Effective Traffic Information Dissemination using Spatial Database

D.SureshKumar [1],S.Anbarasu [2]
[1]CSE, Sasurie College of Engineering India
[2]CSE, Sasurie College of Engineering India
[1] csesuresh58@gmail.com
[2] anbuharrish@gmail.com

Abstract**—** Finding efficient driving directions has become a daily activity and been implemented as a key feature in many map services like Google and Bing Maps. A fast driving route saves not only the time of a driver but also energy consumption (as most gas is wasted in traffic jams).GPS-equipped taxis can be regarded as mobile sensors probing traffic flows on road surfaces, and taxi drivers are usually experienced in finding the fastest (quickest) route to a destination based on their knowledge. In this paper, we mine smart driving directions from the historical GPS trajectories of a large number of taxis, and provide a user with the practically fastest route to a given destination at a given departure time. In our approach, we propose a time-dependent landmark graph, where a node (landmark) is a road segment frequently traversed by taxis, to model the intelligence of taxi drivers and the properties of dynamic road networks. Then, a Variance-Entropy-Based Clustering approach is devised to estimate the distribution of travel time between two landmarks in different time slots. Based on this graph, we design a two-stage routing algorithm to compute the practically fastest route. We build our system based on a real world trajectory simulation dataset and stored in MOD.

Keyword- **Spatial databases and GIS, data mining, GPS trajectory, driving directions, driving behavior.**

I. INTRODUCTION

WHILE location-based services (LBSs) are booming in this decade, many vendors start to provide map and navigation services (e.g., Garmin, GoogleMap, MapQuest, NavTeq, Yahoo! Map) along with convenient geo-tagging tools that enable the content providers (e.g., retail stores, facilities and general users) to publish location-dependent information on digital maps. Here, we refer to location-dependent information (e.g., point of interest, traffic, and local events) as spatial objects (or objects for short). Nowadays, there is a tremendous increase of Moving Objects Databases (MOD) due to, on the one hand, location-acquisition technologies like GPS and GSM networks and, on the other hand, computer vision-based tracking techniques. This explosion of information combines an increasing interest in the area of trajectory data mining and, more generally, knowledge discovery from movement aware data. All these technological achievements require new services, software methods, and tools for understanding, searching, retrieving, and browsing spatiotemporal trajectories content. To realize and identify the traffic structure using trajectories, we propose two novel index structures, namely, Route Overlay and Association Directory. The former manages the recently, in the literature there have been proposed several works that try to either efficiently analyze trajectory data or mine movement-aware patterns. In the domain of trajectory, segmentation-related works deal with the problem locally, by partitioning trajectories in a way as to achieve better database organization or extract more intuitive local patterns for clustering and classification purposes. In processing objects on a road network, two basic operations, namely, network traversal and object lookup, are involved. The former visits network nodes and edges according to network proximity, while the latter accesses and checks the attributes of objects located at traversed nodes or edges against object search criteria.

