performance over various multipath channels. Finally, as compared with the substantial performance improvement offered by the PISL algorithm, the complexity increase to implement the PISL algorithm is very small. Therefore, the PISL algorithm can be considered an enabling technology for monobit IR-UWB receivers that demand high performance, lower power consumption, and low complexity.

REFERENCES

[1] F. Nekoogar, *Ultra-Wideband Communications: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[2] A. E. Molisch, "Ultrawideband propagation channels—Theory, measurement, and modelling," *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1528–1545, Sep. 2005.

[3] M. Z. Win and R. A. Scholtz, "On the energy capture of ultrawide bandwidth signals in dense multipath environments," *IEEE Commun. Lett.*, vol. 2, no. 9, pp. 245–247, Sep. 1998.

[4] R. Hoctor and H. Tomlinson, "Delay-hopped transmitted-reference RF communications," in *Proc. IEEE Ultra Wideband Syst. Technol. Conf.*, Baltimore, MD, USA, May 2002, pp. 265–269.

[5] J. Romme and K. Witrisal, "Transmitted-reference UWB systems using weighted autocorrelation receivers," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 4, pp. 1754–1761, Jun. 2006.

[6] S. Franz and U. Mitra, "Generalized UWB transmitted reference systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 780–786, Apr. 2006.

[7] H. Khani and P. Azmi, "Weighted high data rate UWB-TR system in dense multipath fading channels," *IET Commun.*, vol. 3, no. 4, pp. 571–584, Apr. 2009.

[8] A. A. D'Amico, "IR-UWB transmitted-reference systems with partial channel knowledge: A receiver design based on the statistical invariance principle," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1435–1448, Apr. 2011.

[9] H. Khani, P. Azmi, and H. Nie, "Performance analysis of high rate weighted-TR UWB system in the presence of inter-block and multiuser interferences," *AEU-Int. J. Electron. Commun.*, vol. 66, no. 3, pp. 219–227, Mar. 2012.

[10] S. Hoyos, B. M. Sadler, and G. R. Arce, "Monobit digital receivers for ultrawideband communications," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1337–1344, Jul. 2005.

[11] Q. Zhang, A. Nallanathan, and H. K. Garg, "Monobit digital Eigen-based receiver for transmitted-reference UWB communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2312–2316, May 2009.

[12] A. A. D'Amico, U. Mengali, and L. Taponecco, "Synchronization for differential transmitted reference UWB receivers," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4154–4163, Nov. 2007.

[13] L. Ke, H. Yin, W. Gong, and Z. Wang, "Finite-resolution digital receiver design for impulse radio ultra-wideband communication," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5108–5117, Dec. 2008.

[14] H. Yin, Z. Wang, L. Ke, and J. Wang, "Monobit digital receivers: Design, performance, and application to impulse radio," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1695–1704, Jun. 2010.

[15] H. Khani, H. Nie, W. Xiang, and Z. Chen, "Finite-resolution digital receiver for high rate ultra-wideband weighted-transmitted reference system," in *Proc. Int. Conf. Ultra-Wideband*, Bologna, Italy, Sep. 2011, pp. 200–204.

[16] J. Gil-Pelaez, "Note on the inversion theorem," *Biometrika*, vol. 38, no. 3/4, pp. 481–482, 1951.

[17] Time Domain Corp., P210 Integratable Module Data Sheet, 320-0095 Rev B, Time Domain Corp., Huntsville, AL, USA.

[18] P. C. Richardson, W. Xiang, and W. Stark, "Modeling of ultra-wideband channels within vehicles," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 906–912, Apr. 2006.

[19] R. Schober, *Noncoherent Detection and Equalization for MDPSK and MDAPSK Signals*, B. Girod and J. Huber, Eds. Aachen, Germany: Shaker Verlag, 2000.

[20] R. L. de Lacerda Neto, A. M. Hayar, and M. Debbah, "Channel division multiple access based on high UWB channel temporal resolution," in *Proc. Veh. Technol. Conf.*, Sep. 2006, pp. 1–5.

Fraud Detection From Taxis' Driving Behaviors

Siyuan Liu, *Member, IEEE*, Lionel M. Ni, *Fellow, IEEE*, and Ramayya Krishnan

***Abstract*—Taxi is a major transportation in the urban area, offering great benefits and convenience to our daily life. However, one of the major business fraud in taxis is the charging fraud, specifically overcharging for the actual distance. In practice, it is hard for us to always monitor taxis and detect such fraud. Due to the Global Positioning System (GPS) embedded in taxis, we can collect the GPS reports from the taxis' locations, and thus, it is possible for us to retrieve their traces. Intuitively, we can utilize such information to construct taxis' trajectories, compute the actual service distance on the city map, and detect fraudulent behaviors. However, in practice, due to the extremely limited reports, notable location errors, complex city map, and road networks, our task to detect taxi fraud faces significant challenges, and the previous methods cannot work well. In this paper, we have a critical and interesting observation that fraudulent taxis always play a secret trick, i.e., modifying the taximeter to a smaller scale. As a result, it not only makes the service distance larger but also makes the reported taxi speed larger. Fortunately, the speed information collected from the GPS reports is accurate. Hence, we utilize the speed information to design a system, which is called the Speed-based Fraud Detection System (SFDS), to model taxi behaviors and detect taxi fraud. Our method is robust to the location errors and independent of the map information and road networks. At the same time, the experiments on real-life data sets confirm that our method has better accuracy, scalability, and more efficient computation, compared with the previous related methods. Finally, interesting findings of our work and discussions on potential issues are provided in this paper for future city transportation and human behavior research.**

***Index Terms*—Behavioral science, data analysis, vehicle detection, vehicle dynamics.**

I. INTRODUCTION

Taxi service is one of the major transportation services in the modern urban area, offering great benefits and convenience in our daily life. However, due to the intense competition among taxi service companies, the illegal overcharging problem (or taxi fraud) in taxis causes serious infringement on passengers' interests and a city's reputation.

Example: Reported from the New York Taxi and Limousine Commission, in the past two years (from 2009 to 2011), $8.3 million has been overcharged because a large number of taxi drivers in New York City swindled money out of customers through arbitrarily modifying the taximeter. According to the investigation, overcharging behaviors happened in 1.8 million services out of 3.61 hundred million in total, which brought $4.45 loss in each service for each costumer.

In fact, taxi fraud is not a special phenomenon that is only existing in the U.S. but one that is always happening in the urban transportation. In China, for example, taxi fraud is also a severe problem for the city

Manuscript received January 4, 2013; revised April 8, 2013 and May 30, 2013; accepted July 8, 2013. Date of publication July 11, 2013; date of current version January 13, 2014. This work was supported in part by the Singapore National Research Foundation under its International Research Center at Singapore Funding Initiative and administered by the Interactive Digital Media Program Office, by Huawei Corporation under Grant YBCB2009041-27, and by the University Transportation Center under Grant DTRT12-G-UTC11 from the U.S. Department of Transportation. The review of this paper was coordinated by Prof. A. Jamalipour.

S. Liu and R. Krishnan are with Heinz College, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: siyuan@cmu.edu).

L. M. Ni is with The Hong Kong University of Science and Technology, Kowloon, Hong Kong.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2272792

traffic management. Usually, the more complex the traffic situation of a city, the more taxi fraud. Thus, the threat of taxi fraud is particularly rampant in those cities with a huge population and complicated road networks, e.g., Beijing, Shenzhen, and New York. Authorities nowadays have paid a lot of attention and effort to this practical and serious problem. For example, to prohibit taxi fraud, the local authorities adopt extensive random-selective examination periodically. However, this scheme takes tremendous human resources and economic costs. Moreover, the selective examination is not efficient and effective.

