Detecting Anomaly in Traffic Flow from Road Similarity Analysis

Xinran Liu[1], Xingwu Liu[2], Yuanhong Wang[1],
Juhua Pu[1,3(✉)], and Xiangliang Zhang[4]

[1] The State Key Laboratory of Software Development Environment,
Beihang University, Beijing, China
{liuxinran,lucienwang,pujh}@buaa.edu.cn
[2] Institute of Computing Technology Chinese Academy of Sciences, Beijing, China
liuxingwu@ict.ac.cn
[3] Research Institute of Beihang University in Shenzhen, Shenzhen, China
[4] King Abdullah University of Science and Technology, Jeddah, Saudi Arabia
xiangliang.zhang@kaust.edu.sa

Abstract. Taxies equipped with GPS devices are considered as 24-hour moving sensors widely distributed in urban road networks. Plenty of accurate and realtime trajectories of taxi are recorded by GPS devices and are commonly studied for understanding traffic dynamics. This paper focuses on anomaly detection in traffic volume, especially the *non-recurrent traffic anomaly* caused by unexpected or transient incidents, such as traffic accidents, celebrations and disasters. It is important to detect such sharp changes of traffic status for sensing abnormal events and planning their impact on the smooth volume of traffic. Unlike existing anomaly detection approaches that mainly monitor the derivation of current traffic status from history in the past, the proposed method in this paper evaluates the abnormal score of traffic on one road by comparing its current traffic volume with not only its historical data but also its neighbors. We define the neighbors as the roads that are close in sense of both geo-location and traffic patterns, which are extracted by matrix factorization. The evaluation results on trajectories data of 12,286 taxies over four weeks in Beijing show that our approach outperforms other baseline methods with higher precision and recall.

1 Introduction

With more and more taxies equipped with GPS device, plenty of accurate, realtime vehicle trajectory data are available. Taxi trajectory data are considered to be rich and reliable for studying traffic dynamics in a city, because a large amount of taxies are 24-hour moving sensors widely distributed in urban road networks and reflecting the traffic status.

In this paper, we analyze taxi GPS trajectories for detecting the *non-recurrent traffic anomaly* caused by unexpected or transient incidents, such as traffic accidents, control pretests, celebrations, disasters, emergent road maintenance. This type of traffic anomaly results in sharp changes of traffic status

(speed or volume) in the local area of urban. Unlike the traffic jams caused by the peak hours or holidays, traffic anomaly is difficult to anticipate. It is thus important to detect the anomalies for sensing abnormal events and planing for their impact on the smooth volume of traffic [1].

Generally, there are two different ways for detecting traffic anomaly. The first type of approaches follow an intuitive way of general anomaly detection in any field: model the normal behavior and report abnormal behaviors if they differ much from the normal ones. To detect traffic anomalies, a traffic volume forecasting system is built for modeling normal traffic patterns. Anomalies are reported when the system is not able to predict correctly, i.e., the observed traffic deviates much from what the system predicts as normal traffic [2]. The detection accuracy highly depends on the forecasting system used for normal traffic prediction. A poor prediction algorithm can result in false and missed alarms. In the second type, abnormal traffic patterns are identified based on comparing current observations with those in adjacent days [3] or in history up to this time [1]. An anomaly is reported if a large derivation or difference is found. For example, traffic volume patterns at the link connecting one region to another are compared with those at the same time in adjacent days [3]. The link is suspected to be abnormal if the traffic volume patterns behave differently from other days. To detect topological variation in traffic flow between two regions, [1] takes the distribution pattern of traffic volumes over all paths from one region to another at current time and compares it to the median of all distribution patterns up to this time. More discussion of approaches in this type will be presented in Sect. 5.

