension in a two-dimensional (2D) image which can readily be extended to a 3D image [12]. Fractal dimension was applied to the problem of tumor shape quantification in diagnosis [13]. Many researchers have adopted fractal dimension to measure complexity of the target object in different field of research.

From the above literature review, we can know that GPS data mining is important to the transportation area and that fractal dimension analysis is an appropriate approach to measure tortuosity of GPS trajectory.

III. SYSTEM OVERVIEW

The Logistics business is made up of task such as order, transportation plan, delivery plan, load optimization, route optimization, shipping and calculation, until complete delivery after receiving order of the customer [14]. This business process has been operated with IT system such as logistics management system (LMS) or transportation management system (TMS) to reduce cost and to improve the efficiency of logistics.The core part of TMS is transportation management component (TMC) which performs transportation plan and route optimization.Taxi companies are also operating transportation management system (TMS) which dispatches a taxi to customer.

In this study, we supplemented driving evaluation component (DEC) to TMS. Fig.1 shows the overall system configuration of TMS in logistics business. This configuration is focused on Taxi business. TMC is made up of vehicle's location data and vehicle's dispatch data. Vehicle's location data is built up by GPS receiver module. Vehicle's dispatch data is generated by customer order and dispatch logic in TMC. DEC is made up of vehicle's movement data and vehicle's driving evaluation data. When the transportation is finished, DEC evaluates vehicle's driving by fractal dimension analysis, and sends feedback to vehicle's dispatch logic and the vehicle.

IV. MATERIAL AND METHOD

A taxi trip is composed of one dispatch data and series of GPS coordinate data. One dispatch data set is composed of overall data about each trip from customer's order to driving's evaluation. We acquired the large scaled taxi GPS data and dispatch data from one of four taxi companies in Vancouver, Canada, for three months (October to December 2014).

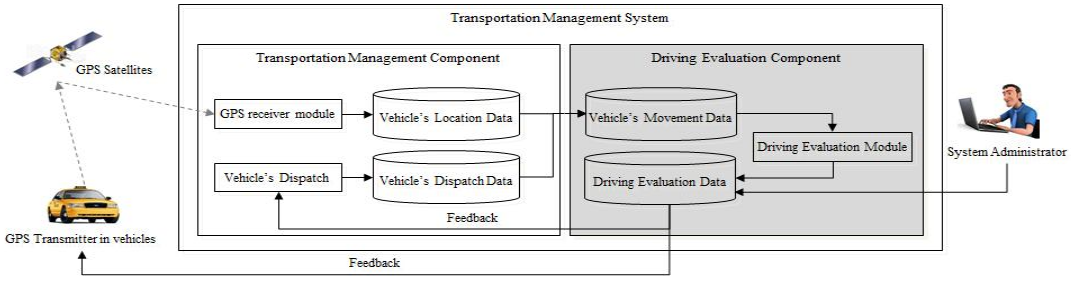


Fig.1. The overall system configuration

A. Data Extraction and Preprocessing

We extracted trip id, vehicle number, driver id, origin latitude, origin longitude, destination latitude, destination longitude, start time and arrival time from taxi dispatch data. And we extracted vehicle number, driver id, created time, latitude and longitude from taxi GPS data. Driving evaluation data and vehicle movement data is mapped with TRIP_ID which is unique key about one taxi trip. After filtering wrong data, we used 81,013 dispatch data and 8,321,572 GPS data to conduct this research. And we generated additional data for fractal dimension analysis. Fig. 2 shows logical data model of DEC.

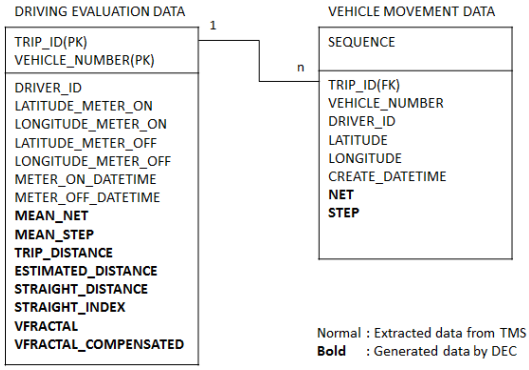


Fig. 2. Logical Data Model

B. Fractal Dimension Analysis

A fractal is a natural phenomenon or a mathematical set that exhibits a repeating pattern that displays at every scale [15]. A fractal dimension is an index for characterizing fractal patterns or sets by quantifying their complexity as a ratio of the change in detail to the change in scale [16]. When a shape is disassembled by the smaller self-similar shape, fractal dimension d is presented as equations (1), (2) [15]. N is a number of disassembled shapes, r is a rate of reduction [15].According to this equation, fractal dimension of straight line is 1 and fractal dimension of surface is 2. Fractal dimension of Koch curve like Fig. 3 is “1.2618”. This number can present the complexity of line. The more complex the shape of line, the fractal dimension is close to 2.