Fortunately, the GPS device is equipped in the taxi, reporting taxis' traces to the companies or authorities. Based on the taxis' traces and the digital road map, hopefully, we can retrieve the reported trajectory and compare it with the actual trajectory to detect the overcharged distances of the taxis. Such distance-based methods are based on map matching [1]–[6] and distance-based clustering [7]–[11]. However, there exist notable location errors and extremely limited reports in the GPS localization data [12], [13]. Such practical issues make it hard for us to retrieve accurate trajectories and then conduct taxi fraud detection in an effective way. At the same time, the complex road networks and dynamic information in the geographical data make the previous methods face big challenges in scalability, data storage, and updating.

Based on careful investigation of taxi fraud behaviors, we noticed that fraud taxi drivers always modify the taximeter to a smaller scale, which means that the taximeter records a longer distance than the actual distance. As a result, the reported speed of the taxi is higher than the actual speed. At the same time, fortunately, the reported speed from the GPS record on a taxi is accurate because it is directly collected from each vehicle's speedometer, not from the GPS measurement [12]. Hence, taking advantage of the given interesting observations on the relationship between the speed information and the fraud behavior, we present a novel method to model taxis' driving behaviors and learn this model from a large-scale real-life data set. First, we propose a new speed-based clustering model to capture the taxi's behaviors from the GPS data. Second, we derive a fraud probability model to detect taxi fraud. Finally, we develop a system, which is called the Speed-based Fraud Detection System (SFDS). Our new system is trajectory-free and map-free for taxi fraud detection.

Our major contributions are summarized as follows.

1) We are the first to model taxi fraud behaviors in a trajectory-free and map-free scenario and construct a model by the speed information instead of the location or distance.
2) We propose a novel speed-based clustering to model taxis' driving behavior, and then, we design a probability model to detect taxi fraud.
3) We evaluate our well-designed system by both synthetic and large-scale real data. The empirical experiment results show that our methods perform better than the state of the art in terms of scalability, efficiency, and effectiveness. Now, our models are already deployed as a taxi fraud system (SFDS) in a city.

The rest of this paper is organized as follows: Section II discusses the related work. The problem and some terminologies will be formally defined in Section III. In Section IV, we will introduce the speed-based clustering model and a probability model to detect taxi fraud. Section V presents the experimental evaluation results. Section VI reports our interesting findings and discussions. Finally, this paper is concluded in Section VII.

II. Related Work

A. Map Matching

Generally, map-matching approaches can be classified into three categories, i.e., local map matching (or incremental map matching) [14]–[16], global map matching [6], [17], [18], and statistical map matching [19], [20]. In local map matching, only a small part of the trajectory that is close to a given position is matched. The local map-matching algorithm proposed in [15] evaluates the candidate edges, judging by two similarity metrics, namely, distance similarity and orientation similarity. Reference [14] assigns confidence values for different sampling points and matches segments in the order of descending confidence. In general, local map-matching algorithms share a common character that the performance in terms of running time and matching accuracy significantly depends on a high sampling rate (e.g., one report per 2–5 s). Otherwise, as the sampling rate decreases, the problem called "arc skipping" described in [15] becomes severe, causing serious degradation in accuracy. Unfortunately, in our scenario, the GPS report rate is one report per certain minutes (1–30 min). Global map matching focuses on matching the entire trajectory with the road network. Reference [18] employs an offline snapping method that aims to find a minimum weight path based on edit distance. While [6] and [17] rely on *Fréchet* distance or its variants. *Fréchet* distance can fulfill trajectory comparison since it takes the continuity of curves into account. The global map-matching algorithm in [17] implements a parametric search over all critical values by finding a monotone path in the free space from the lower left corner to the upper right corner. In addition, this work is extended in [6], in which the average *Fréchet* distance is applied to reduce the effect of outliers. Statistical map matching is based on the Bayesian classifier incorporating a hidden Markov model [20] to model the topological constraints of the road network. An enhanced method based on the extended Kalman filter and cubic spline interpolation is proposed in [19]. Statistical methods amend the errors brought by GPS measurement. However, the computational cost is very high, and scalability is very low.

B. Moving-Objects Modeling

Clustering analysis is a basic model for analyzing moving objects. Li *et al.* [15] proposed a micro moving cluster, which denotes a group of similar objects, both at the current time and at a near future time. Lee [21] presented a trajectory clustering framework in which all the trajectories had been partitioned into segments. This partition scheme improves the clustering accuracy while the computational cost requires for more. After that, Lee [22] proposed an outlier detection framework relying on segmental trajectory clustering. Such trajectory clustering or time-series clustering methods [8]–[10] usually focus on moving objects' location information. Hollmén *et al.* [23] studied spatio-temporal road condition forecasting by Markov chains and artificial neural networks. Castro *et al.* [24] proposed a method to construct a model of traffic density based on taxi traces. Bacon *et al.* [25] tried to use real-time road traffic data to evaluate congestion. Yuan *et al.* [26] presented a Cloud-based system computing customized and practically fast driving routes for an end user using (historical and real time) traffic conditions and driver behavior. Zheng *et al.* [27] detected flawed urban planning using the GPS trajectories of taxicabs traveling in urban areas. Yoon *et al.* [28] tried to detect traffic conditions on surface streets given location traces collected from on-road vehicles, such as GPS location data, and infrequent low-bandwidth cellular updates. In our problem, the location is not accurate, and at the same time, the complex road networks also make the current methods not work well.

III. Problem Definition

Here, we formally define the taxi fraud problem.

Definition 1: Tracking record: Given taxi i, its tracking record is defined as vector $L_i^t = \{i, t, l, v, d, s\}$, where $L_i^t.i$ is the taxi ID, $L_i^t.t$ is the record time, $L_i^t.l$ is the location (longitude, latitude) of the taxi at

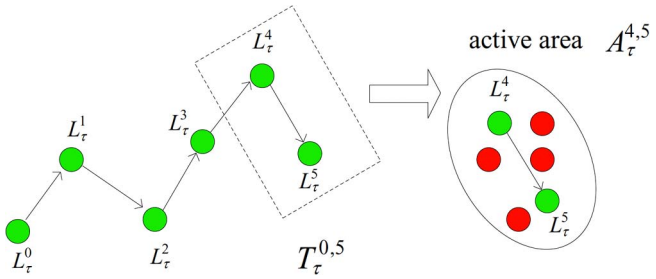

Fig. 1. Trajectory record and active area.

t, $L_i^t.v$ is the instant speed of the taxi at t, $L_i^t.d$ is the moving direction of the taxi at t, and $L_i^t.s$ is the occupied status of the taxi at t.

Definition 2: Trajectory record: Given taxi i, its trajectory record is defined as a sequence of its tracking records $L_i = [L_i^0, L_i^1, \ldots, L_i^t]$ (as shown in Fig. 1). The trajectory record is abbreviated as trajectory for convenience in our paper.

Definition 3: Task: Given taxi i, its task is defined as vector $K_i = \{i, bt, et, at, d, f\}$, where $K_i.i$ is the ID of taxi i, $K_i.bt$ is the begin time of task K, $K_i.et$ is the end time of task K, $K_i.at$ is the waiting time in task K, $K_i.d$ is the distance of task K, and $K_i.f$ is the fee of task K. The total duration of task K_i is $(K_i.et - K_i.bt)$. Note that waiting time is a user-defined parameter.

Definition 4: Average speed: Given task K_i of taxi i, the average speed v_{avg} of task K_i is derived by

$$v_{\text{avg}} = \frac{K_i.d}{(K_i.et - K_i.bt) - K_i.at}. \tag{1}$$

The intuition to define an average speed instead of an instant speed to detect taxi fraud is after each task with a passenger (passengers) is finished, we conduct fraud detection based on this task report. Such design is robust to location errors and report latency.

Definition 5: Active area: Given a trajectory record L_i of taxi i, the active area $A_{t,t+1}$ is defined by a pair of temporal consecutive tracking record L_i^t and L_i^{t+1}, where $[L_i, L_i^{t+1}] \in L_i$.

