GPS – GSM BASED HYBRID SYSTEM

[1] Mr. S.K.CHAUHAN, [2] Dr. K. R.PARMAR,
[1]C.S.E.[Communication System Engineering] Student, Department Of Electronics and Communication, L. D. College of Engineering, Ahmedabad, Gujarat.
[2] Asst.Professor And Head Of the Department , Department Of Electronics and Communication, L. D. College of Engineering, Ahmedabad,Gujarat.

Siddharth_gec@yahoo.co.in , parmar.kr@ldce.in

ABSTRACT: *This paper discusses the tasks of place discovery and place recognition — learning and recognizing places significant to a user -— by analyzing GPS data and partitioning into cluster region using a variant of k means clustering algorithm and build a markov model and travel summary of data to identify significant locations.. I have developed place discovery algorithms (geometric algorithm) .Geometric algorithm are most commonly used in coordinate based systems and identify significant places in terms of circles and polygons in reference to absolute location coordinates. And these clusters are examined in temporal sequence to build a Markov model to predict future transitions between locations. It is a survey and critique of clustering algorithm used in place discovery systems. A prototype system to automatically collect, analyse and build travel summaries of GPS location data. Location provides valuable context into the user's environment, and place-discovery and recognition algorithms enable human centric systems to communicate with the user in human terms. In this paper, we introduce a novel two-phased approach to place-discovery and recognition that uses GPS cell data. Outputs such as clusters demonstrating significant places on Map , application of algorithm on graph, markov model , graph finding optimum radius, removal of transitional clusters , and Travel summary is obtained .Therefore in total 6 outputs are obtained and are shown last.*

Keywords— Global Positioning System(GPS)

INTRODUCTION

Location can provide valuable contextual clues about a user's environment, and research in spatial cognition shows that people naturally structure their experiences in terms of socially and personally meaningful places. In order to communicate with the user on more human terms, human-centric systems stand to benefit from place discovery - the ability to analyze a user's everyday routines and extract the abstract notion of a significant place, and place recognition — the ability to recognize when the user has returned to a significant place.

Place discovery and recognition systems provide valuable functionality with which to manipulate the user's past, present, and future. Place discovery in the past provides a complete, accurate record of places where the user has spent time; querying the past can provide context to deferred events that the user may not remember in detail (e.g. such as what the user was doing during a missed call). Place recognition and discovery in the present localizes the user and enables a device to make intelligent decisions based on the user's current context (e. g. a mobile phone silences itself when it detects that the user has entered a conference room/ auditorium; displaying a to-do list at the shopping center). The past and present can then give context about the user's future, enabling an entire class of "just-in-time" services such as trafiic and schedule updates in navigation systems, location-based reminders, and schemes for reducing communication cost.

Mobile phones are a natural platform for tracking user location for the purposes of place discovery and recognition. Though conceptually simple, it is challenging in practice to rely on any single location-sensing technology for accurate place discovery due to a number of trade-offs inherent in the location data (e.g. loss of GPS signal, location ambiguity due to large cell tower coverage, lack of wireless infrastructure). And a useful place-discovery system must be effective subject to the constraints on the mobile phone's hardware resources (limited battery and storage).
This paper describes a novel approach to place-discovery that meets the following requirements. Our approach uses GPS to provide a first estimate of the state space for a user's significant places.

Automatic collection of location data

The collection of location data must be passive and must not hinder the user's everyday activities with the mobile device.

Compact data storage

Location data can be huge if stored naively, and so we must store a summary of location data as well as statistics of the user's significant places in compact ways.

Minimize the usage of additional service infrastructure

Battery life is a major resource constraint on mobile devices. Factors such as continuous computation and broadcasting on radio frequency mediums affect battery drain rate.

Robustness to noisy data

Certain kinds of data (e.g. GPS) are subject to signal loss and lack of precision (cellular coverage). Our place-discovery algorithm must minimize the effect of noisy data in producing false-positive points of interest.