Our proposed approach in this paper is different from the existing ones on the following aspects:

(1) We detect anomalies in road level, rather than region or link level. It thus improves the spatial accuracy of the detected anomalies from region level to road segment level.
(2) We evaluate the abnormal score of traffic on one road by comparing its current traffic volume with both its historical data and its neighbors. The neighboring roads are selected according to their closeness in geo-location and similarity in traffic patterns. This new evaluation of anomaly reduces false alarms caused by only comparing to the historical data. Traffic often has fluctuations. The day-to-day or day-of-week variability due to temporal factors are easy to caught by monitoring periodically regular changes. However, irregular "surges" or "decreases" caused by circumstantial factors or special events introduce high variation to historical data. For example, the traffic volume sharply increased on the first workday of the Victory Day on September 3rd 2015 resulted from the no limits for driving on roads, while the traffic volume decreased obviously on December 8th 2015 due to the hazy weather in Beijing. By additionally considering the traffic status in neighboring roads, our approach is therefore more robust to irregular variation than the method in [4,5].
(3) We compare one road segment with those that are close in sense of both geo-location and traffic patterns, rather than comparing with only spatially

close roads or all road segments [6] in two reasons. First, comparing with all segments is not computationally efficient, as the number of road segments is large. Second, comparing with only spatially close roads potentially hides anomalies that affect several spatial adjacent roads in a region [6], e.g., a major traffic accident. We factorize the traffic volume matrix by Nonnegative Matrix Factorization (NMF) and extract a pattern matrix and a coefficient matrix, which represent the types of traffic patterns in all road segments and the distribution of these types in each road segment, respectively. Similarity of two roads is measured by what types of traffic running on the roads and how geographically close they are. Bringing in roads similar in traffic patterns for comparison helps on filtering out anomalies that causes traffic variation jointly in spatial adjacent roads.

We evaluate our proposed approach on trajectories data of 12,286 taxies over 4 weeks in Beijing, and compared with baseline methods in [4–6]. The experimental results show that our approach outperforms baseline methods with higher precision and recall, and improves the F1-measure by at least 10 %.

The rest of the paper is organized as follows. Section 2 introduces the traffic pattern extraction. Section 3 presents the proposed traffic anomaly detection method. Section 4 reports the experimental evaluation results. Section 5 discusses the related work. Finally Sect. 6 concludes our paper.

2   Traffic Pattern Extraction

2.1   Basic Traffic Pattern Extraction Based on NMF

Definition 1 (Traffic Volume Matrix): The traffic volume of each road on day d is represented by traffic volume matrix V^d, where each element v_{ij}^d is the traffic volume of i-th road on day d at j-th hour. For example, $V^d \in \mathbb{R}_+^{m \times 24}$ is a traffic volume matrix containing 24 h traffic volumes of m roads.

$$V^d = \begin{bmatrix} 151 & 108 & \cdots & 73 \\ 45 & 29 & \cdots & 50 \\ \ldots & \ldots & \ldots & \ldots \\ 22 & 32 & \cdots & 20 \end{bmatrix}$$

To avoid the disturbance from the noise in the historical volume, traffic volume matrix is not used directly in our study. Inspired by [7] where Nonnegative Matrix Factorization (NMF) was employed to find human mobility patterns and land uses of urban area, we use NMF to extract basic traffic patterns of roads, because it produces two nonnegative matrices with physical meanings that match our analysis requirement [8]. Meanwhile, NMF reduces data dimensionality and thus makes the anomaly detection more efficient.

NMF decomposes a nonnegative matrix into two:

$$V \approx CP \tag{1}$$

where $C \in \mathbb{R}_+^{m \times r}$, $P \in \mathbb{R}_+^{r \times h}$, m is the number of road segments, h is the number of time slots, 24 h of one day in our case, and r is the number of traffic patterns (usually specified according to application need).

To minimize the loss of factorization, an objective function is defined as

$$J(C, P) = \frac{1}{2} \sum_{i,j} [V_{ij} - (CP)_{i,j}]^2 \tag{2}$$

Given r, matrix C and P can be iteratively updated until convergence by

$$C_{ik} \leftarrow C_{ik} \cdot \frac{(VP^T)_{ik}}{(CPP^T)_{ik}}, \quad P_{kj} \leftarrow P_{kj} \cdot \frac{(C^T V)_{kj}}{(C^T CP)_{kj}} \tag{3}$$

The matrix C and P in our application have their own physical meanings.

Definition 2 (Pattern Matrix): P, pattern matrix, has rows each of which corresponds to a basic traffic pattern, representing a stable mode of human collective mobility.

