Turkish Journal of Computer and Mathematics Education Vol.11 No.3 (2020), 1599-1607

Research Article

AN OVERVIEW OF TRAJECTORY DATA MINING TECHNIQUES TO ENHANCE THE MOBILITY DETECTION

[1]B Jyothi, [2]V Rama Krishna, [3]G Balram, [4]G Kiran Kumar

[1,2,3,4]Asst Prof, AnuragUniversity

[1]bjyothicse@cvsr.ac.in [2]ramakrishnacse@cvsr.ac.in [4] kirankumarcse@cvsr.ac.in

ABSTRACT

Trajectory data mining plays a most important role in the real world application which enables to predict the moving location details of humans, vehicles, animals and so on. Trajectory data analysis is a most difficult task due to frequent changing location details of humans due to its continuous mobility. Trajectorydata might consist of more noise details which is more complex to analysis. It is required to have more concentration on the trajectory data analysis to ensure the accurate trajectory data mining outcome. There are various research methods had been proposed by different researchers to perform the analysis of trajectory data. In this analysis work, discussion of various trajectory data analysis techniques has been given. This analysis work provides the detailed discussion of working procedure of each trajectory data analysis method with the samples. And also comparison evaluation of the research techniques in terms of merits and demerits is given in order to evaluate the performance. Finally numerical evaluation of the research techniques is provided based on different performance metrics. From this numerical evaluation, best method has been chosen as base work to enhance the trajectory data mining.

Keywords: Trajectory data,data mining, node mobility, path direction, human locations

I. INTRODUCTION

Movement detection is the most concentrated research issue focused by various researchers to enhance the working of different applications [1]. Nowadays, increased communication technologies requiredetecting the location and movement of objects to make accurate decision [2]. For example Google map direction needs to estimate the movement of objects more accurately to make the correct the prediction about the estimated arrival time and different path to reach the destination [3]. This increased the focus on various researchers towards analysing of trajectory data sets. However it is more difficult to handle and process the time varying location specific data accurately.

The main problem associated with the trajectory data mining is the increased growth of time specific data sets [4]. The increased technologies lead to more data generation for specific time period intervals which is more difficult to store and analyse [5]. The process of trajectory data mining requires more concentration to ensure the accurate and efficient extraction useful information to support the various applications such as urban control, traffic management, public health control, animal protection and so on [6]. The more efficient processing of data needs to be done to accomplish the well defined trajectory data analysis process.

There are various research techniques and tools have been introduced by researchers to perform the trajectory data mining efficiently. However few techniques and tools only recognized as the most successful tool to perform the trajectory data mining process [7].

Many of the methodologies tend to have more drawbacks and as well as some of them leads to guarantee the efficient outcome. It is very complex to predict the advantages and disadvantages of the varying research techniques. Thus it is required to integrate the working procedure and merits and demerits of varying research methodologies. This integrated overview might be helpful for the researchers to predict and extract the working procedure of different research techniques.

The main goal of this analysis work is to build the overall integrated view of the different research methodologies. Thus the researchers can obtain the integrated knowledge about the various research techniques easily without the need of studying varying research techniques. This analysis work provides the complete overview of the research techniques in terms of working procedure, merits and merits. And also the numerical evaluation of the different working techniques has also been carried out in this work to estimate the performance improvement of each technique.

The overall organization of the research work is given as follows: In this section introduction about the need of trajectory data mining techniques and the issues behind them is given. In section 2, discussion about the varying related research techniques whose contribution is to achieve the better performance in trajectory data mining is given. In section 3, comparison analysis of the methodologies discussed in section 2 in terms of merits and demerits has been given. In section 4 numerical assessments of those corresponding techniques has been given. Finally in section 5, conclusion of the analysis work is provided in terms of comparison and numerical assessment.

II. ANALYSIS OF TRAJECTORY DATA MINING METHODS

Trajectory data mining techniques are used to extract the useful information from the time varying geo location information. In this section, working procedure of the different trajectory data mining techniques has been given.