There are two general classes of place-discovery algorithms: ***geometric algorithms*** and ***fingerprinting algorithms*.**

Geometric algorithms are most commonly used in coordinate-based systems and identify significant places in terms of circles and polygons in reference to absolute location coordinates. Fingerprinting algorithms are typically used in landmark/beacon-based systems (in which no explicit geographic topology is known) and identify significant locations in terms of unique "signature" sequences of metrics such as cell tower id or signal strength patterns.

Camparision of different existing technologies

Tech	Precise Locatio n	Low infra Demand	Low Power Usage	Coverage	Work Indoor
GPS	Good	Good		Good	
GSM		Good	Good	Good	Good
802.11 WiFi					Good
Bluetooth			Good		Good

GEOMETRIC PLACE DISCOVERY

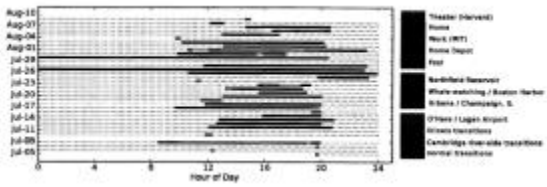
Perhaps the pioneering work on place-extraction is Marmasse and Schmandt's work on the system. Their key insight is that the loss of GPS signal signifies an important place because it indicates that the user has arrived inside of a building. Unfortunately, this approach does not identify significant places in outdoor settings (e.g. parks, camping trails), and is prone to produce false-positives due to "urban canyon" effects occluding the GPS signal. Ashbrook and Starner improves place-extraction by pre-processing GPS data to filter for only when user movement is detected and identifies points as significant if the user's dwell-time is greater than 10 minutes. In their study, a user's significant places are identified using a variant of the well known k-means clustering algorithm, and these clusters are subsequently examined in temporal sequence to build a Markov model to predict future transitions between locations. However, k-means clustering requires the number of clusters to be known before clustering, favors symmetrically- shaped clusters (e.g. circles, spheres) and is sensitive to noise in the data. We follow a similar approach to place-discovery in this thesis, but incorporate the predictive model to further distinguish between clusters that are significant locations and "clusters"
where the user is actually in transit between two locations.

Density-based clustering approaches such as Density-Join clustering (DJ-Cluster) use the density of local neighborhoods of points for place discovery, and is an improvement over k-means clustering due to its robustness to noise and anomalous points. In contrast to k-means clustering, density-based clustering can identify clusters of arbitrary shape. Here, density is defined by two parameters — the radius of a circle and the minimum number of points within the circle. Density-based clusters have a notion of "reachability" that allows a chain of connected, merged circular neighborhoods to form a cluster of arbitrary shape, and clusters identified by DJ- Cluster have considerably higher density than points outside of the cluster. Zhou et al. later devised the Time Density Join (TDJ) and Relaxed Time Density Join algorithms as improvements to DJ-Cluster by adding additional parameters such as relaxed time constraints on the formation of a neighborhood/cluster to capture often-visited locations with small dwell-time. Kang simplifies place extraction by contributing an incremental clustering algorithm that identifies places with arbitrary shape from traces of Wi-Fi location data observed from PlaceLab. The drawback of this algorithm is that its temporal clustering requires frequent data sampling of one GPS reading per second, which can result in significant battery life depletion for a mobile device.

GPS Place Discovery

We are interested in learning locations with high dwell-time — e.g. where the user spends the most time. We use a k-means clustering approach similar to that of Ashbrook and Starnerin order to learn important locations from a user's GPS log traces . A cluster is a circular region defined by a center and a

radius. The k-means clustering algorithm takes as input a radius r and the set of points to be partitioned into clusters of the input radius. The algorithm iteratively forms new clusters and assigns points to each cluster until all points are covered by a cluster.



Travel summary indexed shown per day breakdown of users travel

Cluster formation starts by choosing a data point p at random and defining a cluster with center at p. The algorithm marks all points within r distance of p and calculates the average center of all of the points in this cluster. If the average center changes, the cluster has moved from its initial position, and the algorithm iteratively scans the dataset to mark new points within r of this new center, recalculating the average center at the end of each scan. When the average center is stable, the formation of the current cluster is complete, and there are no unmarked points in the dataset that also belong to the current cluster. The algorithm then removes members of the completed cluster from consideration, and repeatedly forms new clusters until there are no points to consider.