Definition 3 (Coefficient Matrix): C, coefficient matrix, includes the coefficients of each road segment with respect to the corresponding basic patterns. In other words, one row of C indicates the proportion and scale of each basic traffic patterns on a road, and thus can be employed to study road similarity in next section.

The factorization formula $V \approx CP$ indicates that the traffic volume of one road (e.g., i-th row of V) is a linear combination of traffic patterns (P), with coefficient in the i-th row of C. Next, we will demonstrate and discuss the P pattern matrix and C coefficient matrix obtained from our traffic volume data.

2.2 Discussion of Pattern Matrix

Our data set includes traffic of 28 days. We first construct the Traffic Volume Matrix V^d for each day, $d = 1, ..., 28$. The factorization rank r in NMF is a critical parameter. In order to obtain stable and meaningful decomposition results, we tried r to be any integer from 2 to 20. For each setting of r, V^d ($d = 1, ..., 28$) is factorized into C^d and P^d.

Conforming with the finding in [7], when $r = 3$, the obtained matrix P is most stable over time, namely, the row-wise deviation of $\{P^d\}_{d=1}^{28}$ is minimum.

Figure 1(a) shows the basic patterns on workdays, P_1, P_2 and P_3 denoting the three row of P. Solid lines represent the mean value, while dashed lines represent the positive and negative deviations averaged on different days. The basic pattern P_1 is in accord with working or education mobility pattern. The peak hour of P_1 is at 7:00–8:00 am and 5:00–6:00 pm, when people go to work and get off work, or students go to school and leave school. The basic pattern P_2 reveals people moving in daytime, about from 10:00 am to 4:00 pm, which may be the traveling between two workplaces for business. The peak hour of P_3 is at around 8:00 pm, which is related with entertainment and other activities.

2.3   Discussion of Coefficient Matrix

Each row of C contains r coefficients, describing how much different basic patterns contribute to the traffic volume on each road. In our case of $r = 3$ for example, a large value in the first coefficient means the traffic volume of this road is relatively high in 7:00–8:00 am and 5:00–6:00 pm. Figure 1(b) visualizes the coefficients of road segments in C (the three columns corresponding to P_1, P_2 and P_3 are represented in red, blue and yellow, respectively). The major areas in red are Xizhimen and Zhongguancun, which are the vital communication lines in Beijing and have high traffic volume during peak hours. The blue areas include two famous business zones: Sanlitun and Xidan. Areas in yellow are widely distributed over the whole city, as well as roads heading to Beijing Capital Airport. Therefore, we can see roads with similar coefficients have the same property, even though they are not adjacent. Considering the similarity in traffic patterns can compensate for the deficiency of finding similar roads only by closeness in geo-location.
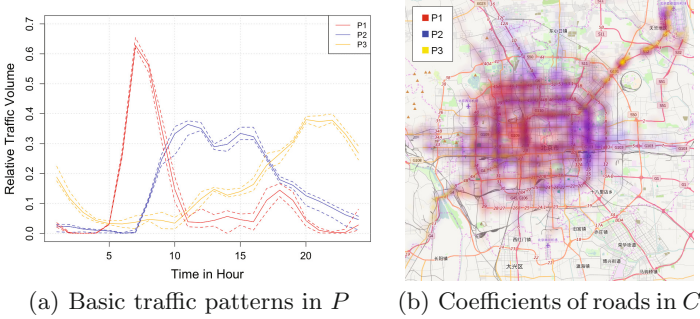

(a) Basic traffic patterns in P (b) Coefficients of roads in C

Fig. 1. The results of NMF (Color figure online)

3   Traffic Anomaly Detection

After applying NMF, we develop our proposed traffic anomaly detection approach. Generally, the approach consists of two steps: first, select similar road segments for comparison; second, calculate the anomaly score.

3.1   Finding Similar Roads by Clusters

As mentioned in the Sect. 2, the traffic volume of a road can be considered as a linear combination of different traffic patterns. The combination coefficients indicate how the traffic on this road relevant to different patterns. In other words, one row of C matrix characterizes each road segment as a new representation vector. Therefore, two road segments i and j hosting similar types of traffic pattern will have a small distance calculated from the i-th and j-th row of C.

