A Recursive Bayesian Filter for Anomalous Behavior Detection in Trajectory Data

Hai Huang, Lijuan Zhang and Monika Sester

Abstract This chapter presents an original approach to anomalous behavior analysis in trajectory data by means of a recursive Bayesian filter. The anomalous pattern detection is of great interest in the areas of navigation, driver assistant system, surveillance and emergency management. In this work we focus on the GPS trajectories finding where the driver is encountering navigation problems, i.e., taking a wrong turn, performing a detour or tending to lose his way. To extract the related features, i.e., turns and their density, degree of detour and route repetition, a long-term perspective is required to observe data sequences instead of individual data points. We therefore employ high-order Markov chain to remodel the trajectory integrating these long-term features. A recursive Bayesian filter is conducted to process the Markov model and deliver an optimal probability distribution of the potential anomalous driving behaviors dynamically over time. The proposed filter performs unsupervised detection in single trajectory with solely the local features. No training process is required to characterize the anomalous behaviors. Based on the results of individual trajectories collective behaviors can be analyzed as well to indicate some traffic issues, e.g., turn restriction, blind alley, temporary road-block, etc. Experiments are performed on the trajectory data in urban areas demonstrating the potential of this approach.

H. Huang (✉)
Institute of Applied Computer Science, Bundeswehr University Munich,
85577 Neubiberg, Germany
e-mail: hai.huang@unibw.de

L. Zhang · M. Sester
Institute of Cartography and Geoinformatics, Leibniz University Hannover, Appelstr. 9a,
30167 Hannover, Germany
e-mail: Lijuan.Zhang@ikg.uni-hannover.de

M. Sester
e-mail: Monika.Sester@ikg.uni-hannover.de

J. Huerta et al. (eds.), *Connecting a Digital Europe Through Location and Place*,
Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-03611-3_6,

© Springer International Publishing Switzerland 2014

1 Introduction

Anomalous behavior detection refers to the problem of finding patterns in data that do not conform to expected behaviors. It is of great interest for the applications of navigation/driver assistant system, surveillance and emergency management.

The techniques employed for anomalous pattern detection in the last years are summarized in (Chandola et al. 2009) with following classes: classification based techniques, parametric or non-parametric statistical techniques, nearest neighbor based techniques, clustering based techniques, spectral techniques and information theoretic techniques. A significant number of works related to automated anomaly detection in trajectory data involve trajectory learning, i.e., cluster models of trajectories corresponding to normal cases are learned from historical trajectories and new trajectories are typically assigned an anomaly score based on the distance to the closest cluster model or likelihood of the most probable cluster model (Morris and Trivedi 2008). Hu et al. (2006) propose an algorithm for automatic learning of motion patterns and use these patterns for anomaly detection and behavior prediction. Trajectories are clustered hierarchically using spatial and temporal information and then use a chain of Gaussian distributions to present each motion pattern. Based on the learned motion patterns, statistical methods are used to detect anomalies and predict behaviors. Besides the cluster based trajectories learning method, Piciarelli et al. (2008) propose a trajectory learning and anomaly detection algorithm based on one-class Support Vector Machine. The algorithm can automatically detect and remove anomalies in the training data. They first evenly sample points from the raw trajectory and then model each trajectory with a fixed-dimensional feature. Bu et al. (2009) build local clusters using continuity characteristics of trajectories and monitor anomalous behavior via efficient pruning strategies. Ma (2009) presents a method of real-time anomaly detection for users following normal routes. Trajectories are modeled as a discrete-time series of axis-parallel constraints ("boxes") and then incrementally compared with a weighted trajectory collected from N norms.

Current approaches include (Kim et al. 2011), in which Gaussian process regression is used for the recognition of motions and activities (also anomalous events given already learned normal patterns) of objects in video sequences. Pang et al. (2011, 2013) adapt likelihood ratio test statistic to learn traffic patterns and detect anomalous behavior from taxis trajectories to monitor the emergence of unexpected behavior in the Beijing metropolitan area.

Recursive Bayesian estimation (or Bayes filter) (Masreliez and Martin 1977), e.g., the Kalman filter (Kalman 1960) for linear and normally distributed variables, are widely used in the areas of signal processing, navigation and robot/vehicle control. A main character of Bayes filters is the dynamic updating (actually two steps: prediction and updating) the estimation of the underlying variable(s) based only on the most recently acquired measurement data. Kalman filter and its extension have been proved appropriate for trajectory analysis. Recent works include (Prevost et al. 2007), which presents an extended Kalman filter to predict the trajectory of a moving object with the measurement data from a moving sensor—an unmanned aerial system (UAS). An Unscented Kalman filter is used in (Sun et al. 2012) for the trajectory tracking based on the satellite data with weak observability and inherent large initial error.

