An online motion pattern recognition algorithm for moving objects considering the characteristics of stay points

Nanping Mao, Yuan Ling, Yan Su, Qiang Wang, Xiaodong Sun, Xudong Zhang, Lanhui Zeng, Guoqing Liu, Haoming Cui

Marine Department of Satellite Tracing and Metering Jiangyin, China

Keywords: transportation modes detection; stay point; trajectory classification

Abstract: The popularity of intelligent devices have promoted the growth of spatiotemporal trajectory data. How to update the rules of residence point cluster, path rule and classifier used in motion pattern recognition with the increase of data volume is an important research content. In this paper, an online motion pattern recognition algorithm for moving objects considering the characteristics of stay points and a parallel online motion pattern recognition algorithm for moving objects considering the characteristics of stay points are proposed. Experiments were conducted using Geolife data provided by Microsoft research and the results were analyzed.

1. Introduction

In recent years, with the popularity of GPS global positioning system, it is convenient for people to obtain the GPS geographical location information of moving objects, resulting in a large number of spatiotemporal trajectory data of moving objects. By analyzing the spatiotemporal trajectory data of moving objects, it can help researchers predict traffic jams to assist the construction of intelligent traffic and recommend routes. Therefore, spatiotemporal trajectory data mining has become a current research hotspot, and its research directions mainly include adjoint pattern mining, frequent pattern mining, anomaly detection and motion pattern recognition of moving objects [1-4].

The basic process of moving object motion pattern recognition algorithm based on spatiotemporal trajectory is generally divided into trajectory preprocessing, trajectory feature extraction and classifier training/recognition, among which the most critical step is trajectory feature extraction. According to the different methods of trajectory feature extraction, the research on the moving object motion pattern recognition algorithm is divided into moving object motion pattern recognition algorithm based on motion feature, moving object motion pattern recognition algorithm based on classification rules and moving object motion pattern recognition algorithm based on image signal analysis [5-8]. At first, researchers collected users' wifi data and analyzed the wifi data to identify the users' movement patterns. Krumm et al. [9] used the hidden markov model to divide the trajectory of moving objects into two categories: stay and movement. Yin et al. [10] proposed the DBN classification model. With the popularity of accurate GPS positioning, researchers began to use GPS data to analyze the motion pattern recognition of moving objects, In 2017, Liang J et al. [11] divided the motion characteristics of trajectory into three categories, namely velocity, acceleration and behavior. The Velocity class includes maximum Velocity, average Velocity, High Velocity Rate (HVR), Middle Velocity Rate (MVR) and Low Velocity Rate (LVR), which respectively represent the proportion of sampling points with High Velocity, sampling points with medium Velocity and sampling points with Low Velocity. Acceleration class includes maximum acceleration, average acceleration, etc. The behavior class includes the proportion of the number of stay points and the proportion of the points with large direction change.

How to update the rules of residence point cluster, path rule and classifier used in motion pattern recognition with the increase of data volume is an important research content [12-15]. Motion pattern online recognition refers to the real-time and rapid updating of classification models in motion pattern recognition for the constantly coming GPS data streams.

  DOI: 10.25236/cseem.2020.176

2. Relevant concepts and algorithm

Definition 1: Given the trajectory data set D, D is divided into three parts D_{Tr}, D_{Te} and $D_{Tr}{}'$. The training track set DTr contains the track set for the category label of the known motion mode, while the test track set DTe contains the track set for the category label of the unknown motion mode. The newly added training track set DTr 'refers to the track set of newly arrived known category labels.

Definition 2: $N_{Eps}(p)$ refers to the set of all points whose distance from point p is less than or equal to Eps, and the core object refers to the points with more than *MinPts* in the radius Eps. $\mathrm{UpdSeeds}_{\mathrm{Ins}}(p)=\{q|\ q$ is the core object in $D\cup\{p\}$, there is o. o is the core object in $D\cup\{p\}$, and $q\in N_{Eps}(o)\}$.

Definition 3: *changeCL* is expressed as $\{(CL_1\text{->}CL_1'),...(CL_n\text{->}CL_n')\}$, refers to the CL_i of the dwell point cluster is merged into a new cluster CL_i' after updating. *changeRL* is expressed as $\{(RL_1\text{->}RL_1'),...(RL_n\text{->}RL_n')\}$, refers to the rule RL_i updated and merged into a new rule RL_i'.

The method is divided into three parts: offline training stage, online training stage and test stage. The off-line training stage is basically the same as the process in the previous chapter. The difference is that the algorithm saves the information of stay point cluster, trajectory network diagram, stay point cluster rules, path rules, classifiers and so on mined in the off-line process, so as to facilitate the update of the online stage. The online training stage is divided into three stages: motion feature extraction, stay point feature extraction and classifier update.

Motion feature extraction is the same as offline phase. In the feature extraction phase of stay points, first extract the stay points for all the newly arrived training tracks, then read the current latest cluster of stay points, update these clusters with the newly arrived cluster and save them. Based on these updated clusters, the current network graph of the latest trajectory is read, updated and saved.

Finally, based on the updated trajectory network graph, the current rules are read and updated (including the rules of the dwell point cluster and the rules of the path).

These rules are used to extract the characteristics of stay points from the newly arrived training tracks and combine the motion features to form the track features. Finally, in the classifier update phase, the updated rule information is used to update the track features of all previous training tracks, and the classifier is updated based on the track features of the newly added training tracks. In the test phase, the latest rules of residence point cluster and path rules were used to extract track features for each test track, and the latest classifier was used to predict the motion pattern type.