$$\mathrm{N} \times r^{d} = 1 \qquad (1)$$

$$\mathrm{d} = \frac{\log N}{\log \frac{1}{r}} \qquad (2)$$

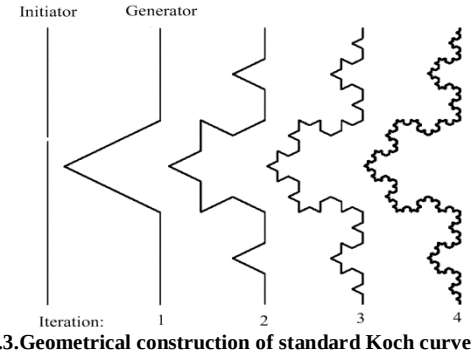


Fig.3.Geometrical construction of standard Koch curve [17]

C. Indicator of Tortuosity

Using fractal dimension, Nams suggested the VFractal as indicator to analyze movement characteristics of animal [10].Fig 5 shows two steps of animal’s movement path shaping "V".The thick curve is the animal’s movement path. The thick dotted lines represent steps of equal size Salong the path, and the two steps form the V. θrepresents oneturn angle between two steps, and Net is the straight distance fromthe start of one to the end of the second step [10].As this shape is repeated through all path, this can be analyzed by fractal dimension.

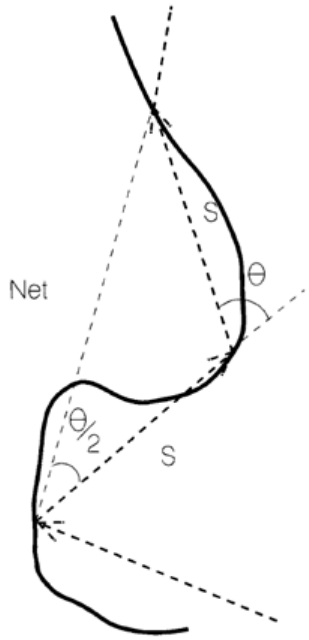


Fig. 4. One individual V from an animal's path [10]
Equation (3) represents an estimator of fractal d (VFractal) for the V formed by one pair of steps [10].

$$d = \frac{\log 2}{\log \frac{Net}{s}} = \frac{\theta}{1 + \log_2(\cos\theta + 1)} \qquad (3)$$

If we use the mean of "Net" and "s", we can calculate the VFractal about overall movement path.ThisVFractal score represents the tortuosity of animal’s movement.VFractal value d is between 1 and 2. If the VFractal valuesare lower than 1, this case has GPS data communication problem or canceled trip. When the VFractal values are higher than 2, we specified maximum value 2 in VFractalbecause GPS trajectory is line.

This equation can be applied at logistics business for evaluation of transportation service quality.

D. Compensation of Road Condition

In case driving along a winding road, the VFractal value is higher than in straight road, but this driving is not fraud.To reflect the real road condition in the VFractal value, we defined the Straightness Index (SI) and made the VFractal_C. Equation (4) represents the SI.

$$\mathrm{SI} = \frac{D}{L} \qquad (4)$$

The SI is the straight distance (D) divided by the real road distance (L) from origin to destination. Fig.5 is a diagram about D and L.

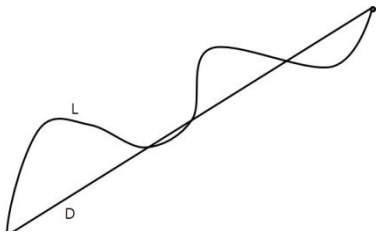


Fig.5. Real Road Shape

The more tortuous road condition is, the SI value is close to zero.The D value can be generated by calculation of straight distance between two GPS coordinates of origin and destination.We obtained L using Google Direction API that provide estimated distance based on real road condition from origin to destination.And we calculated the VFractal_Caccording to equation (5). Because GPS trajectory is a line, d is the value between 1 and 2. We multiplied SI at the part of decimals of VFractal.

$$VFractal_C = (d - 1) \times SI + 1 \quad (5)$$

By compensation with equation (5), in case of curved road condition, VFractal_C value become lower than VFractal d value.