An active area $A_{t,t+1}$ is an area where taxi i possibly moves during time $L_i^t.t$ to $L_i^{t+1}.t$. We illustrate an example of the active area in Fig. 1. The ellipse-shaped circle demonstrates the active area defined by L_i^4 and L_i^5 (green dots); the red dots denote other taxis in $A_{4,5}$ during the time of $L_i^4.t$ to $L_i^5.t$.

Definition 6: Suspicion: Given task K_i and its trajectory record L_i, $\forall L_i^t \in L_i$, $K_i.bt \le L_i^t.t \le K_i.et$, the suspicion of any active area $A_{t,t+1}$ of trajectory record L_i is defined by suspicion function $Sus(L_i^t, L_i^{t+1})$. The suspicion describes the possibility for average speed v_{avg} of task K_i being an outlier in active area $A_{t,t+1}$ after speed-based clustering is implemented.

Definition 7: Fraud probability: Given task K_i and its trajectory record L_i, the fraud of K_i is decided by each suspicion of the active area formed by tracking records in L_i. Formally, the fraud of K_i is defined by fraud function $Fr(K_i)$. The fraud of task K_i depicts the probability of taxi i being a fraudulent taxi judging by i's performance in K_i.

Taxi Fraud Detection: Now, the problem of taxi fraud detection can be defined as follows.

Given task K_i of taxi i, we evaluate its fraud probability for i. Moreover, given a set of tasks and taxis, we return the fraudulent taxis.

The problem is shown in Fig. 2, where trajectory record L_i of task K_i (in red line) contains four tracking records $L_i = \{L_i^0, L_i^1, L_i^2, L_i^3\}$, and the three ellipse-shaped areas with a blue edge demonstrate the active areas of L_i. Moreover, each active area has a suspicion function.

The fraud probability of K_i is defined by three suspicion functions, namely, $Sus(L_i^0, L_i^1)$, $Sus(L_i^1, L_i^2)$, and $Sus(L_i^2, L_i^3)$.

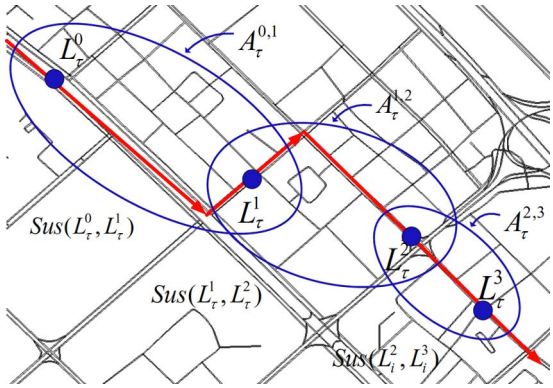

Fig. 2. Speed-based clustering.

IV. Taxi Driving Behavior Modeling

Here, we introduce our taxi driving behavior model via a novel speed-based clustering method. The framework of our model is shown in Fig. 3. First, the task and GPS records are collected. Second, the taxi driving behavior model is constructed by speed-based clustering. Third, the error estimation model is utilized in the taxi driving behavior model to compute the suspicion function. Finally, the fraud probability of the taxi is output. Our speed-based clustering model can be described as follows: The v_{avg} of K_i is put into each active area along L_i to the clustering. The fraud probability is obtained from the clustering result in each active area.

A. Active Area Construction

The GPS data point for each taxi is reported with a period set by the taxi driver. There are more than 70% of taxis reporting their tracking records to the data center more frequently than every 100 s. While in Fig. 4, the cumulative distribution function (cdf) of the task duration (September 14–20, 2009) shows that more than 80% of the tasks lasted longer than 5 min. Statistically, for most tasks, five or more tracking records can be found in the taxi track record.

Although in most cases, several tracking records are contained in a trajectory record of a task, the movement information between any two consecutive points is still missing due to the sampling error. To query a task as continuous as possible, we propose an ellipse model to define the active area. In such an active area, the movement information on a target object can be inferred by its neighbors.

For any two consecutive tracking records of the trajectory, its in-between movement is missing due to the sampling error. In other words, given two tracking records, i.e., L_i^t and L_i^{t+1}, from trajectory L_i of task K_i and theoretical maximum speed $V_{\max}$, the area in which taxi i can move from t to $t+1$ is determined. In this paper, we take this area as the active area [as shown in Fig. 6(a)].

In a city, each road is limited by a maximum speed. We observe our data in the city from September 14 to 20, 2009 to discover the speed distribution. In Fig. 5, the cdf of the speed distribution indicates that $V_{\max} = 100$ km/h is a proper value since more than 99% of taxis are with a lower speed. As the urban street is usually bidirectional, in case objects moving toward different directions would be covered by one active area, the maximum road (50 m) width $W_{\max}$ is defined as the upper bound of b. The ellipse model for constructing the active area is described in Algorithm 1. The major radius is the longest distance that the object can possibly travel during the time interval between j and $j+1$. The minor radius is the physical distance between L_t^j and L_t^{j+1}. It is computed by the Euclidian distance (ED) function.

In Algorithm 1, the continuous active areas for a trajectory of any task are constructed. In the process of constructing the active area

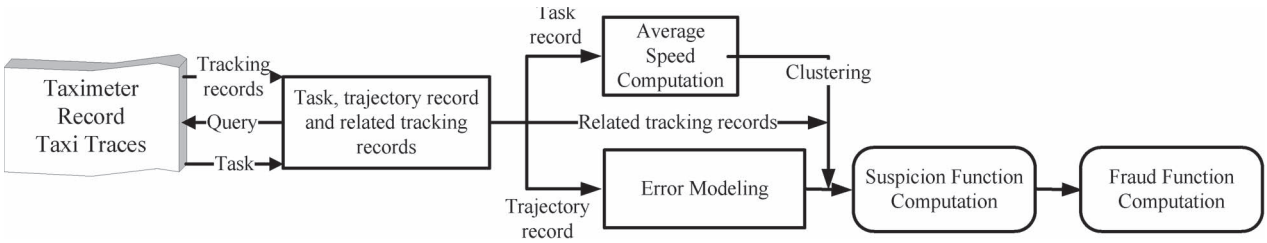

Fig. 3. Framework of the speed-based clustering model.

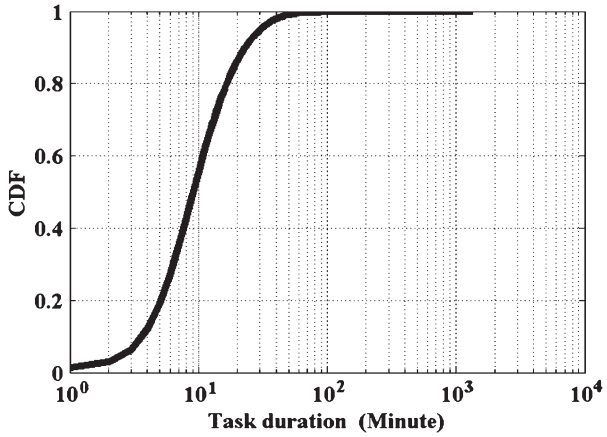

Fig. 4. Task duration, September 14–20, 2009.

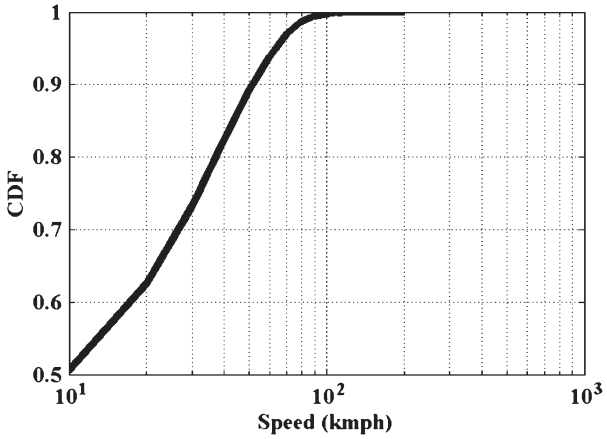

Fig. 5. Speed distribution, September 14–20, 2009.