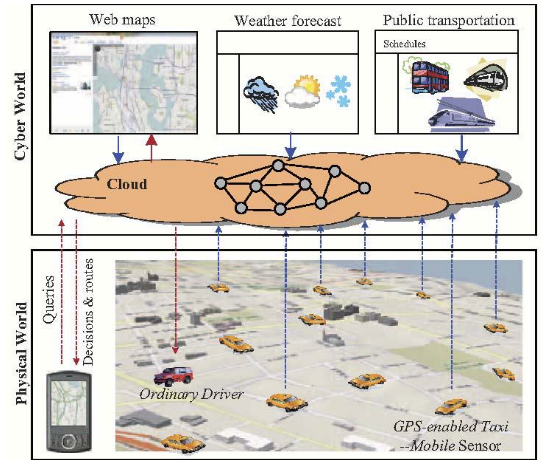

Fig 1.GPS Trajectory

Essentially, the time that a driver traverses a route depends on the following three aspects: 1) The physical feature of a route, such as distance, capacity (lanes), and the number of traffic lights as well as direction turns; 2) The time-dependent traffic flow on the route; 3) A user's driving behavior. Given the same route, cautious drivers will likely drive relatively slower than those preferring driving very fast and aggressively. Also, users' driving behaviors usually vary in their progressing driving experiences. E.g., traveling on an unfamiliar route, a user has to pay attention to the road signs, hence drive relatively slowly. Thus, a good routing service should consider these three aspects (routes, traffic and drivers), which are far beyond the scope of the shortest/fastest path computing. In this paper, we propose a cloud-based cyber physical system for computing practically fast routes for a particular user, using a large number of GPS equipped taxis and the user's GPS-enabled phone. As shown in Fig. 1, first, GPS-equipped taxis are used as mobile sensors probing the traffic rhythm of a city in the physical world. Second, a Cloud in the cyber world is built to aggregate and mine the information from these taxis as well as other sources from Internet, like Web maps and weather forecast. The mined knowledge includes the intelligence of taxi drivers in choosing driving directions and traffic patterns on road surfaces. Third, the knowledge in the Cloud is used in turn to serve Internet users and ordinary drivers in the physical world. Finally, a mobile client, typically running in a user's GPSphone, accepts a user's query, communicates with the Cloud, and presents the result to the user. The mobile client gradually learns a user's driving behavior from the user's driving routes (recorded in GPS logs), and supports the Cloud to customize a practically fastest route for the user.

I.A VE-CLUSTERING

The road network is dynamic, it can use neither the same nor a predefined time partition method for all the landmark edges. Meanwhile, as shown in Fig. 4(a), the travel times of transitions pertaining to a landmark edge clearly gather around some values (like a set of clusters) rather an a single value or a typical Gaussian distribution, as many people expected. This may be induced by

1) The different number of traffic lights encountered by different drivers. 2) The different routes chosen by different drivers traveling the landmark edge. 3) Drivers' personal behavior, skill and preferences. Therefore, different from existing methods regarding the travel time of an edge as a single valued function based on time of

day, consider a landmark edge's travel time as a set of distribution score responding to different time slots. Additionally, the distributions of different edges, such as $e13$ and $e16$,change differently over time.

II. PROBLEM DEFINITION:

The challenge after storing the data is the implementation of appropriate analytics for extracting useful knowledge. However, traditional data warehousing systems and techniques were not designed for analysing trajectory data.

II.A OBJECTIVE:

Modeling traffic flow in a road network. Our framework for mining traffic patterns, Detecting traffic flow in a road network using effective object search methods. To provide better driving experience, This system aims to provide traffic free path information, To mine smart driving directions from a large number of real-world historical GPS trajectories.

III. SYSTEM ANALYSIS

The problem of point-to-point shortest path computation in spatial networks is extensively studied with many approaches proposed to speed-up the computation. Most of the existing approaches make the simplifying assumption that weights (e.g., travel-time) of the network edges are constant.

Disadvantages Of Existing System: Existing TF-OPTICS mainly clusters whole trajectories and is not tailored to identify patterns of sub-trajectories in an unsupervised way.1)The temporal information is not considered in 2)The segmentation is performed per trajectory and it does not use global criteria,3)The identified clusters of segments conform to straight movement patterns and cannot identify complex (e.g., snake-like) patterns, which are usual in real-world applications.4)The proposed algorithm for identifying the representative trajectory is defined per cluster and it is a synthetic trajectory computed by an averaging technique.

This proposed system to find out the practically fastest route for a particular user at a given departure time. Specifically, the system mines the intelligence of experienced drivers from a large number of taxi trajectories and provide the end user with a smart route, which incorporates the physical feature of a route, the time-dependent traffic flow as well as the users' driving behaviors (of both the fleet drivers and of the end user for whom the route is being computed). We build a real system with real-world GPS trajectories generated by over 33,000 taxis in a period of three months, then evaluate the system with extensive experiments and in-the-field evaluations. The results show that our method significantly outperforms the competing methods in the aspects of effectiveness and efficiency in finding the practically fastest routes. Overall, more than 60 percent of our routes are faster than that of the existing online map services, and 50 percent of these routes are at least 20 percent faster than the latter.

Advantages Of Proposed System: 1)Effective implementation of Moving object data management 2)Acquiring and storing trajectories in MODs 3)Location-aware querying and object searching 4)Trajectory indexing 5) The automatic segmentation of the given trajectories into "homogenous" sub trajectories according to their "representativeness" in MOD 6)Sampling of the most representative subtrajectories of the MOD.

III.A FLOW OF PROCESS AND ARCHITECTURE:

1)Moving object data management.2)Acquiring and storing trajectories in MODs.3)Trajectory indexing.4)Store and query trajectory data.5)Reconstruct a trajectory from raw logs- Position devices provide us information just about location points and about trajectories.6)Analyse trajectory data.7)Spatial-temporal patterns in database.