Note, road segments are compared based on C, rather than the original volume matrix V, due to 1) efficiency (C has a smaller size than V); and 2) robustness (C is road's features mapped into stable space P). The neighboring roads are selected according to their closeness in geo-location and similarity in traffic patterns. We thus define the distance between road i and road j as follow:

$$D(i,j) = \alpha D_t^T(i,j) + (1-\alpha)\frac{D_g(i,j)}{\tau} \tag{4}$$

Distance $D(i,j)$ is composed of two parts: D_t^T is the distance measured based on traffic patterns; D_g is the physical distance between two roads. τ is a scaling constant for making D_t^T and D_g comparable. α is a parameter adjusting the weight of two distances, whose setting will be studied in Sect. 4.4. We set $T = 21$ (the number of training days) and $\tau = 10000$ (scaling down the geo-distance measured in meters) in this paper.

While D_g can be computed by the coordinates of two roads, D_t^T is defined as:

$$D_t^T(i,j) = \sqrt{\sum_{d=1}^{T}\sum_{k=1}^{3}(C_{ik}^d - C_{jk}^d)^2} \tag{5}$$

where C_{ik}^d is the k-th coefficient of i-th road on day d in coefficient matrix C^d. $D_t^T(i,j)$ actually measures the difference of coefficients of two roads on T continuous days, for enhancing the stability.

The neighboring roads are formed based on clustering. Given the pairwise distance from Eq. (4), roads can be clustered by exemplar-based method, which takes actual roads as cluster centroids. Affinity Propagation (AP) [9] is the most appropriate algorithm for our problem. It a clustering algorithm proposed to find out the most representative actual items, called *exemplars*, in a data set. It provides optimality guarantee about minimizing the clustering distortion, compared to k-medoids. It can be applied with or without a specified number of clusters [10]. Since our problem has no determined number of clusters, we assign each road the same preference to be an exemplar and let AP method form the optimal number of clusters. To ensure a road has enough neighbors, a cluster with a small number of members will be merged with another cluster to which it has the nearest center.

3.2   Anomaly Scoring

We evaluate the abnormal score of traffic on one road by comparing its current traffic volume with not only its historical data but also its neighbors. Normally, traffic volume at current moment is similar to that at the same moment of previous days. However, irregular factors, such as the atrocious weather and the particular event, cause variations in traffic volume. We define the abnormal score by estimating the existing probability of traffic volume. The abnormal score of road i on day d at time slot t is:

$$S_{it}^d = \beta P_a^T(v_{it}^d) + (1-\beta)P_r(v_{it}^d) \tag{6}$$

where P_a^T is the existing probability among history data (till T days), while P_r is the probability among the neighbors' volume. We use Kernel Density Estimation (KDE) [11] algorithm to evaluate the probability. β is the weight for balancing the two factors, whose sensitivity will be discussed in Sect. 4.4. The P_a^T and P_r are defined as follow.

$$P_a^T(v_{it}^d) = \frac{1}{Th_1} \sum_{j=1}^{T} \phi(\frac{v_{it}^d - v_{it}^j}{h_1}) \tag{7}$$

$$P_r(v_{it}^d) = \frac{1}{Nh_2} \sum_{j=1}^{N} \phi(\frac{v_{it}^d - v_{jt}^d}{h_2}) \tag{8}$$

where v_{it}^d is the volume of road i on day d at time slot t. N is the number of neighbors in the cluster of road i. The KDE kernel ϕ is set to be Gaussian. Bandwidth $h_1 = \frac{5\sigma}{T}$ and $h_2 = \frac{5\sigma}{N}$, and σ is the standard deviation of the history volume or neighboring volume.

4   Experiments and Analysis

4.1   Settings

Trajectory Data: Real world taxi GPS trajectories are used to evaluate the effectiveness and efficiency of the proposed algorithm, with statistics shown in Table 1. As about 20 % of traffic on road surfaces in Beijing is generated by taxicabs, the taxi trajectories represent a significant portion of traffic volume on the network [1]. We used ST-Matching algorithm to process the low-sampling-rate GPS trajectories [12], and counted the traffic volume for each road segment every hour. The first 3 weeks trajectory data are used as training set, while the last 1 week data are used as testing set.