E. Anomalous Trajectory Detection Process

We detected anomalous trajectory by setting the threshold of VFractal_C to "1.8". This value can be adjusted according to policy of each transportation company. Fig.6 shows overall anomalous trajectory detection process. After extracting vehicle's dispatch and GPS data from TMS, additional data is generated by calculating VFractal, SI and VFractal_C. The VFractal_C value is sent to the vehicle and dispatch module in TMS.This process is performed whenever each trip is finished.

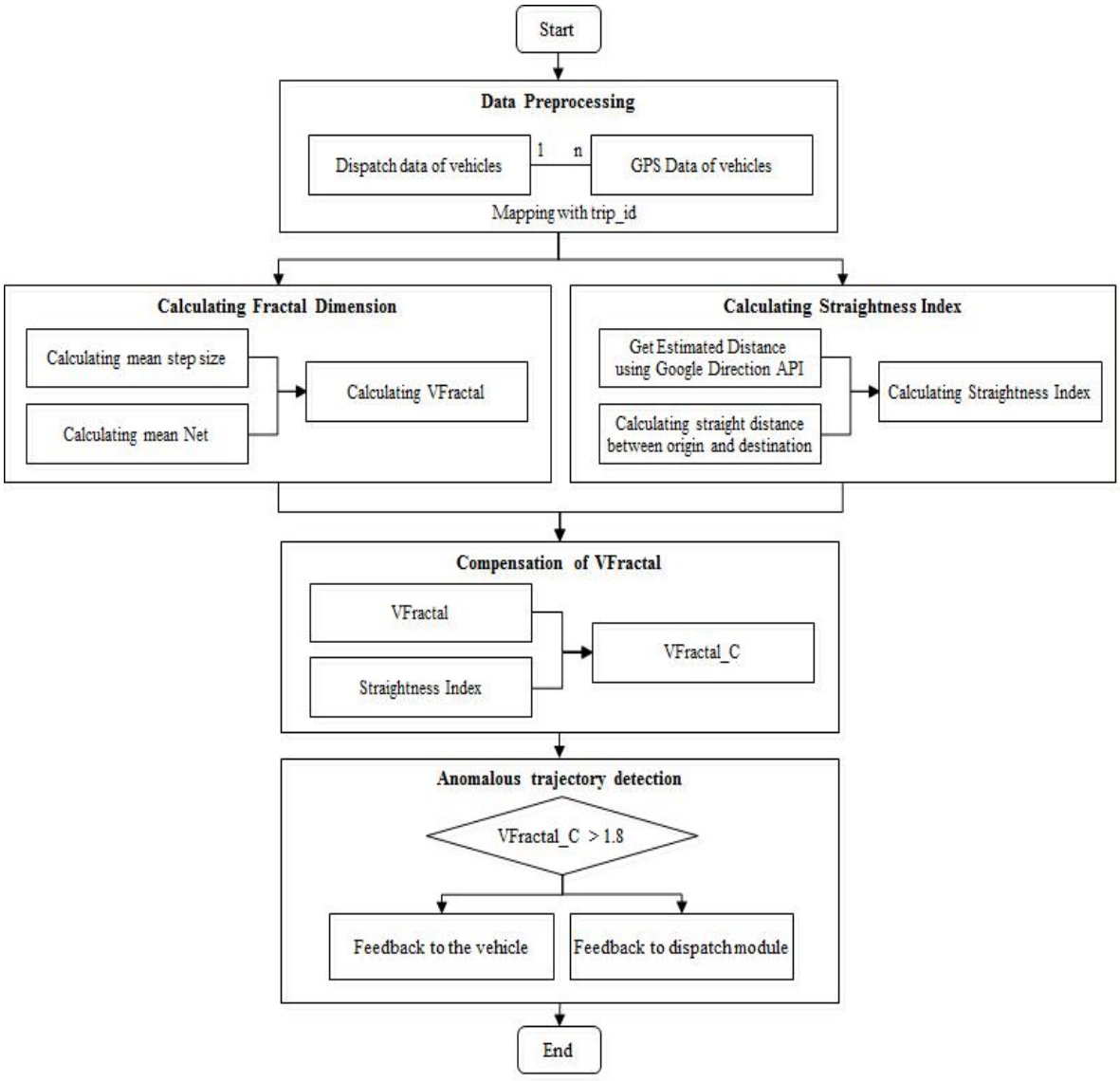


Fig.6. Anomalous Trajectory Detection Process

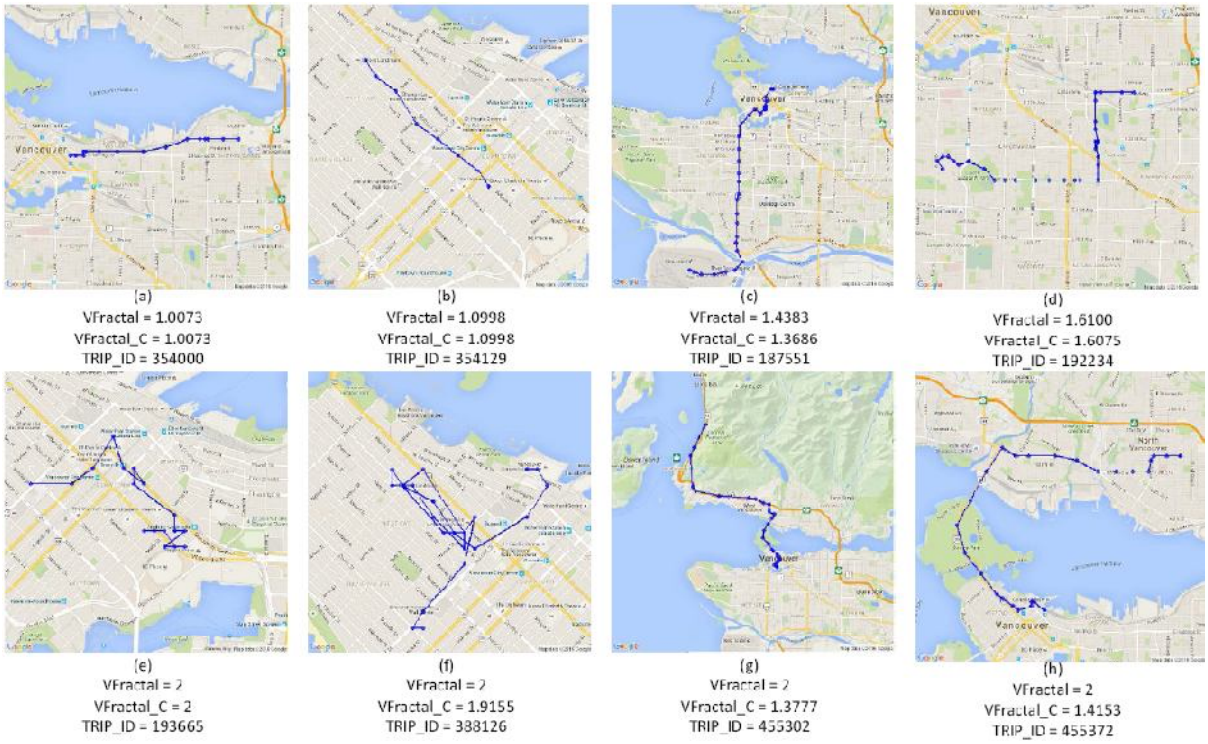


Fig.7. The Result of Driving Evaluation

V. RESULTS

We could detect anomalous trajectory checking VFractal_C value. Fig. 7 shows the real GPS trajectories and VFractal and VFractal_C value. When VFractal_C values come close to 1, trajectories were almost straight line like (a), (b) of Fig. 7. When VFractal_C values come close to 2, trajectories were complex and tortuous line like (e), (f) of Fig. 7. In case of check pattern's road, VFractal_C value comes between 1.3 and 1.6 like, despite driving was honest, shown as (c), (d) of Fig. 7.In case of (g), (h) of Fig. 7, the VFractal value is maximum value 2, but the VFractal_C value became lower than 1.5 by compensation process reflecting real road condition. These were evaluated to normal trajectory, because VFractal_C is lower than threshold value 1.8. In this result, we can know that the VFractal_C is a reasonable indicator for evaluating tortuosity of taxi trajectory.

CONCLUSION AND FUTURE WORKS

In this study, we proposed the VFractal_C as an indicator for measuring vehicle's driving service quality in transportation area. This study has differentiated features which are proposed to anomalous trajectory detection method based on fractal dimension analysis and Google Direction API about single driving trajectory in contrast with previous studies based on accumulated historical data of same origin and destination.