Algorithm 1

Variant of the k-means clustering algorithm. Iteratively partitions the input data into clusters of radius radius.

Compute_clusters (yps_data , radius)
1: Filter gps_data for only points when user speed > 1 mile/hour
2: places <= Filter gps_data for points with dwell time > 10 minutes
3: clusters <= []
4: while places not empty do
5: center <= choose a random point p from places
6: neighbors <= find all points within radius of center
7: avg_center ¢ mean center point of neighbors and center
8: if avg_center == center then
9: add neighbors to clusters
10: remove all members of neighbors from places
11: else
12: center <= avg_center
13: Goto line 6
14: end if
15: end while
16: return clusters

Our clustering approach preprocesses the GPS log data according to the following assumptions in order to reduce the number of irrelevant points in our partitioning:

1. The user does not move at high speeds when at a high dwell-time location. Thus we need not consider points that are far below the average human's walking speed (e.g. less than 3 miles per hour)
2. Filtering by walking speed may create gaps in the dataset where the user stays at a significant location. Our preprocessing step identifies these gaps as points of interest by filtering for points with dwell-times longer than 10 minutes.

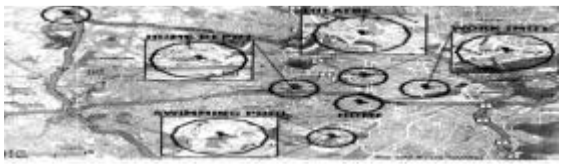
One drawback of this algorithm is that it must iterate until all points in the dataset are assigned to a cluster. As a result, the algorithm produces frequently used transition paths between two significant places as clusters. In order to determine which clusters are meaningful vs. transitional, we employ metrics such as probability of visits and total/ average dwell-times in our prediction modeling to reduce the number of erroneous clusters identified as significant locations.

User Movement Prediction

The clustering algorithm gives a list of clusters that account for all GPS points in our data log. Many of these clusters account for transitions between important locations, and we take further steps to identify locations where the user spends the most time by sequencing the clusters into a travel sequence and building a Markov model for the user's movement.

To obtain the travel sequence, assign each cluster a unique ID, and substitute each point in the original chronologically-ordered GPS trace with the ID of the enclosing cluster. This gives a chronological sequence of visited clusters during the user's travels.

During construction, the following statistics are maintained for each cluster:
- *total_dwell_time* - total amount of time spent in the cluster
- *average_dwell_time* - average dwell time (helps to identify transitional clusters)
- *nvisits* - number of visits to this cluster (transitions between the same cluster do not increment nvisits
- *intervals* - the list of (arriva1_time, exit_time) pairs denoting periods when the user was in the cluster



Formation of Clusters

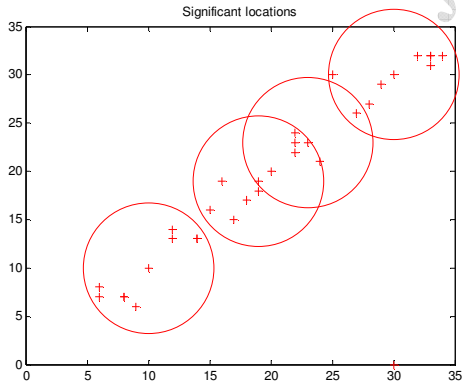
Transition	Relative Frequency	Probability
Home → *Home*	16/44	0.3636
Home → *c1*	1/44	0.0227
Home → *Work*	12/44	0.3636
Home → *c5*	1/44	0.0227
Home → *c9* (false-positive)	2/44	0.0455
Home → *c10*	1/44	0.0227
Home → *c11*	1/44	0.0227
Home → *Pool*	3/44	0.0682
Home → *c15*	1/44	0.0227
Home → *c16*	1/44	0.0227
Home → *Store*	2/44	0.0455
Home → *Airport*	2/44	0.0455
Home → *c27*	1/44	0.0227

Chronological travel sequence

Next, a *Markov model* is built from the chronological travel sequence, in which clusters are states, and transition probabilities between clusters are calculated via frequency counting . The Markov model assists in defining clusters that represent significant places — transitions to and from significant places occur with greater than random probability among all observed transitions in the travel sequence, and clusters that represent significant places have much higher total and average dwell-times.