Road Network: The used road networks of Beijing are shown in Table 1.

Table 1. Statistics of dataset

Trajectories	data duration	2–29 Nov. 2011
	No. of taxies	12,286
	No. of effective days	28
	No. of Avg. sampling intervals(s)	60
Roads	No. of road segments	180,350
	No. of road nodes	132,273
Weibo Reports	data duration	23–29 Nov. 2011
	Avg. reports of day	19

Traffic Anomaly Reports: WeiBo is a Twitter-like social site in China. We extract the traffic anomaly reports from the news posted at WeiBo and use them for evaluating the detected anomalies. The statistics is shown in Table 1.

Baseline: We choose three different methods for comparison. ICDE2009 [6] detects road anomaly based on a comparison with all other roads. ICDM2012 [4] takes all roads and history data into consideration. SPIE2015 [5] applies KDE to investigate the probability value of a road traffic status given history data.

Measurement: The reported traffic accidents at WeiBo are taken as the ground truth. Three metrics are used:

1. **Recall** is the percentage of the number of the correctly detected anomalies over the number of all reported anomalies.
2. **Precision** is the percentage of the number of correctly detected anomalies over the number of all detected anomalies.
3. $F1$-**measure** is the equally-weighted mean of precision and recall:

$$F1 - measure = \frac{2 * recall * precision}{recall + precision} \qquad (9)$$

4.2 Experiments Using Real-World Data

The detection performance of our proposed method is compared with that of baseline methods on the last seven days data. As shown in Fig. 2(a), the recall of the proposed method is 92.76 %, which is about 20 % higher than the highest recall of baseline. The precision of the proposed method is 74.55 % while the highest precision of baseline is 62.11 %. The $F1$-measure of proposed method is the highest.

(a) Real-world data

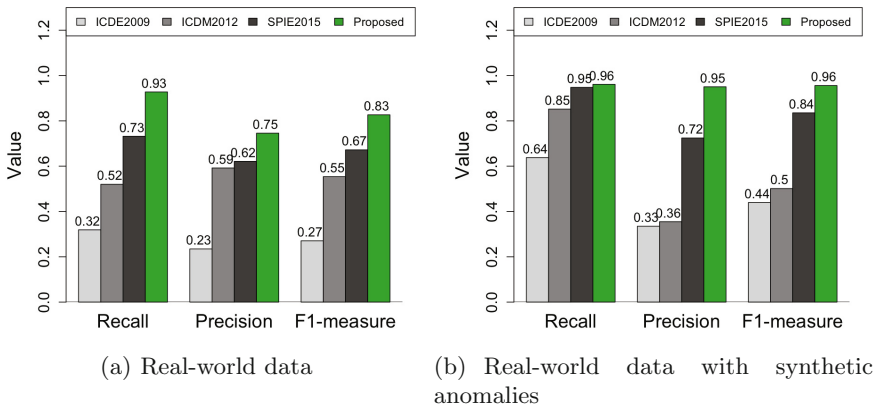(b) Real-world data with synthetic anomalies

Fig. 2. Detection performance comparing to baseline methods (Color figure online)

According to [6], the data in California is considerably stable with an average deviation of less than 10 %. While the data of Beijing is relatively unstable and irregular. Thus, our method performs the best detecting anomalies in traffic with high variations in Beijing, though ICDE2009 performed well with the stable traffic data in California. The performance of ICDM2012 is consistent with [1]. SPIE2015 is a special case of our method when $\beta = 1$.

The reported *Precision* in Fig. 2(a) is not as high as *Recall*, mainly because the ground truth we used from Weibo is just a subset of the true ground truth. The facts are: not all accidents were reported at Weibo; not all anomalies are caused by accidents; a reported anomaly event may have impact on several road segments and last for a long time but only have been recorded at one road and at one time stamp.