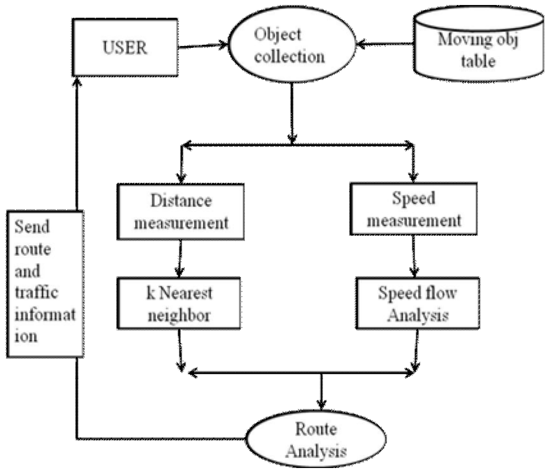

Fig 2.Flow of Process

The challenge after storing the data is the implementation of appropriate analytics for extracting useful knowledge. However, traditional data warehousing systems and techniques were not designed for analysing trajectory data.

IV. MODULES

IV.A QUERYPROCESSING:

First, the user sends a query tuple to the server, while accessing the user details the system will get the time and location details from their presence. Initiating the objects and its location is the first step of the implementation. The followings are the parameters involved in the node processing and its querying value identification process.

The route and object creation will take place the following attributes:1)Number of nodes, 2)Number of edges, 3)Speed, 4)Direction, 5)Signals

The process which included in the route creation and updating process

1)Update process, 2)Signaling, 3)Hash table

IV.B SEGMENTATION

This module introduces an algorithm for the automatic segmentation of trajectories into "homogenous" sub-trajectories according to their "representativeness" in the MOD.The module splits the area into different clusters for effective result comparison and analysis.

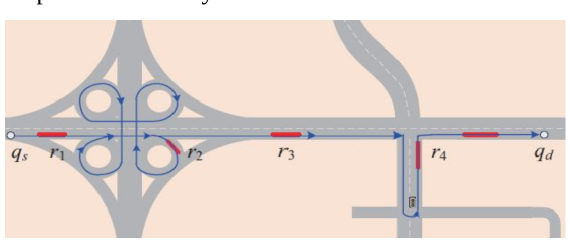

Fig 3 Segmentation

For example, as shown inFig. 3, *r*2 and *r*4 are wrongly mapped road segments, the actual route is along the horizontal road from *qs* to *qd*. The map matching error results in that *r*2 and *r*4are recognized as landmarks and brings noise when estimating the travel time, e.g., the real travel time for *r*2 → *r*3 is very likely to be much longer than the estimated time due to the map matching error, which leads to *r*2 → *r*3 becomes a part of this rough route.

IV.C TRAJECTORY ANALYSIS

The Valuable information like traffic reports must be converted into raw trajectories for decision making purpose. When monitoring traffic, a good indication of behavior is motion. Motion is captured by trajectories which indicate the spatio-temporal characteristics of objects and encode behavior. A key observation for trajectory analysis is that typical actions are repetitive while the unusual do not occur often. This indicates that through sufficient observation one is likely to observe and can learn all the prototypical behaviors for a given scene. In order to learn typical patterns, a training database of trajectories is accumulated. In this module the objects and trajectories were stored in the dataset. From the dataset the trajectory status can be monitored.

IV.C.A ALGORITHM:

In this step, we aim to automatically partition time of a day into several slots according to the traffic conditions reflected by the raw samples pertaining to a landmark edge. Then we estimate the travel time distribution of each time slot for each landmark edge.

```
Algorithm 1: LocalSmoothing
Input: a sequence L = l_1 → l_2 → l_3 → ... → l_{n-1} → l_n,
       dist(l_i,l_j), i, j = 1, 2, ..., n
Output: a subsequence (of L)
       L' = l'_1 → l'_2 → l'_3 → ... → l'_{m-1} → l'_m that satisfies:
       ∀i = 1, 2, ... m − 1, dist(l_i, l_{i+1}) = min_{j>i}{dist(l_i, l_j)}
for i ← 2 to n do
    for j ← i − 1 downto 1 do
        if SL(j) == ∅ then
            Insert the sequence l_j → l_i to SL(j)
        else
            Binary search in SL(j) for the largest integer p such
            that dist(l_j,l_{j_p})≤dist(l_j,l_i), if no such value exists,
            p := 0;
            /* where l_j → l_{j_p} → ··· is the p-th
               sequence in SL(j)                                  */
            Insert l_j → l_i after the p-th sequence of SL(j);
            /* if p == 0, insert l_j → l_i as the first
               element of SL(j)                                   */
            for w ← 1 to p do
                if SL(j)^(w) ⊖^l l_j ⊕^r l_i ∈ SL(j_w) then
                    /* SL(j)^(w) == l_j → l_{j_w} → ··· is the
                       w-th sequence of SL(j)
                       SL(j)^(w) ⊖^l l_j ⊕^r l_i represents that
                       the SL(j)^(w) removes the first
                       landmark l_j from the
                       beginning(left) and adds l_i to
                       the end(right), i.e.,
                       SL(j)^(w) ⊖^l l_j ⊕^r l_i == l_{j_w} → ··· → l_i
                       */
                    SL(j)^(w) := SL(j)^(w) ⊕^r l_i;
                    /* add l_i after the sequence
                       SL(j)^(w), i.e.,
                       SL(j)^(w) := l_j → l_{j_w} → ··· → l_i        */
    return The longest sequence L' in {SL(i)|i = 1, 2, ..., n}
```