RESULTS AND DISCUSSIONS

First GPS coordinates are plotted on ordinary plot in MATLAB and k means clustering algorithm is applied for various cluster radius. The subject strove to carry mobile and a Bluetooth GPS receiver all the times. Below figure shows GPS coordinates with clusters formed.



Clusters formed around significant places by Applying algorithm. Circles represent significant places.

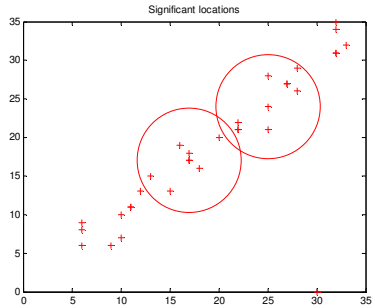


Figure 1 GPS data with clusters formed. Then k-means clustering algorithm is executed for several other cluster radius and graph of no of clusters vs radius is produced.

To obtain the optimal cluster radius, we calculate the number of clusters found for a range of input radii and look for the radius jst before the number of clusters converges to number of points in our dataset.Naturally large cluster radii may result in plac discovery that is too coarse grained and small radii may produce too many individual clusters

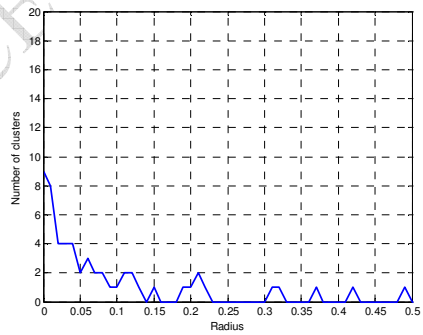


Figure 2 Number of cluster vs cluster radius (Selecting Optimum radius)

It shows as cluster radius increases number of clusters approaches zero.Then GPS data is plotted on the map and then k means clustering applied on that to get clusters.

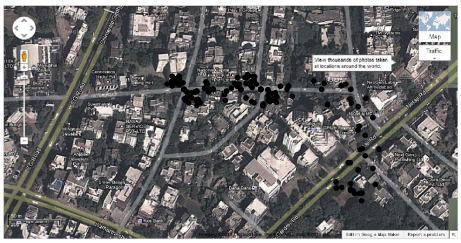
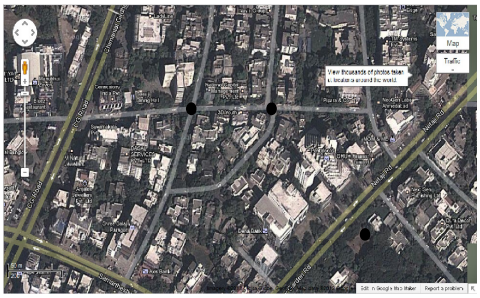


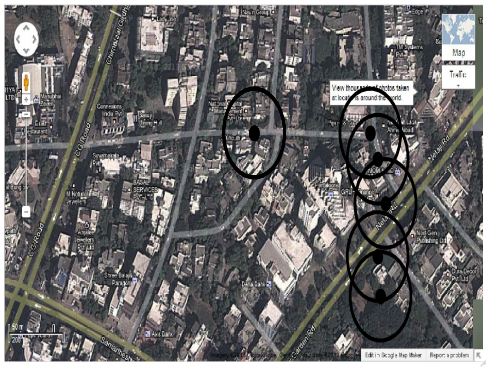
Figure 3 GPS data on map

To obtain travel sequence, assign each cluster a unique ID, and substitute each point in the original chronologically ordered GPS trace with ID of the enclosing clusters.



Places significant to user

In order to determine which clusters are meaningful vs transistional, we employ metrics such as probability of visits and total dwell times in our prediction modeling to reduce the number of erroneous clusters identified as significant locations.