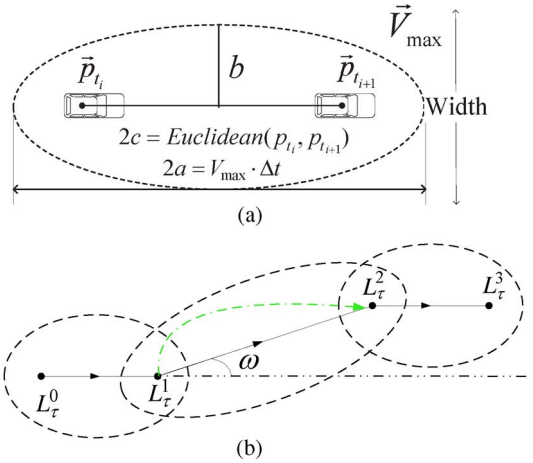

Fig. 6. (a) Active area. (b) Active area sequence.

for two consecutive tracking records L_t^j and L_t^{j+1}, there is mainly a twofold consideration. The first one is to derive minor radius b of the ellipse by the actual distance $ED(L_i^j, L_i^{j+1})$ and the possible maximum distance $V_{\max} \cdot (L_i^{j+1}.t - L_i^j.t)$ between the two tracking records. The second one is if the derived b exceeds the maximum width $W_{\max}$ of the road; then, major radius a will be set as $W_{\max}$ and the actual distance $ED(L_i^j, L_i^{j+1})$. In practice, the width $W_{\max}$ of the road is a clear constraint that is better than the assumed maximum speed $V_{\max}$. For a trajectory [as shown in Fig. 6(b)] that consists of four tracking records, every couple of its consecutive tracking records specifies an active area. The green dashed line shows the actual trace between L_i^1 and L_i^2. Although the ellipse model takes the black line as the distance between the two tracking records, the real trace is still completely covered by active area $A_{1,2}$.

Algorithm 1 Active Area Construction

Input: a task's trajectory $T_\tau^{0,k} = [L_\tau^0, L_\tau^1, \ldots, L_\tau^k]$; the maximum speed $V_{\max}$; the maximum road width $W_{\max}$.
Output: Active Areas of the task
begin
a: the major radius; b: the minor radius; 2c: the focal length;
for $j = 0$ **to** $k - 1$
  pick the consecutive pair of GPS data L_τ^j and L_τ^{j+1};
  $a = (V_{\max} \cdot (t_\tau^{j+1} - t_\tau^j))/2$;
  $c = (ED(L_\tau^j, L_\tau^{j+1}))/2$;
  $b = \sqrt{a^2 - c^2}$;
  if $b > W_{\max}$ **then**
    $a = W_{\max}$;
  else
    $b = \sqrt{a^2 - c^2}$;
  end if
  $\omega = \arctan(|lat_\tau^{j+1} - lat_\tau^j| / |lon_\tau^{j+1} - lon_\tau^j|)$;
  $\begin{cases} x = a\cos(t)\cos(\omega\pi) - b\sin(t)\sin(\omega\pi) \\ y = a\cos(t)\sin(\omega\pi) + b\sin(t)\cos(\omega\pi) \end{cases}$,
  $t \in (0, 2\pi)$;
end for
return active areas of the task.

We use n to denote the number of tracking records in trajectory L_i for task K_i, and m is the number of tracking records for objects in the active areas of L_i. The time complexity of the speed-based clustering algorithm is $O(nm + n)$. Note that our algorithm does not depend on the number of road segments, which is more efficient than the trajectory/map-based methods.

B. Speed-Based Clustering

Since we can define the active area for each pair of neighboring tracking records, for any task whose reliability we want to evaluate, we aim to derive the suspicion function for each active area via speed-based clustering. The average speed v_{avg} of a task is put into each

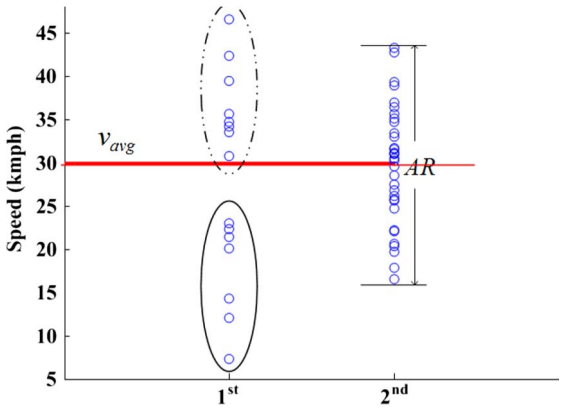

Fig. 7. Two active areas with $v_{\text{avg}} = 30$ km/h.

active area of the task's trajectory to measure its distance to the speed of other objects' tracking records. Particularly, the distance of two objects' speed is measured by the distance function as follows.

Given task K_i and its trajectory $L_i = [L_i^0, L_i^1, \ldots, L_i^t]$, where $K_i.bt \leq L_i^0.t < L_i^t.t \leq K_i.et$. For any active area $A_{j,j+1}$ defined by L_i^j and L_i^{j+1}, where L_i^j, $L_i^{j+1} \in L_i$, the distance between the speed of a tracking record $L_\tau^{j+\delta}(L_i^j \leq L_\tau^{j+\delta} \leq L_i^{j+1})$ for taxi τ in $A_{j,j+1}$ and v_{avg} of K_i is computed as

$$Dist\left(L_\tau^{j+\delta}.s, v_{\text{avg}}\right) = \left(L_\tau^{j+\delta}.s - v_{\text{avg}}\right)^2. \tag{2}$$

Active area $A_{j,j+1}$ is partitioned by v_{avg} into two clusters (as shown in Fig. 7). One includes tracking records within the speed above v_{avg} (in dotted circle), whereas the other includes tracking records within the speed under v_{avg} (in solid circle). We term the distance from v_{avg} to the speed of all the tracking records in a cluster as the dissimilarity between v_{avg} and cluster O. The partial dissimilarity function of v_{avg} and the below cluster O_u can be deduced as

$$Diss'(O_u, v_{\text{avg}}) = \sqrt{\frac{\sum_{t=1}^{m}\left(L_\tau^{j+\delta}.s - v_{\text{avg}}\right)^2}{m}} \tag{3}$$

where m is the total number of tracking records in the below cluster. The suspicion of $A_{j,j+1}$ increases with $Diss'(O_u, v_{\text{avg}})$. That is, the more different v_{avg} is from O_u, the more unreliable v_{avg} is. Therefore, the formal dissimilarity between v_{avg} and O_u is

$$Diss(O_u, v_{\text{avg}}) = m\sqrt{\frac{\sum_{t=1}^{m}\left(L_\tau^{j+\delta}.s - v_{\text{avg}}\right)^2}{m}}. \tag{4}$$

For the entire $A_{j,j+1}$, the proportion taken by the below cluster indicates the suspicion of v_{avg} in $A_{j,j+1}$. Thus, the suspicion function augments with the increasing v_{avg}. In other words, the higher v_{avg} is, the more tracking records would be classified into the below cluster. We define an average radius (AR) for $A_{j,j+1}$. The AR measures the dissimilarity between v_{avg} and the speed of all tracking records in $A_{j,j+1}$. According to (4), the dissimilarity is defined by

$$Diss(O, v_{\text{avg}}) = l\sqrt{\frac{\sum_{t=1}^{l}\left(L_\tau^{j+\delta}.s - v_{\text{avg}}\right)^2}{l}} \tag{5}$$

where l is the total number of tracking records in $A_{j,j+1}$, and O denotes the entire cluster in $A_{j,j+1}$. The AR actually describes the compactness of the speed distribution around v_{avg} in $A_{j,j+1}$. The v_{avg} is worthy-trusted in $A_{j,j+1}$ when $A_{j,j+1}$ is compact.