This chapter presents an original approach to anomalous behavior analysis of GPS trajectory data of vehicles. A variant of recursive Bayesian filter is proposed for a dynamic inference process. One of the original ideas is to find where the driver is meeting navigation problems, i.e., taking a wrong turn, performing a detour or tending to lose his way. Differing from most of the previous approaches:

1. The filtering is conducted for a high-level feature, i.e., the belief of behavior character, instead of the vehicle state, i.e., position and orientation.
2. The pattern detection is performed on single trajectory and no previous learning process is required to distinguish normal and anomalous behaviors.

For this purpose high-level features, i.e., (1) turns, their combination and density, (2) the degree of detour and (3) the route repetition, are required and a long-term perspective is taken to extract them from the original data. We use an extended high-order Markov chain to remodel the trajectory integrating these long-term features. The proposed recursive Bayesian estimator processes the Markov model and deliver an optimal probability distribution of the potential anomalous drive behaviors over time.

The chapter is organized as follows. In Sect. 2 we introduce the anomalous behaviors in the trajectories and the features we employed to recognize them. Section 3 presents the Markov model adapted to the trajectory data and the recursive belief filter. Experiments and results are demonstrated in Sect. 4. The chapter ends up with conclusions in Sect. 5.

2 Anomalous Behaviors and Features

In this work, we focus on the anomalous patterns in driving. In contrast to a normal drive from the start spot to the predetermined destination, anomalous behaviors may happen in many various situations, e.g., taking a wrong turn, getting lost, road-block, temporary stopover, etc. Please note that we, just like other works, also work based on the basic GPS data, i.e., time stamps and the positions. Unlike the conventional measurements of position and velocity, however, the anomalous patterns can usually not be observed at a single time point. High-level as well as long-term measurements, i.e., behavior features, are required as the "observations" of the underlying state.

2.1 Turns Combination and Density

Turning is one of the most basic movements in the trajectory data. Although a single turn is not indicating any anomalous behavior, their combination and density can

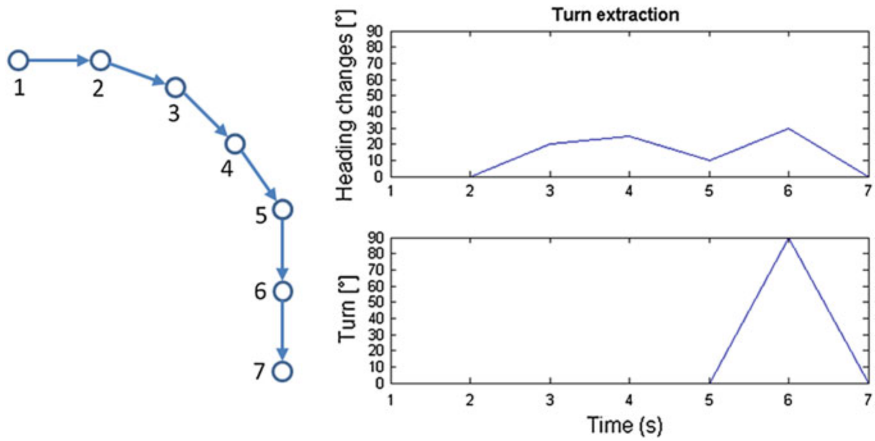

Fig. 1 Turn extraction: the accumulated heading changes are extracted as a turn movement at the last state

deliver some anomalous patterns like forming a detour/loop and unusually intensive turns. In the GPS data a turn is normally finished within several time stamps (several seconds). As shown in Fig. 1, at the first state where a turn starts the values of heading changes are counted and the turn is "marked" at the last state when the turn has been finished. We use the total absolute heading change of 40° as the threshold to determine a turn. Turns right to the previous direction is defined as positive.

In comparison with the detection of a single turn, a perspective with even longer term is taken to observe and evaluate the combination and density of multiple turns. Turns inside a given time interval (a "memory" of normally 1 to 2 min at urban driving speed) are recorded with the directions of turning. Intensive sequential same turns, e.g., double or triple left turns, have more impact on the belief of anomalous behaviors than the sequential different turns because they are implying a potential detour (see below) or the tendency of looping.

2.2 Detour Factor

Detour often happens when the driver meets traffic issues, e.g., road-blocking and traffic jam, or fails to find the correct or best way to the destination. A detour factor is conducted to quantify the degree of detour as an anomalous feature. From a start point, if the trajectory tends to go backwards, or in other words, the heading change is about 180°, it will be treated as a detour and the detour factor will be calculated for all the points in the backward segment. As shown in Fig. 2, the detour factor of an individual point is the ratio of the length of the trajectory from the start point (solid blue) and the direct distance (green) between the start point and the current position.