Fig 4. Algorithm

IV.D LOCATION ANALYSIS

Collected Raw Trajectories represent time stamped geographical locations. Apart from storing raw data in the moving object database it needs to reconstruct the trajectories. Raw points arrive in bulk sets, it needs a filter that decides if the new series of data is to be appended to an existing trajectory or not. The road users too need good quality traffic information in order to plan and adjust their routes. Traffic information has traditionally been collected with inductive-loop detectors. Location updates may help to monitor accurate location analysis. This will helps to gather all relevant information's about the analysis.

IV.E TRAJECTORY UPDATION

Road condition and road network structure change over time. Rather than immediately rebuilding a Route Overlay upon changes, which is expensive, we develop several techniques to incrementally update Route

Overlay for edge distance changes, and network structure changes. The proposed thesis adopts a filtering-and-refreshing approach. In the "filtering" phase, shortcuts possibly affected by an edge change are identified. Only the identified shortcuts are updated in the "refreshing" phase. Collected Raw Trajectories represent time stamped geographical locations. Apart from storing raw data in the moving object database, we also need to reconstruct the trajectories. Raw points arrive in bulk sets, we need a filter that decides if the new series of data is to be appended to an existing trajectory or not. This filter,1)Tolerance distance,2)Temporal gap,3)Spatial gap,4)Maximum

V. CONCLUSION

This paper describes a system to find out the practically fastest route for a particular user at a given departure time. Specifically, the system mines the intelligence of experienced drivers from a large number of taxi trajectories and provide the end user with a smart route, which incorporates the physical feature of a route, the time-dependent traffic flow as well as the users' driving behaviors (of both the fleet drivers and of the end user for whom the route is being computed).We build a real system with real-world GPS trajectories generated by over 33,000 taxis in a period of 3 months, then evaluate the system with extensive experiments and in-the-field evaluations. The results show that our method significantly outperforms the competing methods in the aspects of effectiveness and efficiency in finding the practically fastest routes.14Overall, more than 60% of our routes are faster thanthat of the existing on-line map services, and 50% of these routes are at least 20% faster than the latter. On average, our method can save about 16% of time for a trip, i.e., 5 minutes per 30 minutes driving.

VI. FUTURE ENHANCEMENT

Here objects are considered as small vehicles so it is not much efficiency to find traffic free path, In future work identifies the types and length of vehicles then will give the traffic information to user. And also these extend the time dependent land mark graph to used for comparison between different route traffic data. And then find out the particular vehicle information such as speed, travelling time and also gives movement of that vehicle and its route information with use of GPS.

VII. REFERENCES

[1] Jing Yuan, Yu Zheng, Xing Xie, Guangzhong Sun "T-Drive: Enhancing Driving Directions with Taxi Drivers' Intelligence" in *Proc. MDM*, 2013.

[2] J. Yuan, Y. Zheng, C. Zhang, W. Xie, G. Sun, H. Yan, and X. Xie,"T-drive: Driving directions based on taxi trajectories," in *Proc.GIS*. ACM, 2010.

[3] T. Hunter, R. Herring, P. Abbeel, and A. Bayen, "Path andtravel time inference from gps probe vehicle data," in *Proc.NIPS*, 2009.

[4] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang,"Map-matching for low-sampling-rate gps trajectories," in*Proc. GIS*. ACM, 2009.

[5] J. Yuan, Y. Zheng, C. Zhang, and X. Xie, "An interactive-votingbased map matching algorithm," in *Proc. MDM*, 2010.

[6] E. Kanoulas, Y. Du, T. Xia, and D. Zhang, "Finding fastestpaths on a road network with speed patterns," in *Proc. ICDE*,2006.