We have two case-studies on this unsatisfactory precision issue due to imperfect ground truth. In the first case, an anomaly was detected by our method at Workers' Stadium North Road around 10:00 pm on 11/26/2011. However, it was not reported in the Weibo record. After looking back the historical news, we find this anomaly was caused by a concert. In the evening of 11/26/2011, a vocal concert of Sodagreen (a popular idol groups in Taiwan, China) was hold in Beijing Workers' Sports Complex, which was nearby the road. In the Fig. 3(a), the red mark is the Beijing Workers' Sports Complex, and the blue mark is the Workers' Stadium North Road. The road is in the north of the stadium. Figure 3(c) shows the traffic volume patterns of the road and its neighbors. The traffic patterns of roads are accordant most of the time except the 10:00 pm of 11/26/2011, when the volume of the anomalous road is relatively larger than that of its neighbors. The increasing traffic volume may be caused by the leaving of the people when the concert was over.

In the second case, there was an anomaly on Xueyuan Road at 8:04 am on 11/23/2011 reported by Weibo, while two road segments are identified as anomalies by the proposed method. As illustrated in Fig. 3(b), the red mark is the reported anomaly and the two blue marks are the detected anomalies. Actually, both of the two road segments are the parts of Xueyuan Road. Due to the lack of the perfect ground truth, the proposed method cannot be evaluated
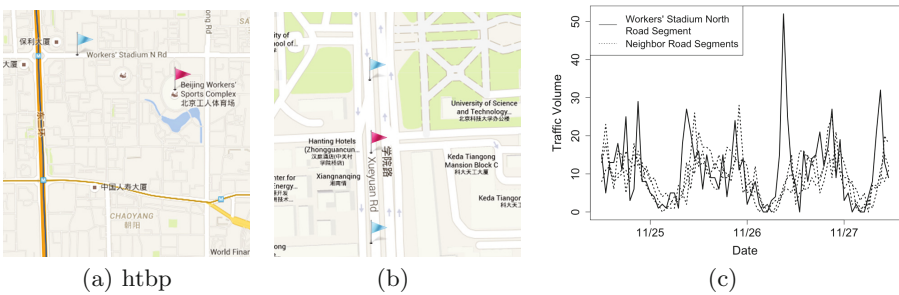

(a) htbp   (b)   (c)

Fig. 3. Two study cases (Color figure online)

Table 2. Results with different parameter α and β tested on real-world data

β \ α	0			0.5			1		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
0	65.38%	43.07%	51.93%	78.86%	63.84%	70.56%	81.22%	48.87%	61.03%
0.5	73.43%	53.10%	61.63%	**92.76%**	**74.56%**	**82.67%**	89.86%	69.40%	78.32%
1	73.20%	62.12%	67.21%	73.20%	62.12%	67.21%	73.20%	62.12%	67.21%

Table 3. Results with different parameter α and β tested on real-world data with synthetic anomalies

β \ α	0			0.5			1		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
0	65.53%	41.20%	50.60%	72.53%	82.29%	77.11%	80.83%	40.86%	54.30%
0.5	81.70%	68.74%	74.66%	**96.12%**	**95.04%**	**95.58%**	93.70%	81.70%	87.29%
1	94.77%	72.44%	82.11%	94.77%	72.44%	82.11%	94.77%	72.44%	82.11%

properly. Therefore, we design the semi-synthetic experiments, which will be discussed in the next section.

4.3   Experiments Using Semi-synthetic Data

To better evaluate our method, the experiment of semi-synthetic data still uses the last week data as testing set but with added synthetic anomalies.

Since the traffic volume on different days at the same time slot demonstrates normal distribution, anomalies are generated with volumes larger than $\mu + 4\sigma$, where μ is the average and σ is the standard deviation of traffic volumes. The results of the all methods are shown in Fig. 2(b). Our proposed method performs the best, while the precision is also improved.

4.4   Parameter Discussion

In this section, we take experiments on both real-world and semi-synthetic data to explore how parameters α and β have an influence on anomaly detection, as well as the clustering results.

The α and β. α is the weight to balance the distance of geo-location and of traffic patterns in Eq. (4). β is the weight to balance the history score and neighbor score in Eq. (6). Tables 2 and 3 show the results of different α and β tested on real-world and real-world data with synthetic anomalies, respectively.