Liu et al [8] introduced the technique of utilizing RF tag array monitoring for the detection of moving objects. This method reduces the burden of fixing the RFID for single object by fixing the set of RFID sensors on the fields. Here the object movement is detected based on signal interference where the object which is crossing through the fixed sensor would affect the signal relay. These signal information would be collected for the periodic time which will then be analysed to predict the information. This gathered signal will be learned by using training and testing phase to predict the object movement information accurately.

Wang et al [9] proposed an interactive system for the traffic jam detection which is done visually. To perform this trajectory information are extracted from the network which are pre-processed first to remove the noise information. And the data cleaning is performed to remove the unwanted and irrelevant information from the data. This cleaned data is then mapped with the road map to detect the traffic jam information. This method ensures the accurate and efficient prediction of the traffic jam information. And also this method enables proposed system to explore the traffic jam information of the entire city.

Zheng et al [10]introduce the new system namely Gathering. This research work attempted to gather the various kind of trajectory datasets from online such as traffic monitoring, urban control, public health management and so on. This system attempted to

introduce the online pattern discovery system through which trajectory data set will be processed and the information can be extracted efficiently. Authors mainly attempted to contribute the online discovery based trajectory management system through which various trajectory dataset can be processes and the information can be mined. This is accomplished by introducing the different optimization techniques thus the newly arrived trajectory data can also be processed.

He et al [11] proposed an intelligent routing scheme to enhance the performance of the car pooling process. The main goal of this research work is to introduce the trajectory based intelligent routing scheme through which car location and movement can be detected more accurately. Thus the more efficient decision making about the car pooling process can be taken. This research method enables the route path merging by finding the most suitable alternate route paths to reach the destination and also can pick up the customers on its way. This method also concentrates and gains better performance in case of increased traffic rate.

Lee et al [12]introduced the Unifying Framework to support the pattern mining of different types of trajectory data's. This method reduces the gap between the different trajectory data sets which are varying with each other. This method accomplishes this by following out two steps namely initial pattern discovery and granularity adjustment. This research method can attain the improved performance and can support the varying kind of trajectory data sets efficiently. This research method also supports to enhance the performance by constructing the pattern forest structure through which information can be extracted.

d'Acierno et al [13] focused on massive generation of trajectory data due to increased technology usage such as digital camera, mobile phones and so on. Authors introduced the Novel Storage Management System for handling the increased arrival of data sets. This research work ensures the optimal storage of data contents in the database system. This work also ensured the optimal querying and extraction of information from the data sets. This method works over 3D queries for retrieving the data contents stored in the database. Experimentation of this work ensures that this research work can be applied for the extended real time applications and also can guarantee the extraction most useful information.

Qiao et al [14] introduced the three in one trajectory prediction model to evaluate the trajectory pattern of moving objects dynamically. This research methods attempts to reduce the research gap of eliminating the working linearly to predict the moving object locations. It is done by performing pattern mining in three steps. Those are constructing the dynamic index structure of moving objects which is names as constrained network R tree. And then region of interest part is extracted to learn about the different trajectory data point information. Finally future location prediction of moving objects is done to improvise the trajectory pattern mining result.

Zhou et al [15] proposed a Multi-Context Trajectory Embedding Model, called MC-TEM, to investigate settings efficiently. MC-TEM is created in the disseminated portrayal learning system, and it is adaptable to describe different sorts of valuable settings for various applications. To the best of the information, it is the first occasion when that the disseminated portrayal learning strategies apply to direction information. It formally joins different setting

data of direction information into the proposed model, including client level, direction level, area level and transient settings. All the setting data is spoken to in the equivalent implanting space. It applies MC-TEM to two testing undertakings, in particular area suggestion and social connection expectation.

Zhang et al [16] attempted to improve the bus traffic information by learning and predicting the geo location information of bus movement continuously. This research method ensures the accurate prediction outcome of the bus traffic information by learning and predicting the bus movement behaviour based on extracted trajectory data. This improved the accuracy of the pattern mining outcome by analysing the historical data also. Based on this information this method selects the most optimal routing path, thus the optimal vehicle movement can be guaranteed. The experimental outcome of this research work proves that it can ensure the increased packet delivery rate by selecting the most optimal route path.