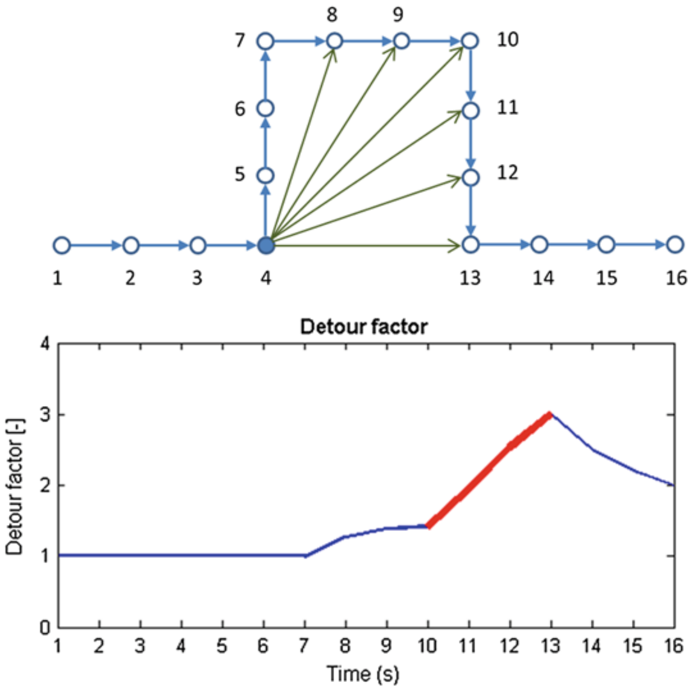

Fig. 2 Detour factor: calculated from the first turn of turn-combination of U-form. The detour factor values are only given to a certain segment (*bold*, *red*) after the heading change is accumulated to about 180° until the value decreases. A value of 1 is then given to the rest states meaning no detour

2.3 Route Repetition

The most prominent feature in a one-way trajectory is the route repetition, i.e., the driver goes back to the same road part, from either same or opposite direction, on his way to the destination. Route repetition with the opposite direction is mostly the result of performing an U-turn while that with same direction often happens after driving a loop. The current trajectory segment, i.e., between the current and the last steps, is repeating a former route when any prior trajectory segment(s) fall inside the buffer of the current segment and approximately parallel to it.

3 A Recursive Belief Filter

A Bayes filter is conducted to estimate an unobserved state, i.e., the belief of anomalous behaviors, recursively over time. The extracted features mentioned above can be considered as the measurements/observed states in the Hidden Markov Model (HMM).

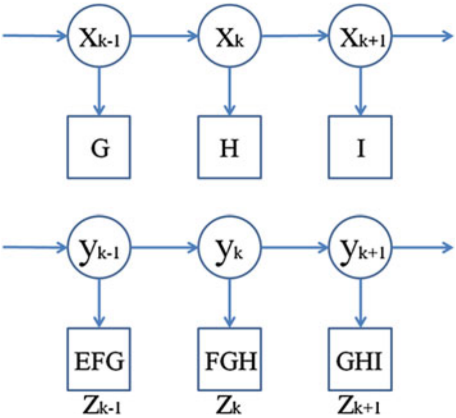

Fig. 3 An example of remodel the high-order Markov chain with the "memory" $m = 3$. The last 3 observations of each state in X is integrated into new observations Z for the new chain Y

3.1 The High-Order Markov Chain

Markov chains are commonly used in the modeling of state changes over time sequence. The first order Markov chain is the basis of most Bayes filter variants, e.g., Kalman filter, and can be easily considered as an appropriate model for trajectory data. In this work we use an extended high-order Markov chain to integrating the long-term features mentioned above. Let x be the unobserved state (here the probability of anomaly) and z the measurement from long-term observations, the HMM as the process model is presented in Fig. 3.

The proposed high-order Markov chain $X = \{x_0, ..., x_n\}$ still follows the Markov assumption, i.e., the probability of the current state given a limited number of previous ones is conditionally independent of the other earlier states:

$$p(x_k|x_{k-1}, x_{k-2}, ..., x_{k-m}, ..., x_0) = p(x_k|x_{k-1}, x_{k-2}, ..., x_{k-m}) \tag{1}$$

with $m < k$. The measurement $Z = \{z_0, ..., z_n\}$ at each state is dependent not only on the corresponding state, but also several previous states:

$$p(z_k|x_k, x_{k-1}, ..., x_{k-m}, ..., x_0) = p(z_k|x_k, x_{k-1}, ..., x_{k-m+1}) \tag{2}$$

