Enhancement of Trajectory Mining for Map Pattern using GPS/GIS
Ms. Rupali H. Asai1
P.G. Student
Department of Computer Engineering
D.P.C.O.E. Savitri Bai Phule Pune University Pune, Maharashtra, India

Abstract— A new method has been introduced which determine and investigate the unknown hidden information of large object route data across the world (e.g., vehicle). GPS technology is used for calculating the geographic coordinates (i.e., latitude and longitude) to road names reflecting the related some semantic information which is transforms by this approach. One article is considered as a object movement that consist some hidden information like path name that enable the semantic study of huge object data sets as a article amounts. Textual area modeling techniques are used for invisible the themes like object covering area, are documented but in secure manner which is invisible for third party. In system there is an interactive visualization tools, like vehicle area maps, area routes, street clouds and parallel coordinates, to visualize the probability-based current information. Travelers, and drivers can carry out their mixture of knowledge discovery tasks with direct semantic and visual assists by using these visual analysis system. The use of Global Positioning System technology is that collecting the GPS trajectory data and visualize the map of any object moving in that trajectory. The system has been also used to show changes in position, speed, time and direction of object moving in particular path. These visual analytic system provides security which is more important for any person location and vehicle location. Geographic components in collected GPS data and visualizing it by mapping provides a complete view of any area where the object is moving. Some secure algorithms provide security to hidden information to a particular vehicle.

Key words: GPS, Pattern Mining, Object Mapping, Speed Analysis, Direction, Latitude, Longitude, Road Usage Pattern Privacy Preserving and Trajectory

I. INTRODUCTION

A new visual analytic system is developed. Trajectory is the path where any object can moved and mining means extracting hidden data from object. The use of GPS technology is that collecting the GPS related trajectory data & visualize the map of object moving in that trajectory. Here tracking of any object is basic research but the disadvantage is that it is not secure it is usable by all people in which the user can use it for wrong intension. Therefore here the main research is on location security and its hidden information security which is not accessible by third party. These system is also used to show changes in position, speed, time & direction of object moving in a particular path. Using the geographic components in the collected GPS data & visualizing by mapping, provides a clear view of any region in which the object is moving. When any object moving in a particular path it will be visualize in visual analysis system. For visualizing object it is necessary to track those object. For tracking the object GPS technology is used here. The original map of object from source to destination is saved in database & current map followed by object will be visualize on the system.

II. PROBLEM STATEMENT

The aim of the system is to provide practical privacy preserving techniques to solve the FRVP problem, such that neither a third-party, nor participating users, can learn other users locations; participating users only learn the optimal location.

A new visual analysis system is developed for a particular organization that visualize the different route map of object travelling in whole day & night, but here visualization is allowed for a particular group of member in organization so that third party can not hack the object location, for that here the main research is on security provided to the system, also it determined the direction, speed, distance of that object & time travel from source to destination.

III. PROPOSED SOLUTION

By using GPS technology visualizing the map of any object moving in a specific path. For knowing the position of object it is necessary to track the location. By mapping any object it will be easy to determine the original path and current path of any object. According to that, speed and direction of object can also be determined. From this it will be analyze that the original time required to object when moving from source to destination is similar to time required to the same object moving from same source to destination. If yes then no problem if no then some problem can occurs in path where the object is moving. Here there are two possibilities can happened.

1) Object can reach to their destination in less time period compared than original time period. In these case object can apply shortest path to reach their destination.

2) Object can reach to their destination in more time period compared than original time period. In these case object can face the traffic while moving in link.

And main research is on location security and its hidden information security which is not accessible by third party.

A. Approaches for solving the problem and efficiency issues:

1) The privacy preserving FRVP (PPFRVP) problem is general enough and nicely captures the computations and privacy requirements. The user privacy of any PPFRVP algorithm measures a probabilistic advantage that an adversary a gains towards learning the preferred location of at least one other user, except the final rendezvous location, after all users have participated in the execution of PPFRVP protocol.
2) The Secure Hash Algorithm is a family of cryptographic hash functions which provides security.
3) Privacy-preserving algorithms for determining an optimal meeting location for a group of users.
4) GPS algorithm is used for determining geo-coordinates i.e. latitude and longitude.
5) BGN Algorithm is Boneh-Goh-Nissim algorithm which is also used for privacy preserving concept. Here the distance computation module uses BGN algorithm. It is used for encrypting the data.

IV. LITERATURE SURVEY

In visualizing hidden themes of taxi movement with semantic transformation, Ding Chu [1] developed a new visual analytics system in which A new vehicle GPS device on a taxi can trace its real time moving path (i.e. trajectory) as a series of positions sampled with a small periodic interval. At each location, GPS information and meta information are stored such as time, geographical coordinates (latitude, longitude), speed, direction.