From the results, we can see $\alpha = \beta = 0.5$ has the best performance. That is to say, considering both closeness in geo-location and traffic patterns finds better neighboring roads for comparison, and combining the comparison with history traffic data (time domain) and neighboring road traffic status (space domain) results in better abnormal score.

To further study the impact of these two parameters, we fix $\beta = 0.5$ and change α. The F1-measure with different setting of α is shown in Fig. 4(a).

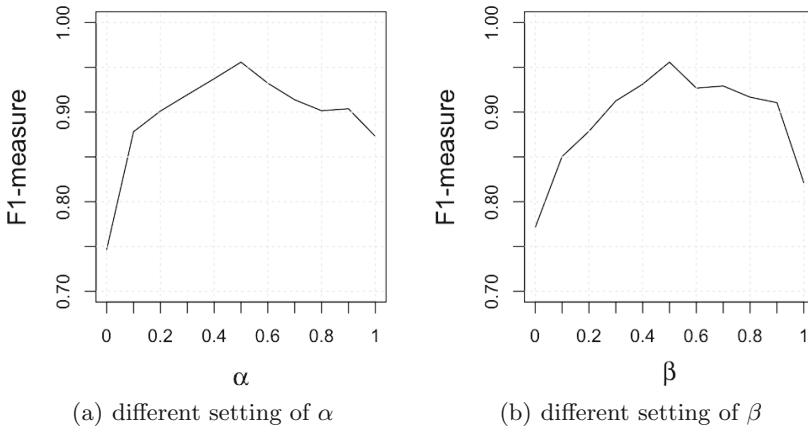**Table 4.** Sensitivity analysis of parameter p in AP clustering

Setting of p	0.5*m	m	2*m
No. of clusters	638	376	220
Recall	90.42 %	**96.12 %**	92.84 %
Precision	91.83 %	**95.04 %**	89.99 %
F1-measure	91.12 %	**95.58 %**	91.39 %

Figure 4(b) presents the results when changing β with $\alpha = 0.5$. As shown in the figures, when $\beta = 0.5$ and $\alpha = 0.5$, the performance is the best.

When $\alpha = 0$, the distance measure in Eq. (4) takes only the road closeness in geo-location, while when $\alpha = 1$ the distance relies only on the closeness in traffic patterns. Neither of these two settings can select a proper neighboring road cluster as a reference for anomaly detection. Bringing roads with similar traffic patterns together with spatial adjacent roads, anomalies affecting regions from a single road or a large district can be detected.

The β parameter controls how much the anomaly score relying on neighboring data (solely, when $\beta = 0$) and historical data (solely, when $\beta = 1$). Relying on only historical data causes false alarms or miss anomalies when traffic volume varies with irregular "surges" or "decrease" caused by circumstantial factors or special events. Neighboring (close in space and in road characteristics) traffic status helps the detection system resist such variations.

The Parameter of Clustering. We adjust the preference p in AP clustering method, and obtain different clustering results. Table 4 shows the results of anomaly detection with different parameter p. When $p = m$, we can obtain the best result, where m is the median of all pairwise distance values and is popularly

(a) different setting of α (b) different setting of β

Fig. 4. Sensitivity analysis of parameter α and β

suggested as the best setting of p in many applications [9]. When p changes, the number of obtained clusters varies a lot. However, the detection performance only slightly affected. Generally, the proposed method is not sensitive to the clustering results.

5   Related Work

The study of anomaly detection on trajectory data can be classified into two categories: **trajectory** anomaly detection and **traffic** anomaly detection. [13,14] focus on the problem of detecting anomalous trajectories by comparing with the historically "normal" routes. [1] identifies traffic anomalies by investigating drivers' routing behaviors that significantly differ from their original patterns. [15] detects traffic anomalies using microscopic traffic variables, such as relative speed, inter vehicle time gap, and lane changing. [3] proposes the STOTree algorithm based on both spatial and temporal properties of detected outliers. [16] adapts likelihood ratio tests to describe traffic patterns. Since these two methods detect anomalies on the level of regions, they may only find the very large scale events and miss the road segment-level traffic anomalies, such as traffic accident. Kuang et al. propose an anomaly detection method based on wavelet transform and PCA [17]. Even though [17] are efficient, its recall value is just 74.62 %. The work in this paper focuses on traffic anomaly detection on road segment level. The proposed detection method identifies anomalies by taking into account historical data in time and neighboring traffic status in space and in road characteristics.