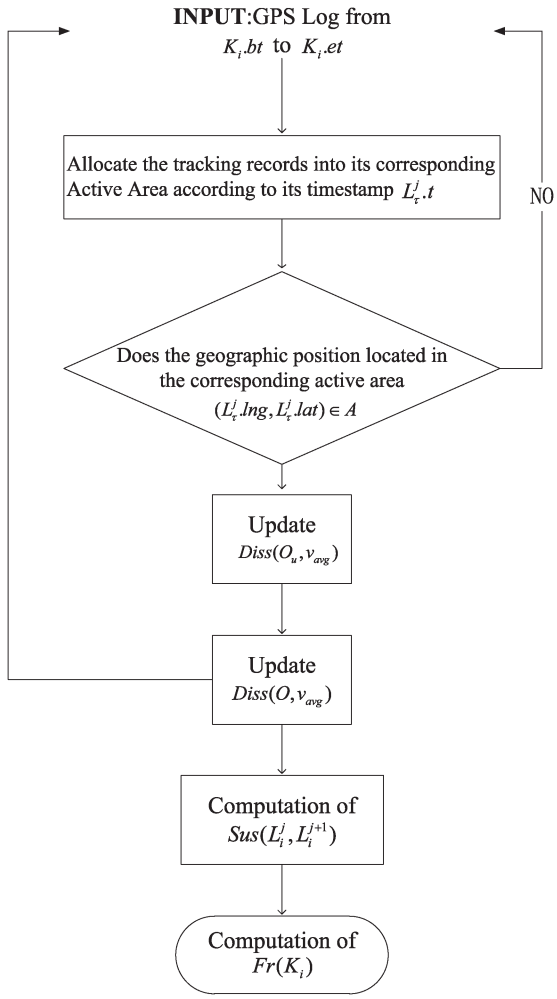

Fig. 8. Workflow of speed-based clustering.

Combining (4) and (5), the suspicion function for $A_{j,j+1}$ is defined as

$$Sus\left(L_i^j, L_i^{j+1}\right) = \frac{Diss(O_u, v_{\text{avg}})}{Diss(O, v_{\text{avg}})}. \tag{6}$$

For task K_i with its trajectory $L_i = [L_i^0, L_i^1, \ldots, L_i^t]$, where $K_i.bt \leq L_i^0.t < L_i^t.t \leq K_i.et$, the fraud function of K_i is defined as

$$Fr(K_i) = \sum_{j=0}^{t-1} Sus\left(L_i^j, L_i^{j+1}\right). \tag{7}$$

The basic procedure of the speed-based clustering model is summarized in Fig. 8.

V. System Evaluation

Here, we evaluate our speed-based clustering model by synthetic and real-life data. First, we describe the experiment including the experiment environment, data sets, and comparison methods. Second, we evaluate the robustness of the methods. Finally, we evaluate the accuracy and efficiency of the methods.

A. Experiment Setup

Data Set Description: We employ two different data sets to conduct the experiments, i.e., synthetic data set and real data set.

Synthetic Data Set: The synthetic data set used in our experiments consists of average speed v_{avg}, trajectory L_i of the task, and the related tracking records in active areas along L_i. To observe the outcome of the suspicion function, v_{avg} is assigned with the value from 10 to 80 km/h. Trajectory L_i is assigned with two motion types, i.e., the uniform motion and the variable motion. In the uniform motion, we simulate L_i to be within a constant speed all along its trip. While in the variable motion, we simulate the speed of each tracking records in L_i varying in a certain pattern, which is predefined. The number of tracking records for other objects in each active area follows a normal distribution $O(\mu, \sigma^2)$. The speed distribution for tracking records in each active area follows a normal distribution $S(\mu, \sigma^2)$.

Real Data Set: The data set we utilize in this paper is collected from a large city in China. The data are from August 1, 2009 to September 30, 2009 and consist of the taxi tracking record and the taximeter record. The taxi tracking record stores the driving status of a taxi, such as the taxi ID, instant location, time stamp, instant speed, and occupied status. The taximeter record stores the operation information on a taxi, such as the taxi ID, the trip fee, the trip distance, the trip start/end time, and the waiting time during the trip. The city contains 44 793 road segments and 10 208 taxis. To clarify, we use a task to denote a trip served by a taxi. An operation record would be generated to the taximeter record as long as a taxi finishes a task.

According to the result of the investigation on the illegal overcharging of taxi fee in August 2009, 320 taxis were identified as fraud taxis, which are taken as the ground truth in our experiments.

Experiment Environment: A server with four Intel Core Quad CPUs, Q9550 2.83 GHz, and 32-GB main memory.

Baseline Methods: We evaluate our fraud detection system, which is called the SFDS, with the map-matching method (ST-matching) in [29], UMicro, one of the latest representative method for clustering uncertain data streams [7] and a taxi driving fraud detection system [11]. (We also conducted similar clustering to identify 50 sites as in [11] and in our work; to simplify, we call it the FDS.)

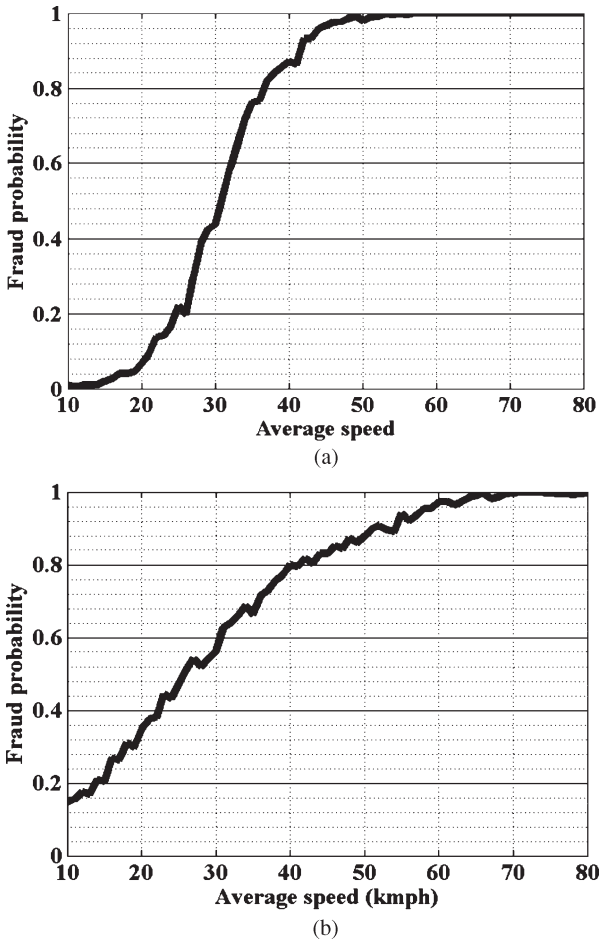

Fig. 9. Fraud detection in different driving behaviors. (a) Uniform motion. (b) Variable motion.

B. Scalability Evaluation and Analysis

There are three parameters in our speed-based clustering model, namely, the speed distribution in each active area $S(\mu, \sigma^2)$, the distribution for the number of objects (related tracking records) in each active area $O(\mu, \sigma^2)$, and the number of tracking records for a trajectory record of a task, which is denoted by N. We conducted the experiments by synthetic data, respectively. The robustness of our model is measured by the variation of the output with the varying parameters.