In Using Incomplete Information for Complete Weight Annotation of Road Networks, Bin Yang, Manohar Kaul, and Christian S. Jensen [2] presents a general framework that takes as input a collection of (trip, cost) pairs and assigns trip cost based weights to a graph representing a road network, where trip cost based weights may reflect GHG emissions, fuel consumption, or travel time. By using the framework, edge weights capturing environmental impact can be computed for the whole road network, thus enabling eco-routing.

The problem of mining Frequent Sequential Patterns (FSPs) from deterministic databases has attracted a lot of attention in the research community due to its wide spectrum of real life applications. For example, in mobile tracking systems, FSPs can be used to classify or cluster moving objects. In Mining Probabilistically Frequent Sequential Patterns in Large Uncertain Databases [3], Zhou Zhao, Da Yan, and Wilfred Ng studied the problem of mining FSPs in the context of uncertain sequence data.

Yanmin Zhu, Yuchen Wu, Bo Li [4] from Trajectory Improves Data Delivery in Urban Vehicular Networks demonstrate the strong spatiotemporal regularity with vehicle mobility by entropy analysis with the extensive data sets of real taxi and bus traces. This demonstrates that predicted trajectories do help data delivery in vehicular networks. The location tracking functionality of modern mobile devices provides unprecedented opportunity to the understanding of individual mobility in daily life.

In The Places of Our Lives: Visiting Patterns and Automatic Labeling from Longitudinal Smartphone Data, Trinh Minh Tri Do and Daniel Gatica-Perez [3] presents a study on place characterization in peoples everyday life based on data recorded continuously by smartphones. This study contributes the understanding of places in peoples daily lives and to the possibility of inferring place categories from smartphone data.

In T-Watcher: A New Visual Analytic System for Effective Traffic Surveillance [6], Jiansu Pu, Siyuan Liu presents an interactive visual analytics system, T-Watcher, for monitoring and analyzing complex traffic situations in big cities. The movement patterns are important for traffic analysts to understand the behaviors of moving objects especially in transportation management. Monitoring and analyzing trajectory data could be used to reason about mobility and to support experts in traffic analysis with reliable information.

Tarik Crnovrsanin, Chris Muelder, Carlos Correa, Kwan-Liu Ma [7] from Proximity-based Visualization of Movement Trace Data there are the visualization of movement traces in time is a challenging problem. Traditional approaches to the problem map 2D and 3D space directly and use time as either an extra dimension, with coordinated views, or animation.

In A Survey Paper on Trajectory Pattern Mining for Pattern Matching Query, S. R. Ghule & S. M. Shinde [8] presents the frequent pattern-based classification could exploit the state-of-the-art frequent pattern mining algorithms for feature generation. Once frequent patterns are found out, these patterns can be used for various applications. So, on finding frequent path travelled by vehicle, we can publish advertisements in that area with the help of two components, one is gateway to store information of vehicles and second one is advertisement server which will send the data to gateway and finally reaches to user, may be through mobile phone.

Mr. M. Jayakameswaraih & Prof. S. Ramakrishna from Mining Frequent Trajectory Pattern Co-ordinates in Spatial-Temporal databases using Apriori Algorithm [9] studied that trajectory is defined as the path of flying object; the path that a projectile makes through space under the action of given forces such as thrust, wind, and gravity. It is also defined as the curve intersecting at constant angle; a curve or surface that intersects all of a family of curves or surfaces at a constant angle.

The implementation uses a Global Positioning System (GPS) which helps in identifying the area in which the object is currently navigating. P. Hemalatha [10], from Implementation of Data Mining Techniques for Weather Report Guidance for Ships Using Global Positioning System presents the weather report of the ships location(example) is made to traverse the Decision Tree and the corresponding decision is passed to the ship for its safe navigation. Thus this implementation, which uses many advanced concepts such as Data Mining and Global Positioning Systems, can also be extended for Aircrafts, Vehicle Tracking, Submarines, etc.

V. SYSTEM ARCHITECTURE

A detail system architecture is shown in the below figure. It shows the complete architecture of visualizing system and flow of data. These system consist of two main modules that are Server side and User side.

In server side there is an AMP mail port in which each user in group have to create his/her own email-id. If there is any new users then he/she have to register by giving some important information. After creating email-id, on AMP mail port it consist different options to user like Private code, Mail box, Archives. In private code it create security code for each user which is essential for user when he/she enter in the visualizing system. Then mail box contains the mail received from receiver. That mail box contain decryption key for each user who is in location searching group. These decryption key is used for visualizing the location of a particular vehicle which is send
by sender. After entering that decryption key in visualizing system a member or user in that group can see current vehicle location and its hidden information.

VI. MATHEMATICAL MODEL

At high traffic densities (signifying traffic jams), links have very low operational exit rates and at low densities, the exit rate varies linearly with the traffic density. Each traffic link reaches an optimal capacity at a corresponding optimal operating density. To formalize this notion, there are a few parameters that describe the state of a link $l$:

A. Capacity:

This is the maximum number of vehicles that the link can hold. It does not vary over time, and is represented as $MAX_l$.