Fu et al [17] proposed novel methodology for the disclosure of continuous course designs dependent on direction reflection. In the first place, direction parcel, area extraction, information disentanglement and basic fragment revelation are utilized to digest direction information, convert these directions into regular section fleeting groupings (STS) and produce 1-visit itemsets. At that point, an example mining calculation is proposed dependent on the spatial-worldly nearness relationship (STAR). This calculation utilizes the imperative system and bidirectional anticipated database to mine regular course designs from STS. In light of the genuine GeoLife direction information, the test results show that the proposed strategy has better execution and can discover longer course designs than other right now accessible techniques.

Altomare et al [18]displayed a work process based parallel methodology for finding examples and tenets from direction information, in a Cloud-based system. Test assessment has been completed on both genuine world and engineered direction information, up to one million of directions. The outcomes demonstrate that, because of the high multifaceted nature and expansive volumes of information associated with the application situation, the direction design mining process exploits from the versatile execution condition offered by a Cloud engineering regarding both execution time, accelerate and scale-up.

Liu et al [19]introduced the novel interactive visual system namely SmartAp which can select the good solutions accurately. This framework integrate the visual analyse with the data mining techniques for the accurate and efficient prediction of trajectory patterns. And also this research method proposed the group of couple visualization sets to predict the correlation between the different solutions. This research method ensures the accurate prediction of bill board locations with ranked details. This method rank the different solution sets based on reviews gathered from the domain experts.

Qi et al [20]accomplished the cost-effective and dependable directing in DTN-empowered VANETs, a novel convenience mindful direction information mining calculation is proposed to anticipate hubs' future positions. A meager time-space diagram is then gotten, in light of which, two directing heuristics are proposed. Recreation results show that our proposed steering calculations guarantee ease and high dependability after some time.

Muzammal et al [21]thought about indeterminate sensor information and change this to probabilistic direction information utilizing pre-handling schedules. Next, creators

demonstrate this information as tuple level dubious information and propose dynamic programming based calculations to mine fascinating directions. An exhaustive observational investigation is performed to assess the adequacy of the methodology. The outcomes demonstrate that the directions could be displayed and filled in as probabilistic information and that the outcomes could be figured effectively utilizing dynamic programming.

III. COMPARISON EVALAUTION OF TRAJECTORY DATA MINING TECHNIQUES

In this section, comparison evaluation of the trajectory data mining techniques has been given in terms of their merits and demerits.

Table 1. Comparison evaluation of the research techniques

S.No	Author	Method	Merits	Demerits
1	Liu et al [2012]	RF tag array monitoring system	It reduces the processing cost considerably by avoiding the RFID sensor integration for individual objects	This method is only suitable for the particular field whereas it cannot be applied for the high application with increased objects
2	Wang et al [2013]	Interactive system for the traffic jam detection	Accurate and faster detection of traffic jam information It is more suitable for the real time traffic monitoring system	Visual encoding are not performed accurately which might reduce the system accuracy
3	Zheng et al [2014]	Gathering	Supports dynamic updation of newly arriving the trajectory data Reduced time consumption for the prediction of the trajectory data set	Required more processing overhead to accomplish the trajectory data mining
4	He et al [2014]	Intelligent routing scheme	Better traffic management Shortest path selection	This scheme might affect the personalized requirements of the individual users
5	Lee et al [2015]	Unifying Framework	It enables to extract the different level of details from the varying kind of datasets Efficient mining of different trajectory data types Time constrained	It required complete information to ensure the accurate mining result
6	d'Acierno et al [2015]	Novel Storage Management System	Guaranteed useful information extraction from large volume of database Efficient handling of	Increased burden while storing the different kind of trajectory data in the single database