3. Algorithm description

This section lists the most important part of the online training process compared to the off-line algorithms in the previous chapter. The updating of the dwell point cluster, the updating of the trajectory network graph, the updating of the dwell point cluster rules and the updating of the path rules and the updating of the classifier are briefly described.

Since the number of dwell point cluster rules and path rules in this online algorithm keeps changing with the addition of training trajectories, this will lead to the change of trajectory eigenvector dimensions, and the previous dimensions may be merged into one dimension, and multiple new dimensions will be added at the same time. Bagged the decision tree classification model selection model, as a result of multiple dimension may be consolidated into one dimension, can lead to multiple branch of the decision tree is merged, and the division of each node attributes (according to the attribute of the node split) changes, so using the incremental decision tree algorithm's time complexity is $O(N*M*H)$, where N is the train track number, M is the dimension of trajectory characteristics, H is the depth of the tree. It has the same complexity as the reconstruction decision tree. Due to the incremental decision tree to be stored by each node of the track samples, and direct decision tree to store it only once all training trajectory, space complexity is less than the incremental algorithm, so considering the above situation, directly to the trajectory characteristics of the original training by updating the mapping relation reconstruction after the old and new rules, according to the trajectory characteristics of the new training path to reconstruct the classification model.

The algorithm is divided into two parts: first, *changRL* is used to modify the track feature of the original training track, and then the Bagged decision tree is reconstructed by combining the track feature of the new training track.

Input: the new feature vector *Fe'* of the training track, the feature vector *Fe* of the original training track, the updated mapping relation *changeRL* of the old and new rules, and the height threshold *H* of the decision tree

Output: *Bagged* decision tree *baggedDT*

1. Foreach $fe \in Fe$ do {
2. If cantain(*fe*,*changRL*) then
3. Update(*fe*,*changeRL*)// update eigenvalue
4. $Fe= Fe \cup Fe'$
5. baggedDT=baggedDTBuild(*Fe*,*H*,*gini*) // construct a new classification model Bagged decision tree
6. Return baggedDT

4. Conclusion

This paper lists the most important part of the online training process compared to the off-line algorithms in the previous chapter. The updating of the dwell point cluster, the updating of the trajectory network graph, the updating of the dwell point cluster rules and the updating of the path rules and the updating of the classifier are briefly described. The algorithm is divided into three parts: offline training phase, online training phase and test phase. The off-line training stage is basically the same as the process in the previous chapter. The difference is that the algorithm saves the information of stay point cluster, trajectory network diagram, stay point cluster rules, path rules, classifiers and so on mined in the off-line process, so as to facilitate the update of the online stage.

References

[1] Jeung H, Shen H T, Zhou X. Convoy Queries in Spatio-temporal Databases[C]. IEEE 24th International Conference on Data Engineering,Cancun: Mexico, 2008: 1457-1459.

[2] Li Z H, Ding B L, Han J W, et al. Swarm: Mining Relaxed Temporal Moving Object Clusters[J]. Proceedings of the VLDB Endowment, 2010, 3(1-2): 723-734.

[3] Tang L A, Zheng Y, Yuan J, et al. On Discovery of Traveling Companions from Streaming Trajectories[C]. IEEE 28th International Conference on Data Engineering, Washington: USA, 2012: 186-197.

[4] Vieira M R, Bakalov P, Tsotras V J. On-line Discovery of Flock Patterns in Spatio-temporal Data[C]. Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Washington: USA, 2009: 286-295.

[5] Kalnis P, Mamoulis N, Bakiras S. On Discovering Moving Clusters in Spatio-temporal Data[C]. International Symposium on Spatial and Temporal Databases, Berlin:Germany, 2005: 364-381.

[6] Jeung H, Yiu M L, Zhou X, et al. Discovery of Convoys in Trajectory Databases[J]. Computer Science, 2010, 1(1):1068-1080.

[7] Gudmundsson J, van Kreveld M, Speckmann B. Efficient Detection of Motion Patterns in Spatio-temporal Data Sets[C]. Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems, Washington: USA,2004: 250-257.

[8] BenkertM, Gudmundsson J, Hübner F, et al. Reporting Flock Patterns[C]. European Symposium on Algorithms, Zurich: Switzerland, 2006: 660-671.

[9] Krumm J, Horvitz E. LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths[C]. International Conference on Mobile and Ubiquitous Systems: Networking and Services, Boston: USA,2004:4-13.

[10] Yin J, Chai X, Yang Q. High-Level Goal Recognition in A Wireless LAN. [C]. 16th Conference on Innovative Applications of Artificial Intelligence, California: USA. 2004: 578-584.

[11] Liang J, Zhu Q, Zhu M, et al. An Enhanced Transportation Mode Detection Method Based on GPS Data[C]. International Conference of Pioneering Computer Scientists, Engineers and Educators, Changsha: China, 2017: 605-620

[12] Patel D. Incorporating Duration and Region Association Information in Trajectory Classification[C]. IEEE 28th International Conference on Data Engineering, Arlington: USA, 2012:1132-1143.

[13] Patel D, Sheng C, Hsu W, et al. Incorporating Duration Information for Trajectory Classification[J]. Journal of Location Based Services, 2013, 7(4):246-271.

[14] Cheng H, Yan X, Han J, et al. Discriminative Frequent Pattern Analysis for Effective Classification[C]. IEEE International Conference on Data Engineering, Istanbul: Turkey,2007:716--725.

[15] Cheng H, Yan X, Han J, et al. Direct Discriminative Pattern Mining for Effective Classification[C]. IEEE 24th International Conference on Data Engineering, Cancun: Mexico,2008: 169-178.