The first parameter is the speed distribution $S(\mu, \sigma^2)$. The speed distribution in each active area is related to the realistic traffic condition. Generally speaking, the better traffic in one active area, the higher speed for most taxis in that active area. For a fraudulent task K_i, its average speed v_{avg} computed by the taximeter is always higher than its actual speed no matter how the traffic condition changes in the task. To observe the outcome of the fraud function given by different v_{avg}, we preset the number of active areas for task K_i to 20, that is, 21 tracking records are contained in trajectory L_i for K_i. In addition, the number of taxis in each active area for K_i follows a normal distribution $O(\mu, \sigma^2)$, where μ is set to 20, and σ is 10. For each different average speed in Fig. 9(a), we first assume that L_i obeys a uniform motion. The speed of each tracking record in L_i is constant. The speed is set to 30 km/h. The speed distribution in each active area for a task is described as a normal distribution $S(\mu, \sigma^2)$, in which μ is all set to 30, and σ is 10. The average speed v_{avg} varies from 10 to 80 km/h, as shown in the x-axis in Fig. 9(a). The fraud probability reaches to 1 from $v_{\text{avg}} = 45$ km/h. To describe a more practical case, we also generate another type of motion (the so-called variable motion) for L_i. The speed distribution is $S(\mu, \sigma^2)$, where μ varies in each active area. For any active area $A_{j,j+1}$ in L_i, its corresponding $\mu_{j,j+1}$ is defined as

$$\mu_{j,j+1} = \frac{L_i^j.s + L_i^{j+1}.s}{2} \tag{8}$$

where $L_i^0.t \le L_i^j.t < L_i^{j+1}.t \le L_i^{21}.t$. The fraud probability varies with v_{avg}, as shown in Fig. 9(b). In Fig. 9, we can conclude that there is a higher probability that there exists fraud if the speed is over 35 km/h.

The second parameter is the number N of active areas for task K_i. This is decided by two factors, i.e., the sampling rate of a taxi and the task duration. We have studied the sampling rate and task duration distribution in the previous section. To obtain the result purely under the impact of N, we first assume that the tracking records in trajectory L_i for K_i are within a constant speed as 30 km/h. The speed distribution in each active area still follows a normal distribution $S(\mu, \sigma^2)$, where $\mu = 30$, and $\sigma = 10$. Additionally, the number of objects in each active area obeys the normal distribution $O(\mu, \sigma^2)$, where $\mu = 20$, and $\sigma = 10$. The fraud probability for each task with the different v_{avg} is shown in Fig. 10(a). The x-axis denotes the number of active areas included in K_i. We can see that the fraud probability varies with the increasing number of active areas slightly, which indicates that our speed-based clustering model is stable when dealing with the task consisting of different number of active areas.

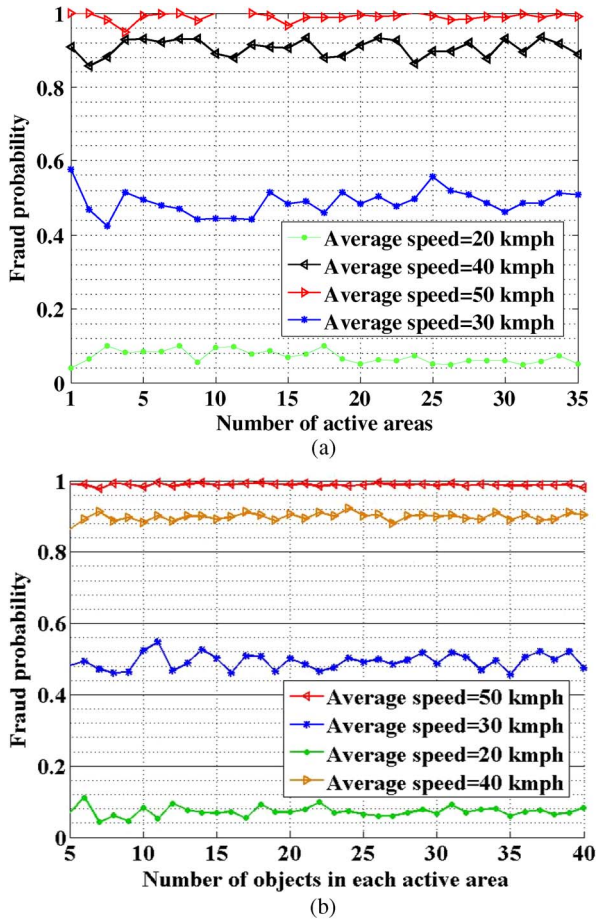

Fig. 10. Scability of fraud detection. (a) Number of active areas. (b) Number of objects.

The third parameter is the distribution for number of objects $O(\mu, \sigma^2)$ in each active area. On one hand, it depends on the realistic situation (e.g., road segment and time interval). On the other hand, the number of objects in each active area relates to the size of the active area. However, at least each active area should cover two tracking records that define the active area. We still first assume that the tracking records in trajectory L_i for K_i are within a constant speed as 30 km/h. The speed distribution in each active area is described as the normal distribution $S(\mu, \sigma^2)$, where $\mu = 30$, and $\sigma = 10$. Moreover, the number of objects in each active area varies from 5 to 40 in each task, as shown by the x-axis in Fig. 10(b). Given the different average speed v_{avg}, the fraud probabilities of all the tasks are shown in Fig. 10(b). Within the comparably low v_{avg}, such as $v_{\text{avg}} = 20$ km/h or $v_{\text{avg}} = 30$ km/h, the confidence would be faintly infected by the small number of objects in each active area. While for a higher v_{avg}, the fraud probability stays steady without the influence of the number of objects in the active area. Thus, our model is also resistant to the variation of the number of objects for each active area.

Based on our analysis on those three parameters in the speed-based clustering model, we can see that our model is robust to both the variation of the number of active areas for a task and the number of objects in the corresponding active area. The robustness of our model is an important feature when we implement the speed-based clustering on the real data.

C. Accuracy Evaluation and Analysis

We set $V_{\max} = 100$ km/h and the road width as 50 m for the ellipse model. According to the city's investigation on local taxi management,

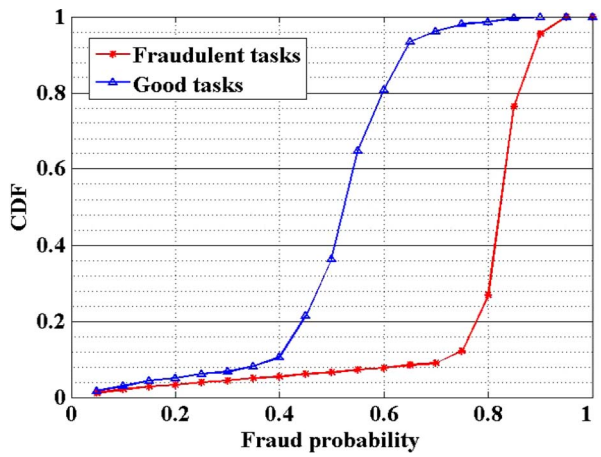

Fig. 11. Fraud probability: fraudulent task versus good task.

TABLE I
ACCURACY EVALUATION

Approach	Precision	Recall	F-score
SFDS	0.748	0.820	0.782
ST-matching	0.012	0.015	0.013
UMicro	0.232	0.137	0.172
FDS	0.423	0.328	0.369

320 taxis in the city have been identified as illegally overcharging taxis in August 2009. We implement our model on the taxi track record and the taximeter record of them. The 320 fraudulent taxis totally took 7041 fraudulent tasks in that month. Thus, we derive the fraud probability for the tasks by the 320 taxis and give the cdf of the fraud probability. More than 95% of the fraudulent tasks are within a probability larger than 0.7. The fraud probability of good tasks is used to compare with the 320 identified fraudulent taxis. The cdf of the fraud probability for good tasks is shown in Fig. 11 in blue line. We can see that the fraud probabilities for most tasks (95%) are under 0.7. In addition, combining the two cdf curves in Fig. 11 into consideration, the distribution of fraudulent tasks and that of good tasks can be evidently distinguished. Therefore, we choose 0.7 as the threshold for taxi fraud detection. The probability threshold is defined usually depending on user's specific requirement. The fraud detection result based on this threshold (0.7) is reported below.