			large volume of incoming data More time consuming	
7	Qiao et al [2015]	Three in one trajectory prediction model	Accurate and efficient trajectory pattern mining outcome Efficient future transportation network by estimating the future location information	Personalized requirements of users is not concentrated
8	Zhou et al [2016]	Multi-Context Trajectory Embedding Model	It provides very promising results for the trajectory data pattern mining	Less accuracy due to not learning the trajectory data patterns which affects the incoming data
9	Zhang et al [2016]	Vela	Increased packet delivery ratio Stronger scalability	More computation overhead
10	Fu et al [2017]	spatial-temporal adjacency relationship	Guaranteed selection of the longer route paths Increased scalability Accurate prediction of trajectory pattern mining outcome	Reduced accurate by not considering the semantic information
11	Altomare et al [2017]	Workflow-based parallel approach	Reduced execution time Increased speed up and scalability Flexible environment	Reduced performance by not considering the future estimation information
12	Liu et al [2017]	SmartAp	This research method adapts the positive feedbacks which leads to accurate prediction outcome	More computation overhead
13	Qi et al [2017]	Cost-efficient and reliable routing in DTN-enabled VANET	Cost efficient link connectivity Increased reliability	More computational time
14	Muzammal et al [2018]	Trajectory data mining for uncertain data	Computationally efficient Increased scalability	More computation overhead

IV. NUMERICAL EVALUATION

In this section performance evaluation of the different research methodologies has been done in terms of accuracy of mined result. The comparison evaluation has been made between the different research techniques. The methodologies that are compared in this research work are Interactive system for the traffic jam detection (ISTJD), Unifying Framework (UF), SmartAp and Trajectory data mining for uncertain data (TDMUD).

The accuracy is defined as the correct prediction of the trajectory patterns from the different period of times. The accuracy value is calculated in terms of the prediction system's true positive, false positive, true negative and false negative values. The accuracy is calculated as like as follows:

$$\text{Accuracy} = \frac{T_p}{(T_p + F_p + F_n)}$$

The graphical representation is given as like as follows:

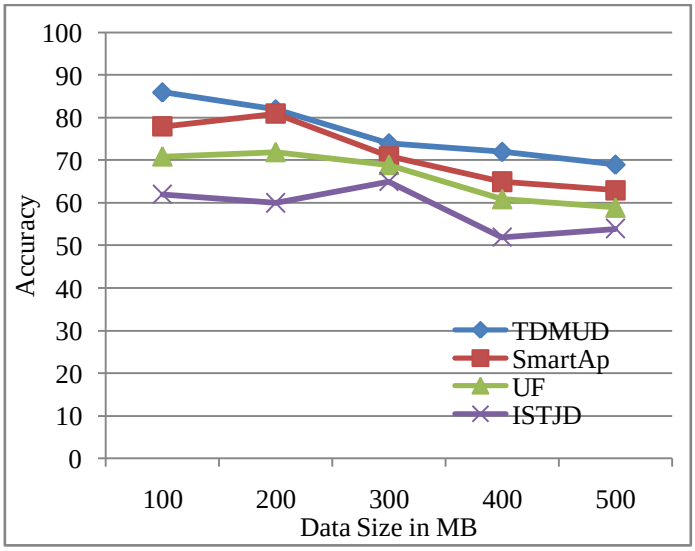

Figure 1. Accuracy Vs data size

From the figure 4, it is clear that the accuracy of the proposed methodology namely TDMUD is higher than the existing methodologies namely SmartAp, UF and ISTJD. The accuracy value increased in the proposed method linearly than the existing methods for the different set of data sizes.

V. CONCLUSION

In this analysis work comparison evaluation of the various trajectory data mining research methodologies is performed in terms of their efficiency and accuracy. This work provided the detailed summary of the research techniques along with their working procedure. This work also provided merits and demerits of different research techniques to predict the most suitable method. This is confirmed by evaluating the research techniques based on numerical from which it is proved that the most latest research method namely Trajectory Data Mining for Uncertain Data performs better than the other existing research methods.

REFERENCE

1. Yun, J., & Lee, S. S. (2014). Human movement detection and identification using pyroelectric infrared sensors. Sensors, 14(5), 8057-8081.

2. Wohl, M.A. and O'hagan, J.J., ZIH Corp, 2014. Method, apparatus, and computer program product for performance analytics determining location based on real-time data for proximity and movement of objects. U.S. Patent Application 13/942,585.