6   Conclusion

In this paper we have studied the problem of detecting anomalies. We evaluate the abnormal score of traffic on one road by comparing its current traffic volume with not only its historical data but also its neighbors. The neighboring roads are selected according to their closeness in geo-location and high similarity in traffic patterns. We evaluated our method with GPS trajectories generated by 12,286 taxies in Beijing. The recall, precision and $F1$-measure of our method with real word data are 92.76 %, 74.56 % and 82.26 %, which is better than the baselines. In the next, we will focus on how the anomalies propagate.

References

1. Pan, B., Zheng, Y., Wilkie, D., Shahabi, C.: Crowd sensing of traffic anomalies based on human mobility and social media. In: ACM SIGSPATIAL GIS (2013)

2. Guo, J., Huang, W., Williams, B.M.: Real time traffic flow outlier detection using short-term traffic conditional variance prediction. Transp. Res. Part C: Emerg. Technol. **50**, 160–172 (2015)
3. Liu, W., Zheng, Y., Chawla, S., Yuan, J., Xing, X.: Discovering spatio-temporal causal interactions in traffic data streams. In: SIGKDD, pp. 1010–1018 (2011)
4. Chawla, S., Zheng, Y., Hu, J.: Inferring the root cause in road traffic anomalies. In: ICDM, pp. 141–150 (2012)
5. Ngan, H.Y., Yung, N.H., Yeh, A.G.: A comparative study of outlier detection for large-scale traffic data by one-class SVM and kernel density estimation. In: IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, pp. 94050I–94050I-10 (2015)
6. Li, X., Li, Z., Han, J., Lee, J.-G.: Temporal outlier detection in vehicle traffic data. In: ICDE, pp. 1319–1322. IEEE (2009)
7. Peng, C., Jin, X., Wong, K.-C., Shi, M., Liò, P.: Collective human mobility pattern from taxi trips in urban area. PloS One **7**(4), e34487 (2012)
8. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. Nature **401**(6755), 788–791 (1999)
9. Frey, B.J., Dueck, D.: Clustering by passing messages between data points. Science **315**(5814), 972–976 (2007)
10. Zhang, X., Wang, W., Nørvåg, K., Sebag, M.: K-AP: generating specified K clusters by efficient affinity propagation. In: ICDM, pp. 1187–1192 (2010)
11. Parzen, E.: On estimation of a probability density function and mode. Ann. Math. Stat. **33**, 1065–1076 (1962)
12. Lou, Y., Zhang, C., Zheng, Y., Xie, X., Wang, W., Huang, Y.: Map-matching for low-sampling-rate GPS trajectories. In: ACM SIGSPATIAL (2009)
13. Chen, C., Zhang, D., Samuel Castro, P., Li, N., Sun, L., Li, S.: Real-time detection of anomalous taxi trajectories from GPS traces. In: Puiatti, A., Gu, T. (eds.) MobiQuitous 2011. LNICST, vol. 104, pp. 63–74. Springer, Heidelberg (2012)
14. Chen, C., Zhang, D., Castro, P.S., Li, N., Sun, L., Li, S., Wang, Z.: iboat: Isolation-based online anomalous trajectory detection. IEEE Trans. Intell. Transp. Syst. **14**(2), 806–818 (2013)
15. Barria, J.A., Thajchayapong, S.: Detection and classification of traffic anomalies using microscopic traffic variables. IEEE Trans. Intell. Transp. Syst. **12**(3), 695–704 (2011)
16. Pang, L.X., Chawla, S., Liu, W., Zheng, Y.: On mining anomalous patterns in road traffic streams. In: Advanced Data Mining and Applications, pp. 237–251 (2011)
17. Kuang, W., An, S., Jiang, H.: Detecting traffic anomalies in urban areas using taxi GPS data. Math. Probl. Eng. **501**, 809582 (2015)