If taxi companies adopt this method, they can detect abnormal driving patterns of taxi driver and they can make method to reduce these abnormal behavior.As a result, they can improve taxi service quality.In case of utilizing this method in logistics business, they can detect abnormal driving behavior about transportation and delivery. According to driving evaluation, driver's work performance can be measured by this VFractal_C value. By feedback to transportation and delivery plan, this method can contribute to enhance the logistics work efficiency.If this fractal dimension analysis applies to autonomous vehicle, the driving quality can be enhanced by sending periodically feedback "VFractal_C" to the vehicle.

ACKNOWLEDGEMENT

This work was supported by UBC (The University of British Columbia), NSERC (Natural Sciences and Engineering Research Council of Canada), the BK21 Plus (Big Data in Manufacturing and Logistics Systems, Korea University) and WISET (Center for Women In Science, Engineering and Technology).

REFERENCES

[1] L. Liu, C. Andris, A. Biderman and C.Ratti, "Uncovering Taxi Drivers' Mobility Intelligence through His Trace," IEEE Pervasive Computing, vol. 160,pp. 1-17, 2009

[2] Li. Liu, C. Andris, C. Ratti, "Uncovering cabdrivers' behavior patterns from their digital traces," Computers, Environment and Urban Systems,vol. 34,no. 6, pp. 541–548,November 2010.

[3] J. Tang, F. Liu, Y. Wang and H. Wang, "Uncovering urban human mobility from large scale taxi GPS data,"Physica A: Statistical Mechanics and its Applications,vol. 438, pp. 140–153,November 2015.

[4] Y. Ge, H. Xiong, C. Liu and Z. H. Zhou, "A Taxi Driving Fraud Detection System,"in 2011 11th IEEE International

Conference on Data Mining (ICDM), Vancouver, Canada, pp. 181–190, 2011.

[5] D. Zhang, N. Li, Z. H. Zhou, C. Chen andL. Sun, "iBAT: Detecting Anomalous Taxi Trajectories from GPS Traces," in 2011 13th International Conference on Ubiquitous Computing (UbiComp'11), Beijing, China, pp. 99-108,September 2011.

[6] C. Chen, D. Zhang, P. S. Castro, N. Li, L. Sun, S. Li andZ. Wang, "iBOAT: Isolation-Based Online Anomalous Trajectory Detection," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 2, pp. 806–818, June 2013.

[7] S. Liu, L. M. Ni andR. Krishnan, "Fraud Detection from Taxis' Driving Behaviors," IEEE Transactions on Vehicular Technology, vol. 63, no. 1, pp. 464–472, January 2014.

[8] J. Lee, J. Han andX. Li, "Trajectory Outlier Detection: A Partition-and-Detect Framework," in: 2008 IEEE 24th International Conference on Data Engineering (ICDE), Cancun, Mexico, pp. 140-149, April 2008.

[9] S. Benhamou, "How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension?"Journal of Theoretical Biology,vol. 229, no. 2, pp. 209–220, July 2004.

[10] V. O. Nams, "The VFractal: a new estimator for fractal dimension of animal movement paths," Landscape Ecology, vol. 11, no. 5, pp. 289-297, October 1996.

[11] L. J. Hadjileontiadisa, E. Doukab andA. Trochidisc, "Fractal dimension analysis for crack identification in beam structures," Mechanical Systems and Signal Processing,vol. 19, no. 3, pp. 659–674, May 2005.

[12] N. Sarkar and B. B. Chaudhuri, "An efficient approach to estimate fractal dimension of textural images," Pattern Recognition, vol. 25, no. 9, pp. 1035-1041,September 1992.

[13] G. Landini andJ. W. Rippin,"How important is tumour shape? Quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis," Journal of Pathology, vol. 179, no. 2, pp. 210-217, June 1996.

[14] M. J. An and H. C. Lee, "Design of the Integrated Logistics Information System based on Cloud Computing," Lecture Notes on Software Engineering, vol. 3, no. 1, pp. 31-34, February 2015.

[15] B. B. Mandelbrot, "The fractal geometry of nature," Vol. 173, Macmillan, 1983.

[16] B. B. Mandelbrot, "How long is the coast of Britain? Statistical self-similarity and fractional dimension," Science,vol. 156, pp. 636-638, 1967.

[17] K. J. Vinoy, J. K. Abraham and V. K. Varadan, "On the relationship between fractal dimension and the performance of multi-resonant dipole antennas using Koch curves,"IEEE Transactions onAntennas and Propagation, vol. 51, no. 9, pp. 2296-2303, September 2003.

★★★