The higher order implies also, however, (1) the number of to be solved parameters grows exponentially with the order $O(|x|^m)$ (with $|x|$ the number of possible states of x and m the order) and (2) the reliability of parameter estimation decreases. We then remodel the trajectory by constructing a new chain $Y = \{y_0, ..., y_n\}$ with a m-tuple of x states:

$$y_k = (x_k, x_{k-1}, ..., x_{k-m+1}) \tag{3}$$

so that the new chain Y over the m-tuple is equivalent to a first order Markov chain keeping the conventional Markov property

$$p(y_k|y_{k-1}, y_{k-2}, ..., y_0) = p(y_k|y_{k-1}) \tag{4}$$

with a "memory" of m, and

$$p(z_k|y_k, y_{k-1}, ..., y_0) = p(z_k|y_k) \tag{5}$$

as shown in Fig. 3 (bottom).

3.2 The Belief Filter

The proposed filter is a simple variant of recursive Bayesian estimator keeping the dynamic property and the prediction/updating scheme. Although normally the prediction and updating steps work alternately and provide required inputs for each other, either of them has also the probability to be skipped. In this work both of these cases will happen:

- We are using multiple measurements (behavior features) for the updating. Sometimes more than one feature can be extracted at the same time stamp. With the assumption that the features are independent to each other (simplified assumption) the updating step is performed multiple times before the next prediction.
- These long-term features, however, cannot be continuously observed. In the interval of the given observations the prediction will be performed solely for multiple times.

Prediction
The prediction step calculates the total probability, i.e., the integral of the products of the transition probability $p(x_k|x_{k-1})$ and the probability of the previous state $p(x_{k-1}|z_{k-1})$ over all possible x_{k-1}. In this case we have only one variable, i.e., the belief, to be estimated and in principle it cannot be predicted based on any current measurements. We assume that the anomalous behaviors are rather "transitory" than the normal drive and, therefore, use a simple exponential decay to predict the belief of the next state:

$$x_k = x_{k|x_{k-1}} = F \cdot x_{k-1} + w_k \tag{6}$$

where

$$F = e^{-s \cdot k};\ w_k \sim \mathcal{N}(0, \sigma^2). \tag{7}$$

F simulates the decay given to the belief of anomaly along with the driving. A Gaussian noise is added by w_k. k is used to count the number of previous state(s) without new anomalous feature(s) being reported. The accumulation of k makes sure that the belief decays rapidly after the driver performs normally. The decay tendency can actually be tuned by the factor s. Generically we give no weight to k for each step, i.e., $s = 1$, in the urban area. Fine tuning of s can adapt the filter to:

- the driving in suburban areas with higher speed and sparse street crossings, in which case two potential anomalous behaviors may have longer time interval and the belief decay should be set slower to keep the pattern being found, and
- the pedestrian trajectory in dense urban areas, where the decay speed may need to be further enhanced to avoid continuous accumulation of the belief.

The predicted state estimate is taken as prior estimate for the current state.

Updating
The update step uses Bayes rules. The prior estimate is refined with the observation on the current state and deliver a posterior estimate.

$$p(x_k|z_k) = \frac{p(z_k|x_k) \cdot p(x_k|z_{k-1})}{p(z_k|z_{k-1})} \tag{8}$$

where the prior distribution is actually

$$p(x_k|z_{k-1}) = \prod_{i \in \mathcal{V}} p(z_{i,k}|x_k) \cdot p(x_k|x_{k-1}) \tag{9}$$

with multiple measurements. $p(x_k|x_{k-1})$ is the initial distribution after prediction. z_i with $i \in \mathcal{V}$ the observation(s) that have been previously integrated ($\mathcal{V}$) in the current updating phase.

3.3 *Belief Inference*

We employ two simple typical cases: detour and wrong turn with simulated data to present the inference process using the proposed belief filter. Besides the information of turns these two cases have their particular features that another one does not have i.e., the detour case has only detour factor and no repeated route while the wrong turn case has the latter only. So that the influence of the individual features can be well demonstrated. Figure 4 presents a simple simulated trajectory with detour (left) and the extracted high-level features plotted over time (right). The bold red line shows the inferred belief of anomalous patterns over time. The possibility values are also presented in the trajectory with scaled colors. Please note that the green circle and red asterisk are used to mark the start and end positions of the trajectory, respectively. The value/color of each line segment is determined by its start point. We use these two examples to demonstrate some typical situations in the inference process.