B. Buffer Size:

This is the number of vehicles on the link at any given point of time. However, traffic density is defined as number of vehicles per unit of road. Thus, buffer size is simply a multiple of traffic density($x$ link length). Hence, the terms buffer size and traffic density are used interchangeably, represented as $Bl$.

This is the rate at which a vehicle can exit a link, and is represented as $C_l$. As mentioned earlier, it depends on $MAX_l$, as well as $B_l$. The point at which this transition takes place has the highest exit rate, denoted by $C_l$. The corresponding traffic density is denoted by $B_l$. Previous research has shown that $B_l$ is approximately equal to $\frac{MAX_l}{x}$. Consider the case when the traffic density is $B_l$, and hence the exit rate is optimal at $C_l$. If a short burst of traffic enters the link and temporarily pushes the traffic density $B_l$ to be more than $B_l$, the exit rate $C_l$ will drop below $C_l$. This decreased exit rate will further increase $B_l$. This domino effect leads to the exit rate decaying rapidly, and we call it congestion collapse. An important point to note is that even if the input rate is greater than the maximum exit rate $C_l$, congestion does not take place until the traffic density becomes greater than $B_l$. Consider a congested link whose $B_l$ value is greater than $B_l$. Even if the input rate is reduced to below $C_l$, we are not guaranteed to get rid of the congestion in a short period of time, because the exit rate $C_l$ at a traffic density of $B_l > B_l$ will be lesser than $C_l$. This is the reason that once congestion collapse takes place, it stays that way for long periods of time.

We primarily concentrate on two specific parameters of $Haft$: $C_l$ $(Haft)$ and $B_l$ $(Haft)$. If the sum total of the exit rates of $Hbef$ and $L$ is always less than the optimal exit rate $C_l$ $(Haft)$, then the merging never faces a congestion problem. If however, the sum of the input rates of $L$ and $Hbef$ is larger than $C_l$ $(Haft)$, then the buffer size of $Haft$ grows. If the buffer of $Haft$ grows beyond $B_l$ $(Haft)$, then the exit rate of $Haft$ begins to drop thereby, triggering the spiralling effect. A small burst in traffic beyond $C_l$ $(Haft)$ is sufficient to drive the system to operate at a low-capacity point. Once reached, the system will continue at this operational point and if the total input rate is greater than the operational exit rate, the congestion increases and spreads into the links $Hbef$ and $L$.

The traffic density in night time is given by:

$$Density = \sum_{i} \frac{H}{count(i) - \left(\frac{P_i}{P_2}\right)h_{i}}$$

where $i$ is the $i$th pixel line and $count(i)$ represents the number of white pixels in that pixel line. $H$ denotes the actual height of the camera, $P_i$ and $P_2$ represents the projection of the $i$th pixel line and the last pixel line on the camera approximately. $h_{i}$ Represents the observed height of complete road length in the image and is given by:

$$\frac{X_{\text{max}}}{X_{\text{max}} + d} = \frac{h_{i}}{H}$$

Where $X_{\text{max}}$ represents the farthest visible point on the road segment and $d$ represents the nearest visible point of the road segment measured relative to the position of the camera.
We also place two emergency signals on L and Hbef a little distance before the actual merge point. When there is no congestion on Haft, these signals are green, thereby allowing the free flow and merging of traffic. As soon as the traffic density $B_t (\text{Haft})$ begins to approach $B_T (\text{Haft})$, we need to throttle down the incoming traffic. The signals at L and Hbef therefore change to red for certain durations to allow the density $B_T (\text{Haft})$ to reduce. The traffic density at an epoch $t + 1$ is related to the that at epoch of time $t$ by the following simple equation:

$$B_t (t + 1) = B_t (t) + \text{input} - \text{output}$$

This equation is used to determine how much traffic can enter in an epoch from both L and Hbef; which determines the amount of time that the signals can be kept green. In order to calculate how to divide the green time between the signals, a simple (yet, effective) method is to divide the total permissible green time amongst them in a ratio that is proportional to their traffic densities (which in turn, are monitored by sensors). This mechanism guarantees that the link Haft will never go into congestion collapse.

VII. CONCLUSION

Finally it is conclude that these system is very helpful in now a days, these system track the vehicle location with providing security to all passengers. It is useful from criminal attacks also, it also reduces time and improve efficiency and fulfill customer satisfaction.

VIII. FUTURE ENHANCEMENT

In the future, there will be investigation for security providing to all vehicle except that for a particular vehicle in a particular private agencies. Also enhance the models for taxi data, and study about semi-automatic parameter for optimal number of topic. Also will integrate more visualization techniques into our system with providing more security for general vehicles also with special vehicles.

REFERENCES


[3] Zhou Zhao, Da Yan, and Wilfred Ng, Mining Probabilistically Frequent Sequential Patterns in Large Uncertain Databases, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 5, MAY 2014.