In our method, we utilize the threshold derived from the one week (September 14–20, 2009) data to detect taxi fraud in August 2009. The accuracy is measured by precision, recall, and F-score. Precision means the percentage of detected tasks by fraudulent taxis in the detected tasks. Recall means the percentage of detected tasks by fraudulent taxis in all true fraud tasks. F-score is the weighted harmonic mean of precision and recall. There are totally 7401 fraudulent tasks conducted by the 320 fraudulent taxis in August 2009. We implement the map-matching method and UMicro to reconstruct the traces for the 7041 tasks via their corresponding tracking records. The distance of those tasks' reconstructed traces is contrasted to the distance collected from the taxis taximeters. Tasks with an error out of $[-1\%, 4\%]$ are defined as the fraud tasks. Based on the whole taxi traces data, we conduct the fraud detection accuracy evaluation. Table I lists the experiment results of our speed-based clustering, the ST-matching method, Umicro, and FDS. The results show that our model achieves better accuracy than the state of the art. The main reason for this performance is that in a large city, it is hard for us to retrieve accurate trajectories from the taxis' GPS reports and update the maps all the time. (There are a number of road changes in the city that can hardly be captured by digital map or GPS reports.)

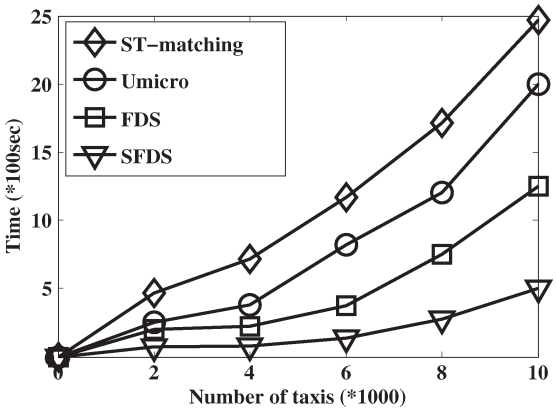

Fig. 12. Efficiency evaluation.

D. Efficiency Evaluation and Analysis

As previously discussed, we conducted the fraud detection accuracy evaluation. At the same time, we also record the time cost of a system. The time cost is defined as the running time of a system to cost to run though all the taxi traces. In Fig. 12, we illustrate the experiment results after we test 10 000 traces. Based on the running results, we can find out that our method, i.e., SFDS, performs better than the current methods. The main time costs of other methods are from constructing trajectories or retrieving locations, which is definitely very time consuming; however, our method has no such concerns. Hence, our method scales very well along the number of taxis.

VI. Findings and Discussions

Here, we provide the following findings about spatial and temporal features about taxi fraud behaviors.

Spatial features: In our detected fraud behaviors, 90.7% of the total fraud behaviors happened in transportation intersections (airports, railway stations, and ports), particularly when there are extremely many passengers in a given location and time (crowded areas in the city). It is interesting that among all the given fraud behaviors, 60.3% of the trips were ended in hotels, and 20.4% of the trips were ended in passengers' next transportation intersections. Because we cannot access passengers' detailed information (e.g., the location of a passenger's home, whether a passenger is visiting this city for a short trip, etc.), it is hard for us to give insights or solid explanations to the correlation between passengers' features and taxi's fraud behaviors; however, our guess is that the fraud taxis prefer to fraud foreign passengers, that is, tourists. On the other hand, among all the trips to resident areas, only 0.8% of the fraud behaviors happened.

Temporal features: In our detected fraud behaviors, 47.2% of the total fraud behaviors happened during rush hours (in our study city, from 8 A.M. to 10 A.M. and from 5 P.M. to 8 P.M. are defined as rush hours by the city transportation authority), and 34.7% happened from 11 P.M. to 2 A.M. On the other hand, 68.9% of the detected fraud behaviors happened on a Friday. Hence, the fraud behaviors happen not only in the time period when there are crowded people but also during late night. Our guess is that this is because fraud behaviors are mostly due to change in the taximeter wheels, and during rush hours, it is hard for passengers to notice taxi's speed changes, and the low-speed charging fee also gives chances to such fraud behaviors. Currently, it is hard for our method and the system to interpret such findings.

VII. Conclusion and Future Work

In this paper, we have studied taxi fraud detection in a trajectory-free and map-free scenario. Due to the challenges from massive errors in the position data and complex city road networks, the traditional trajectory or map-based methods do not work well. We have proposed a novel taxi driver behavior model based on a speed-based clustering method. Our method was evaluated by not only the synthetic data but also the real data. Our methods are better than the state of the art in terms of scalability, accuracy, and efficiency.

In the future work, we will try to propose more accurate speed estimation models and utilize our driving behavior model to help taxi drivers with decision making. At the same time, as our model is deployed as a taxi fraud system in a city, we are going to collect more user feedback data to make our system able to solve real problems in the city. Our methods can be also used in other moving objects related research and applications.

References

[1] D. Pfoser and C. S. Jensen, "Capturing the uncertainty of moving-object representations," in *Proc. 6th SSD*, 1999, pp. 111–132.
[2] H. Zhu, Y. Zhu, M. Li, and L. Ni, "SEER: Metropolitan-scale traffic perception based on lossy sensory data," in *Proc. IEEE INFOCOM*, 2009, pp. 217–225.
[3] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-matching for low-sampling-rate GPS trajectories," in *Proc. 17th ACM SIGSPATIAL*, 2009, pp. 352–361.
[4] G. Taylor, C. Brunsdon, J. Li, A. Olden, D. Steup, and M. Winter, "GPS accuracy estimation using map matching techniques: Applied to vehicle positioning and odometer calibration," *Comput. Environ. Urban Syst.*, vol. 30, no. 6, pp. 757–772, Nov. 2006.
[5] J. Yang, S. Kang, and K. Chon, "The map matching algorithm of GPS data with relatively long polling time intervals," *J. Eastern Asia Soc. Transp. Studies*, vol. 6, pp. 2561–2573, 2005.
[6] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk, "On map-matching vehicle tracking data," in *Proc. VLDB*, 2005, pp. 853–864.
[7] C. Aggarwal and P. Yu, "A framework for clustering uncertain data streams," in *Proc. 24th IEEE ICDE*, 2008, pp. 150–159.
[8] Y. Li, J. Han, and J. Yang, "Clustering moving objects," in *Proc. 10th ACM SIGKDD*, 2004, pp. 617–622.
[9] L. Li and B. A. Prakash, "Time series clustering: Complex is simpler!" in *Proc. ICML*, 2011, pp. 185–192.
[10] H. Güting, T. de Almeida, and Z. Ding, "Modeling and querying moving objects in networks," *VLDB J.*, vol. 15, no. 2, pp. 165–190, Jun. 2006.
[11] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in *Proc. 11th IEEE ICDM*, 2011, pp. 181–190.
[12] S. Liu, Y. Liu, L. Ni, J. Fan, and M. Li, "Towards mobility-based clustering," in *Proc. 16th ACM SIGKDD*, 2010, pp. 919–928.
[13] S. Liu, C. Liu, Q. Luo, L. Ni, and R. Krishnan, "Calibrating large scale vehicle trajectory data," in *Proc. 13th IEEE MDM*, 2012, pp. 222–231.
[14] S. S. Chawathe, "Segment-based map matching," in *Proc. IEEE Intell. Veh. Symp.*, 2007, pp. 1190–1197.
[15] J. Greenfeld, "Matching GPS observations to locations on a digital map," in *Proc. 81st Annu. Meet. TRB*, 2002, pp. 1–13.
[16] C. Wenk, R. Salas, and D. Pfoser, "Addressing the need for map-matching speed: Localizing global curve-matching algorithms," in *Proc. 18th SSDM*, 2006, pp. 379–388.
[17] H. Alt, A. Efrat, G. Rote, and C. Wenk, "Matching Planar Maps," *J. Algorithms*, vol. 49, no. 2, pp. 262–283, Nov. 2003.
[18] H. Yin and O. Wolfson, "A weight-based map matching method in moving objects databases," in *Proc. 16th SSDBM*, 2004, pp. 437–438.
[19] J. M. Patel, Y. Chen, and V. P. Chakka, "Stripes: An efficient index for predicted trajectories," in *Proc. 30th ACM SIGMOD*, 2004, pp. 637–646.
[20] O. Pink and B. Hummel, "A statistical approach to map matching using road network geometry, topology and vehicular motion constraints," in *Proc. 11th IEEE ITS*, 2008, pp. 862–867.
[21] J. Lee, J. Han, and K. Whang, "Trajectory clustering: A partition-and-group framework," in *Proc. 33th ACM SIGMOD*, 2007, pp. 593–604.
[22] J. Lee, J. Han, and X. Li, "Trajectory outlier detection: A partition-and-detect framework," in *Proc. 24th IEEE ICDE*, 2008, pp. 140–149.
[23] K. Sirvio and J. Hollmén, "Spatio-temporal road condition forecasting with Markov chains and artificial neural networks," in *Proc. 3rd HAIS*, 2008, pp. 204–211.
[24] P. S. Castro, D. Zhang, and S. Li, "Urban traffic modelling and prediction using large scale taxi GPS traces," in *Proc. Pervasive*, 2012, pp. 57–72.