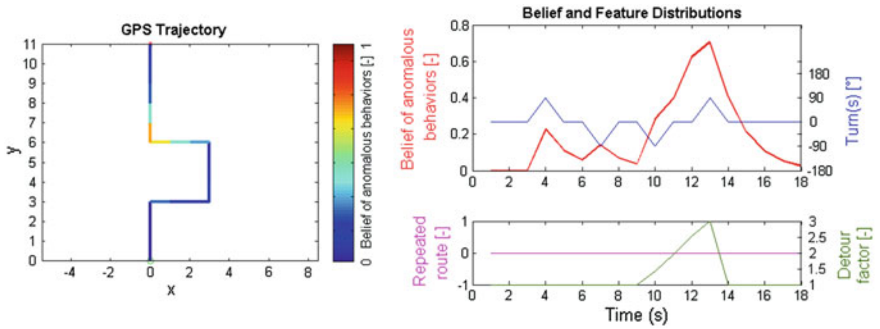

Fig. 4 Simulated trajectory of detour (*left*) with start position (*green circle*), end position (*red asterisk*) and the belief of anomalous behavior shown with scaled color. Three high-level features: Turns (*blue*), repeated route (*magenta*) and detour factor (*green*) are plotted together with the belief of anomaly over time (*right*)

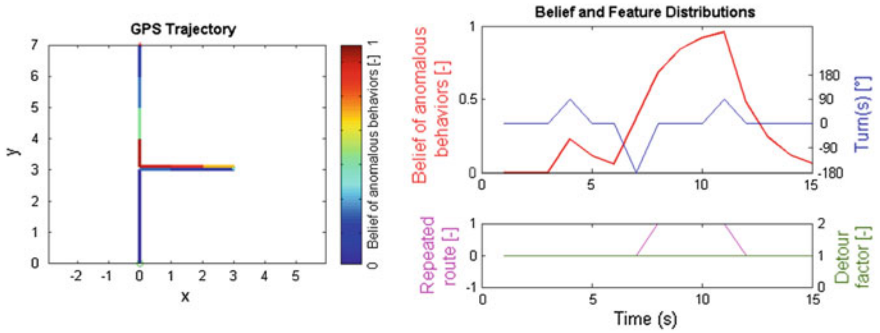

Fig. 5 Simulated trajectory of detour: trajectory with colors indicating the belief of anomaly (*left*) and the distributions of the belief and behavior features over time (*right*)

- Double different turns is considered normal. If the current turn has the different direction to the previous one, less probability is given to guarantee the continuous decay of the belief of anomaly.
- Double same turns, in the contrast, mean potential detour or even looping. Probability gain is added when the second turn happens.
- Detour factor increases and reach the maximum value when the detour is finished. The belief of anomaly has the peak value at this time as well.
- After the detour the belief of anomaly has a fast decay.

Another typical case of wrong turn is shown in Fig. 5. All the states of repeated route have the same feature value of 1. The belief increases as long as the vehicle stays in the wrong way and reaches the peak value at the spot where the wrong turn started. The belief decays to normal value when the vehicle goes back to the previous road.

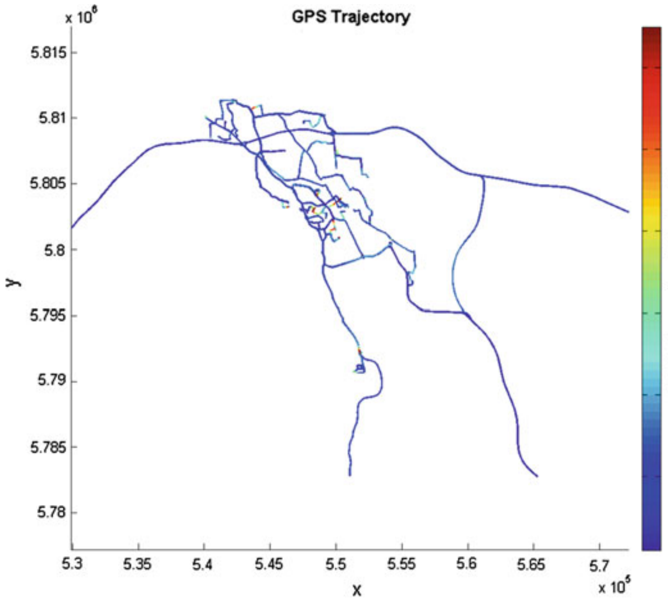

Fig. 6 Experiment on a VGI dataset with 100 Trajectories. Colors of the trajectory indicate the belief of anomaly

4 Experiments

Experiments are performed on the volunteered geographic information (VGI) data, an open trajectory dataset as well as the trajectories from self-acquisition. Figure 6 shows an experiment on a VGI dataset, which has been gathered by one business commuter in two years inside the city of Hannover, Germany. 100 trajectories are randomly selected to test the presented algorithm. Anomalous behaviors, with a probability of anomaly over 85 %, are detected in 12 trajectories (12 % of the total number). 368 out of total 68136 (0.54 %) GPS nodes are labeled presenting potential anomaly. Detailed analyses on several individual trajectories can be found in the follow-up figures.