Significant Clusters formed along with transitional clusters

Cluster formation starts by choosing a data point p at random and defining a cluster with center at p. The algorithm marks all points within r distance of p and calculates the average center of all of the points in this cluster. If the average center changes, the cluster has moved from its initial position, and the algorithm iteratively scans the dataset to mark new points within r of this new center, recalculating the average center at the end of each scan. When the average center is stable, the formation of the current cluster is complete

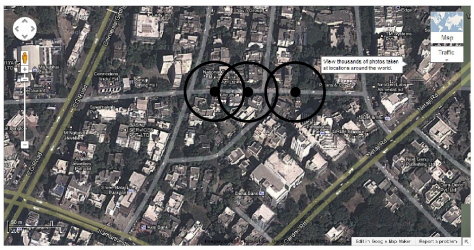
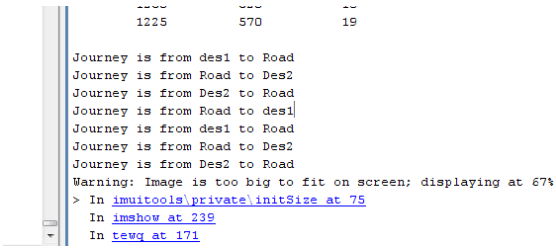


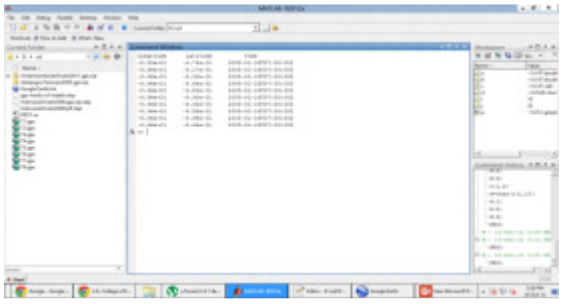
Figure 4 Clusters significant to a user
In practice the algorithm converges very quickly eg the number of iterations is much less than the size of GPS datasets to be clustered.



Output showing transition journey of user obtained in Matlab. The information yields a travel summarizes when the user has spent time in each clusters

Other Snapshots

GPX filed downloaded and plotted on google map.While the clustering algorithm works especially well when spherical clusters are naturally available in the data the amount of overlapping shown suggests that location data may not conform to spherical clusters.



Output of matlab showing latitude longitude co ordinates along with time

CONCLUSION

We design, implemented Geometric clustering algortihm and evaluate a system that learns places significant to a user by analyzing a log of the users

GPS locations. We were able to plot graph of no of clusters vs clusters radius based on which optimal radius was calculated. We were also able to apply algorithm through matlab programming and able to obtain clusters around places significant to a user. Further we were also able to eliminated transistional clusters by applying 10 min dwell time concept. Also Markov model was build based on travel sequence summary which also helped to eliminate transistional clusters. GPX files where also plotted on google map along with matlab output of latitude and longitude. We have also shown output of matlab code showing travel path of user ie journey of user. Place discovery in past provides a complete record of places where the user has spent time, in present make intelligent decisions based on users context and in future enalbling services such as traffic and schedule updates.GPS data alone can't be of any use to the user.

REFERENCES

1)Xiao Yu at Department of Electrical and Computer Engineering, LEARNING SIGNIFICANT USER WITH GPS-GSM LOCATIONS, 2011 at Massachusetts Institute Of Technology.

2) Mohammad Khedher, Mechatronic Dept Jordan Hybrid GPS-GSM Localization of Automobile Tracking System , *International Journal of Computer Science & Information Technology (IJCSIT)* Vol 3, No 6, Dec 2011.

3) D.Krishna Reddy ,P.S.Ravikanth ,A.D.Sarma.Integration of GPS and GSM for determination of cellular coverage area, *R & T Unit for Navigational Electronics (NERTU), Osmania University, Hyderabad, 2011*
Chaitanya Bharathi Institute of Technology, Hyderabad , 2011,India

4) Daniel Ashbrook and thad starner. Learning significant locations and predicting user movement with gps. In ISWC, pages 101-108 *IEEE* Computer Society 2009

5) Len liau, Dieter fox and henry A. Kantz Learning and Inferring transportation routines. In Deborah L. Mc Guinness and George Ferguson, editors, AAAI, pages 348-353, AAAI Press. The MIT Press, 2010.
.