[25] J. Bacon, A. I. Bejan, A. R. Beresford, D. Evans, R. J. Gibbens, and K. Moody, "Using real-time road traffic data to evaluate congestion," in *Proc. Dependable Historic Comput.*, 2011, pp. 93–117.
[26] J. Yuan, Y. Zheng, X. Xie, and G. Sun, "Driving with knowledge from the physical world," in *Proc. 17th ACM SIGKDD*, 2011, pp. 316–324.
[27] Y. Zheng, Y. Liu, J. Yuan, and X. Xie, "Urban computing with taxicabs," in *Proc. UbiComp*, 2011, pp. 89–98.
[28] J. Yoon, B. Noble, and M. Liu, "Surface street traffic estimation," in *Proc. 5th ACM MobiSys*, 2007, pp. 220–232.
[29] X. Li, W. Shu, M. Li, P.-E. Luo, H. Huang, and M.-Y. Wu, "Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1647–1653, May 2008.

A Layered Coalitional Game Framework of Wireless Relay Network

Xiao Lu, Ping Wang, *Member, IEEE*, and Dusit Niyato, *Member, IEEE*

***Abstract*—The wireless relay network (WRN) has recently emerged as an effective way to increase communication capacity and extend a coverage area with low cost. In the WRN, multiple service providers (SPs) can cooperate to share their resources (e.g., relay nodes and spectrum) to achieve higher utility in terms of revenue. Such cooperation can improve the capacity of the WRN and, thus, throughput for terminal devices (TDs). However, this cooperation can be realized only if fair allocation of aggregated utility, which is the sum of the utility of all the cooperative SPs, can be achieved. In this paper, we investigate the WRN consisting of SPs at the upper layer and TDs at the lower layer and present a game-theoretic framework to address the cooperation decision-making problem in the WRN. Specifically, the cooperation of SPs is modeled as an overlapping coalition formation game, in which SPs should form a stable coalitional structure and obtain a fair share of the aggregated utility. We also study the problem of allocating aggregated utility based on the concept of the Shapley value, which stabilizes the cooperation of SPs in the WRN. The cooperation of TDs is modeled as a network formation game, in which TDs establish links among each other to form a stable network structure. Numerical results demonstrate that the proposed distributed algorithm obtains the aggregated utility approximating the optimal solutions and achieves good convergence speed.**

***Index Terms*—Layered coalitional game, network structure formation, overlapping coalition formation, Shapley value.**

I. Introduction

Multihop relay transmission has potential to offer low-cost, high-speed, and long-range wireless communications. The multihop relay transmission will be adopted in the new standards (e.g., WiMAX or IEEE 802.16m). In such standards, multihop relay capability is enabled for legacy network terminal devices (TDs; e.g., access points and routers). In this context, the wireless relay network (WRN) has recently emerged as one of the focuses on next-generation wireless networks. In the WRN, relay transmission based on cooperation among TDs can mitigate channel fading and exploit spatial reuse to accommodate concurrent transmissions, thus potentially increasing network capacity and coverage. Existing works [1], [2] have demonstrated that a significant performance improvement can be achieved in the WRNs in terms of higher throughput, lower bit error rate, and better capacity, compared with traditional single-hop transmission used in the current cellular network [3].

This paper considers the multihop network structure formed by TDs from multiple service providers (SPs) with the purpose of capacity improvement. Multiple SPs may form coalitions to let their TDs serve the transmission demand of each other, with the goal of increasing the SPs' aggregated revenue [4]. The utilization of available resources, i.e., vacant TDs (which can relay the data for other TDs) and network capacity, can be substantially improved by forming SP coalitions. This is beneficial to SPs because, in some WRNs, TDs are dispersedly deployed. For each SP, if there is no cooperation, the limited number of available vacant TDs and their dispersive locations largely confine the link capacity for data transmission. The cooperation among SPs could increase the number of available vacant TDs for each SP, which, in turn, can improve network capacity. Accordingly, a larger amount of traffic demand can be served by more efficient resource utilization in the WRN, which leads to higher aggregated revenue for SPs. The cooperation in the WRN also benefits TDs because a cooperative relay may substantially lead to improved network capacity and, thus, increased throughput for the TDs. We call the formation of this interrelated SP cooperation and TD cooperation as the layered cooperation problem, which is the main focus of this paper. Previous work has also considered a game-theoretic framework with layers/hierarchies, e.g., cognitive radio networks in [5], two-tier femtocell networks in [6], and WRNs in [7]. However, most of the works considered the competition relationship between different layers, which belongs to the Stackelberg game concept [8]. In our work, different layers interact to improve the benefit of each other cooperatively.

Although cooperation in the WRN brings significant benefits, three major challenges arise. First, what is the coalitional structure that is desirable for all SPs? Second, how shall the aggregated utility (i.e., revenue) be allocated to each SP in the coalition in a fair manner so that none of the SPs is willing to leave the coalition? Third, what is the stable network structure for TDs of cooperative SPs to perform uplink transmission? This paper addresses these three challenges.

II. Layered Coalitional Game of Service Providers and Terminal Devices in Wireless Relay Network

We consider a multicell WRN consisting of multiple base stations (BSs) and TDs deployed by different SPs. Let $\mathcal{M} = \{1, 2, \ldots, M\}$ and $\mathcal{B} = \{1, 2, \ldots, B\}$ denote the sets of SPs and BSs, respectively. $\mathcal{N}^{(m)} = \{1, 2, \ldots, N_m\}$ is the set of TDs of SP m and N_m is the total number of TDs belonging to SP m, for $m \in \mathcal{M}$. We call the TD with and without transmission demand at a specific time as the source TD and the vacant TD, respectively. In the WRN, multihop relaying is used to route the flow from the source TD to the destination BS. We assume that each TD i is connected to a BS through, at most, one path, and each vacant TD can only work as a relay for one TD.

We formulate a layered game-theoretic model, which is referred to as the layered coalitional game, to model the decision-making process of cooperation between the SPs and TDs. The SPs aim to maximize their individual utility, whereas the TDs attempt to maximize their end-to-end throughput with the help of TDs from cooperative SPs. In the proposed game, SPs need to decide on the coalitional structure to form to improve their individual and aggregated utility. In addition, the TDs need to make decisions to form a relay network structure to improve throughput. The utility of an SP is defined as the revenue (earned

Manuscript received July 4, 2012; revised September 21, 2012, February 3, 2013, and July 11, 2013; accepted July 17, 2013. Date of publication July 24, 2013; date of current version January 13, 2014. The review of this paper was coordinated by Dr. S. Zhong.

The authors are with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: Luxiao@ntu.edu.sg; Wangping@ntu.edu.sg; Dniyato@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TVT.2013.2274533