In comparison with the two simulated cases (cf. Figs. 4 and 5) Figs. 7 and 8 show the detour and wrong turn detections on the actual VGI trajectories.

A more complicated case with turns, loop and an incomplete detour is given in Fig. 9 (left). As shown in the belief and feature distributions (Fig. 9, right), the influence of the first two turns decrease rapidly along the driving and present actually a normal segment. Forming a loop is in contrast a prominent anomalous behavior, which consists of three sequential turns and route repetition when the vehicle goes back to the former road. A small segment containing turns combination is detected afterwards and part of it is recognized as a detour (second peak of the green line) with the proposed method.

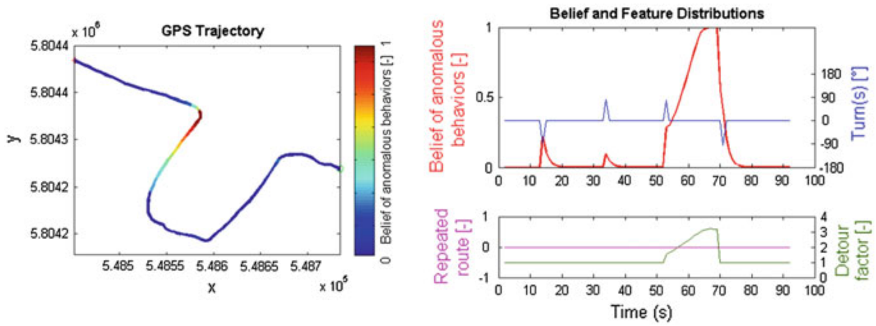

Fig. 7 An example of detour: trajectory with colors indicating the belief of anomaly (*left*) and the distributions of the belief and behavior features over time (*right*)

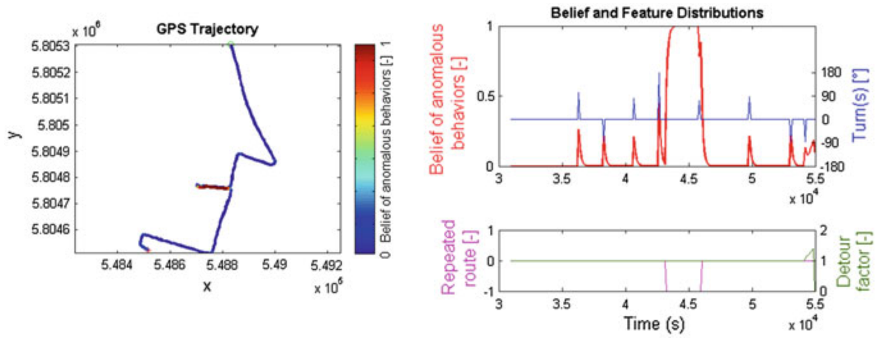

Fig. 8 An example of wrong turn: trajectory with colors indicating the belief of anomaly (*left*) and the distributions of the belief and behavior features over time (*right*)

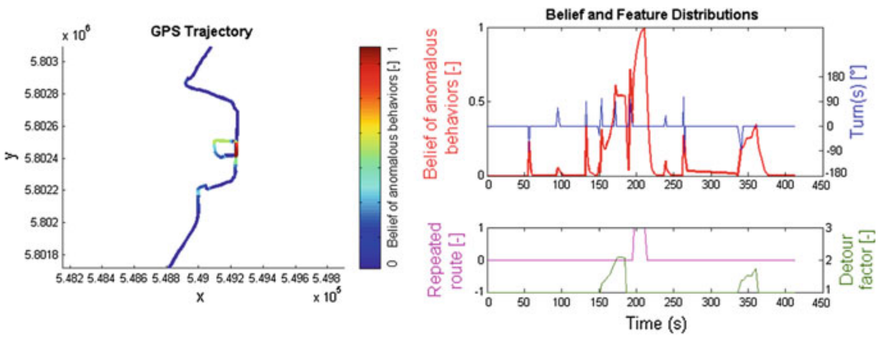

Fig. 9 A single trajectory with turns, loop and detour: trajectory with colors indicating the belief of anomaly (*left*) and the distributions of the belief and behavior features over time (*right*)

Figure 10 demonstrates a trajectory which is correctly recognized as normal driving, i.e., no obvious anomalous pattern are found. The belief distribution function