3. Ramsya, K. S., Reshma, S., & Venkatesan, D. S. (2015). Direction with Way Point Maker using Google Map API. *Internasional Journal on Applications in Information and Communication Engineering, India*.

4. Zheng, Y. (2015). Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *6*(3), 29.
5. Lv, M., Chen, L., Xu, Z., Li, Y., & Chen, G. (2016). The discovery of personally semantic places based on trajectory data mining. *Neurocomputing*, *173*, 1142-1153.
6. Jarvis, L. S., & Valdes-Donoso, P. (2018). A selective review of the economic analysis of animal health management. *Journal of Agricultural Economics*, *69*(1), 201-225.
7. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., &Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and structural biotechnology journal*, *15*, 104-116.
8. Liu, Y., Zhao, Y., Chen, L., Pei, J., & Han, J. (2012). Mining frequent trajectory patterns for activity monitoring using radio frequency tag arrays. *IEEE Transactions on Parallel and Distributed Systems*, *23*(11), 2138-2149.
9. Wang, Z., Lu, M., Yuan, X., Zhang, J., & Van De Wetering, H. (2013). Visual traffic jam analysis based on trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2159-2168.
10. Zheng, K., Zheng, Y., Yuan, N. J., Shang, S., & Zhou, X. (2014). Online discovery of gathering patterns over trajectories. *IEEE Transactions on Knowledge and Data Engineering*, *26*(8), 1974-1988.
11. He, W., Hwang, K., & Li, D. (2014). Intelligent carpool routing for urban ridesharing by mining GPS trajectories. *IEEE Transactions on Intelligent Transportation Systems*, *15*(5), 2286-2296.
12. Lee, J. G., Han, J., & Li, X. (2015). A unifying framework of mining trajectory patterns of various temporal tightness. *IEEE Transactions on Knowledge and Data Engineering*, *27*(6), 1478-1490.
13. d'Acierno, A., Saggese, A., & Vento, M. (2015). Designing huge repositories of moving vehicles trajectories for efficient extraction of semantic data. *IEEE Transactions on Intelligent Transportation Systems*, *16*(4), 2038-2049.
14. Qiao, S., Han, N., Zhu, W., & Gutierrez, L. A. (2015). TraPlan: An Effective Three-in-One Trajectory-Prediction Model in Transportation Networks. *IEEE Trans. Intelligent Transportation Systems*, *16*(3), 1188-1198.
15. Zhou, N., Zhao, W. X., Zhang, X., Wen, J. R., & Wang, S. (2016). A general multi-context embedding model for mining human trajectory data. *IEEE transactions on knowledge and data engineering*, *28*(8), 1945-1958.
16. Zhang, F., Jin, B., Wang, Z., Liu, H., Hu, J., & Zhang, L. (2016). On geocasting over urban bus-based networks by mining trajectories. *IEEE Transactions on Intelligent Transportation Systems*, *17*(6), 1734-1747.
17. Fu, Z., Tian, Z., Xu, Y., & Zhou, K. (2017). Mining frequent route patterns based on personal trajectory abstraction. *IEEE Access*, *5*, 11352-11363.
18. Altomare, A., Cesario, E., Comito, C., Marozzo, F., & Talia, D. (2017). Trajectory pattern mining for urban computing in the cloud. *IEEE Transactions on Parallel and Distributed Systems*, *28*(2), 586-599.

19. Liu, D., Weng, D., Li, Y., Bao, J., Zheng, Y., Qu, H., & Wu, Y. (2017). Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE transactions on visualization and computer graphics*, *23*(1), 1-10.

20. Qi, W., Song, Q., Wang, X., & Guo, L. (2017). Trajectory data mining-based routing in DTN-enabled vehicular Ad Hoc networks. *IEEE Access*, *5*, 24128-24138.

21. Muzammal, M., Gohar, M., Rahman, A. U., Qu, Q., Ahmad, A., & Jeon, G. (2018). Trajectory Mining Using Uncertain Sensor Data. *IEEE Access*, *6*, 4895-4903.