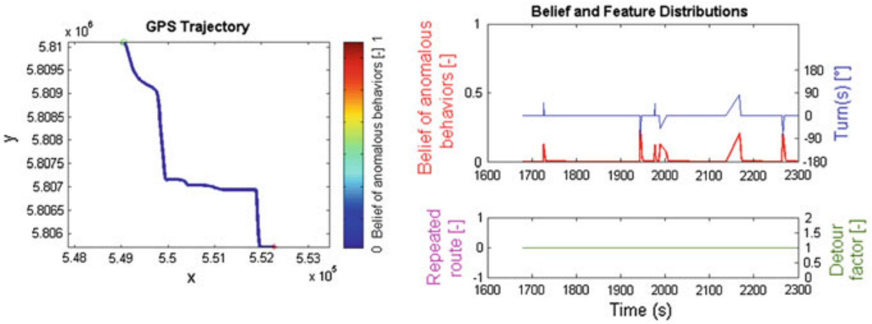

Fig. 10 An example of normal driving: trajectory with colors indicating the belief of anomaly (*left*) and the distributions of the belief and behavior features over time (*right*)

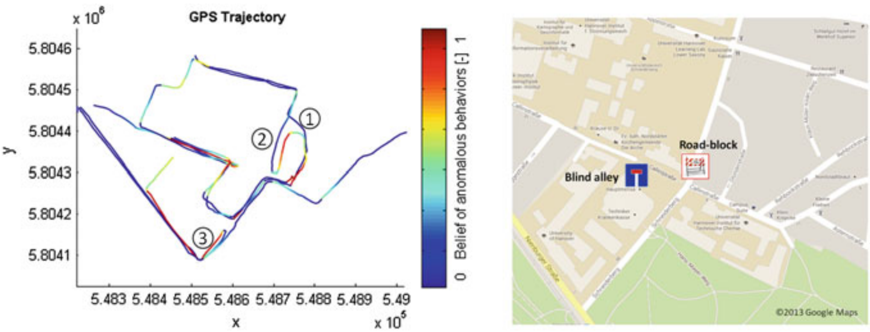

Fig. 11 Collective behaviors in the cases of road-blocking and blind alley: a collection of trajectories in the same area and similar time period (*left*) and the street map with the locations of the road-block and the blind alley being manually labeled (*right*)

shows robustness with only slight fluctuates, even though the trajectory also contains a few large turns.

Surely the anomalous behaviors are not frequent in the usual trajectory data. In some cases of urban traffic, e.g., road-blocking, blind alley or turn restriction, however, anomalous traces will be often found and concentrated in a certain area. These collective behaviors reflect to a certain extent the traffic issues mentioned above.

Figure 11 shows trajectories from self-acquisition (with known traffic conditions and driver behaviors) in an urban area, where a temporary road-block as well as a blind alley nearby (right) near a road crossing can be found. Anomalous patterns are found at the end of the blind alley and from multiple sides of the road-block. As shown in the trajectories (left), driver 1 from the north saw the sign of road-block and made a detour, driver 2 missed the warning sign of road-block, had to make a U-turn right before the road-block and then performed a detour to go on with the same direction. Driver 3 from the south turned around even earlier because of the warning sign and observing a traffic jam before the crossing. Although the blind

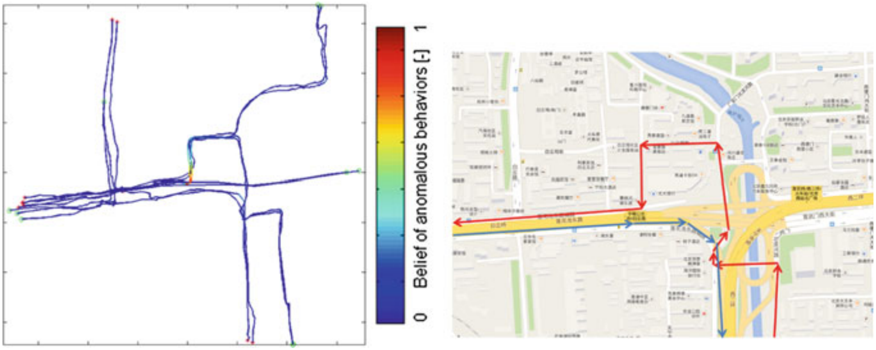

Fig. 12 Collective behaviors in the case of turn restriction: trajectories (*left*) passing the road intersection shows no anomaly except that from the bottom to the left and the street map (*right*) with the routes from bottom to left (*red*) and reversed (*blue*) indicating the turn restriction

alley on the west side is not a temporary setup, it causes U-turns sometimes for the drivers who are not familiar with this area.

Figure 12 presents an experiment on the open trajectory dataset: "GeoLife GPS Trajectories" (Zheng et al. 2008, 2009, 2010) of Beijing, China. Inside the shown segment trajectories from the bottom to the left show coincident detour while the reversed (from the left to the bottom) trajectories have no anomaly. We assume that such phenomena may indicate a potential left turn restriction, which is proven by the street map shown in Fig. 12 (right), i.e., no left turn is possible here because of the cloverleaf junction and the direction restrictions in the streets.

5 Conclusion

This chapter presents an original approach to anomalous behavior detection in the trajectory data by means of a recursive Bayesian filter. The main contributions of this work can be summarized as follows:

- A recursive belief filter is conducted for the dynamic detection of anomalous patterns.
- Long-term behavior features are integrated using high-order Markov model.
- Unsupervised detection in single trajectory with local features.

By these means the belief of anomalous behaviors can be inferred dynamically over time. In single trajectory, the result indicates where the driver is likely meeting navigation problem and an assistant is needed. Furthermore, a potential of reflecting traffic issues, e.g., turn restrictions, unexpected blind alleys and temporary road-blocks, is shown as well by analyzing the collective behaviors of multiple trajectories. We are, however, actually aware that the geometric features only are still limited for a plausible anomaly detection. Further semantic information and background knowledge might be helpful to estimate the anomalous behaviors more precisely.

In this chapter we demonstrate one application of the proposed filter—the detection of specific driving behaviors in GPS trajectory data. We assume that this scheme can also be (1) extended to other trajectory patterns given corresponding features and (2) adapted to the trajectories of pedestrian or animals, which are derived from the other sensors like cameras and trackers.

References

Bu Y, Chen L, Fu AWC, Liu D (2009) Efficient anomaly monitoring over moving object trajectory streams. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '09. ACM, New York, NY, USA, pp 159–168. doi:10.1145/1557019.1557043

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv 41(3):15:1–15:58. doi:10.1145/1541880.1541882

Hu W, Xiao X, Fu Z, Xie D, Tan T, Maybank S (2006) A system for learning statistical motion patterns. IEEE Trans Pattern Anal Mac Intell 28(9):1450–1464. doi:10.1109/TPAMI.2006.176

Kalman RE (1960) A new approach to linear filtering and prediction problems. Trans ASME-J Basic Eng 82(Series D):35–45

Kim K, Lee D, Essa I (2011) Gaussian process regression flow for analysis of motion trajectories. In: Proceedings of IEEE international conference on computer vision (ICCV). IEEE computer society

Ma TS (2009) Real-time anomaly detection for traveling individuals. In: Proceedings of the 11th international ACM SIGACCESS conference on computers and accessibility, Assets '09. ACM, New York, NY, USA, pp 273–274. doi:10.1145/1639642.1639712

Masreliez C, Martin R (1977) Robust bayesian estimation for the linear model and robustifying the kalman filter. IEEE Trans Autom Control 22(3):361–371. doi:10.1109/TAC.1977.1101538

Morris B, Trivedi M (2008) A survey of vision-based trajectory learning and analysis for surveillance. IEEE Trans Circuits Syst Video Technol 18(8):1114–1127. doi:10.1109/TCSVT.2008.927109

Pang LX, Chawla S, Liu W, Zheng Y (2011) On mining anomalous patterns in road traffic streams. In: Proceedings of the 7th international conference on advanced data mining and applications–volume Part II, ADMA'11. Springer, Berlin, Heidelberg, pp 237–251

Pang LX, Chawla S, Liu W, Zheng Y (2013) On detection of emerging anomalous traffic patterns using gps data. Data Knowl Eng 87:357–373. http://www.sciencedirect.com/science/article/pii/S0169023X13000475

Piciarelli C, Micheloni C, Foresti G (2008) Trajectory-based anomalous event detection. IEEE Trans Circuits Syst Video Technol 18(11):1544–1554. doi:10.1109/TCSVT.2008.2005599

Prevost C, Desbiens A, Gagnon E (2007) Extended kalman filter for state estimation and trajectory prediction of a moving object detected by an unmanned aerial vehicle. In: American control conference, ACC '07, pp 1805–1810. doi:10.1109/ACC.2007.4282823

Sun L, Li D, Yi D, Liu J (2012) Trajectory tracking based on iterated unscented kalman filter of boost phase. In: 2012 IEEE International conference on service operations and logistics, and informatics (SOLI). pp 232–235. doi:10.1109/SOLI.2012.6273537

Zheng Y, Li Q, Chen Y, Xie X, Ma WY (2008) Understanding mobility based on gps data. In: ACM conference on ubiquitous computing (UbiComp 2008). Korea, ACM Press, Seoul, pp 312–321

Zheng Y, Zhang L, Xie X, Ma WY (2009) Mining interesting locations and travel sequences from gps trajectories. In: International conference on World Wild Web (WWW 2009). ACM Press, Madrid Spain, pp 791–800

Zheng Y, Xie X, Ma WY (2010) Geolife: a collaborative social networking service among user, location and trajectory. IEEE Data Eng Bull 33(2):32–40