www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume- 3 Issue -3 March, 2014 Page No. 5000-5003

Mining and Clustering of mobile transaction sequences in location based service environment

Vaishnavi M. Naik, Ketki S. Kulkarni, Dipali M. Chordiya, Sreeja Kumar, Prof. Poonam Chaudhari
Dept. of Computer Engineering GES's RHS COE MSR
Nashik, India
v.vaishnavi19@gmail.com
k.ketki4@gmail.com
Chordiya.dipali@gmail.com

Abstract— **The improved capabilities of the recent cellular phones compared to older ones had lead the telephone companies offer a lot of value added services. Emergence of new technologies will lead to emergence of new services increasing the percentage of users for such services. To improve the quality of services provided for the users, mining and clustering of different mobile transaction behavior are useful. For clustering we have LBS based algorithm. LBS based services which are in demand are used to find a location, people, services and to find similarities to predict next behavior of mobile user.**

Keywords- Mobile commerce environment, Location Based services.

Introduction

Wireless networking is becoming a reality , by having recent advances in devices , various middleware and network infrastructures. Tremendous growth in wireless communication implies that people are able to do business in mobile commerce environment. With introduction to such advance technologies, consumers of mobile devices will be able to access content and services anytime and anywhere.[3]

Mobile banking, mobile browsing, location based services etc are products of mobile commerce, location based service being a major product, attracting interest from both industries and research. Location based service not only enable the mobile users to identify the information about geographical point but also determine the location of user.[6].In location based services, one must reveal their exact location, the services accessed from their private information. [5].

Service providers offer various services to the user. These services include finding route, calculation of position, searching specific information of objects of user's interest where geographical based data are requested by business and industry partners. The major work deals with prediction of services which are requested by mobile user in LBS environment. The system should predict the next behavior of the user using techniques like clustering, segmentation of time interval setting a various prediction strategies providing users a precise and efficient mobile behaviors prediction system.[6]

Related Work

Today due to the advancement in wireless technology the concept of LBS has come into existence. Thus the Application providers are trying to provide various advanced and useful services for the users. Focus is been given by the researchers to generate mobile patterns from the whole data logs.

Initially in 1990's the concept of data mining emerged with rapid advancements. There was a wide increase in computing for various applications and hence data mining grown rapidly. . It was at this time Chen at al introduced the path traversal patterns for mining user behaviors. Many researchers worked on and proposed various algorithms. [4]

-To search for time ordered patterns , Sequential pattern mining was introduced named " Sequential Patterns within transaction Databases".[1]

- To discover movements of the user Tseng and Lin proposed SMAP-Mine for perfect mining using sequential mobile patterns.[3]

-Taking into consideration different time intervals to find various behaviors of the users the TMAP concept was proposed by Lee at al.[1]

-To predict the future locations of the mobile user the concept of Hybrid prediction Algorithm was put forth realizing the difficulties of object movements.[3]

- Various clustering techniques were also introduced which included the K_means Algorithm. Then the Cluster Affinity search technique (CAST) was put forth by Ben-Dor and Yakhini..[1]

-For segmentation of time intervals the concept of genetic Algorithm was used and it was introduced by Holland. The quality of the chromosomes is considered. [1]

-Monreale et al proposed a prediction model named Where Next which utilizes the trajectory patterns predicting the next location of objects.

LBS Alignment Algorithm

For clustering of mobile transactions sequences of different users ,we need to find similarity between the users.For this purpose we use LBS alignment algorithm.

In LBS Alignment Algorithm, parameters such as user's time, location, services, age and profession are taken into consideration.

When users move within the mobile network, their locations, service requests and time at which user is requesting for the service are stored in a centralized mobile transaction database. Some users may request services in morning, afternoon or at any time and hence it becomes necessary to consider similarity regarding time along with location and service .As per the human tendency, thinking of a person belonging to certain age group is almost similar and similarly, people with same profession may need same type of services.Hence, to increase accuracy of prediction we have consider two additional parameters named age and profession.

The concept of dynamic programming is used in this algorithm. According to this concept, the matrix is generated with column and rows in dataset according to the No. of user's transactions. Here similarity score is computed for two users. The algorithm is as follows for computing similarity score between two users.

LBS alignment is based on the consideration of two mobile transactions and computing their similarity matrix on the basis of location, service, profession, age and timestamps. The entry M(i,j) in the matrix M represents the similarity of mobile transaction sequences i and j in the database.

The base similarity score is set as 0.5. Two mobile transactions can be aligned if their locations are same then we generate Time Penalty(TP) and Service Reward(SR). Otherwise, a location penalty is generated to decrease their similarity score.

```
Input : Two mobile transaction sequence u1 and u2.
Output : The similarity score between u1 and u2.
p ← 0.5/(u1.length + u2.length)   /* p is the location penalty */
M_{0,0} ← 0.5
M_{i,0} ← M_{i-1,0} − p   ∀ i = {1, 2, ..., u1.length}
M_{0,j} ← M_{0,j-1} − p   ∀ j = {1, 2, ..., u2.length}
For i ← 1 to u1.length
   For j ← 1 to u2.length

      If u1.prof == u2.prof
         profs = 0.05   /* Static Profession Score */
      End If

      If ag1 == ag2   /* Group no. of u1 and u2 a/c to age */
         ages = 0.1   /* Static Age Score */
      Else
         If | ag1 − ag2 | == 1
            ages = 0.05
         End If
      End If

      If u1.location == u2.location
         TP ← p* | u1.time − u2.time | /TS   /* TimePenlty */
         SR ← p * {(u1.service ∩ u2.service)/(u1.service ∪ u2.service)} /* ServiceReward */
         M_{i,j} ← Max(M_{i-1,j-1} − TP + SR, M_{i-1,j} − p, M_{i,j-1} − p)
      Else
         M_{i,j} ← Max(M_{i-1,j} − p, M_{i,j-1} − p)
      End If
   End For
End For
Return M_(u1.length,u2.length) + profs + ages
```

Fig:1 LBS-Alignment algorithm

The location penalty is defined as 0.5/(|u1.length|+|u2.length|),where |u1.length|and |u2.length| are the lengths of sequences u1 and u2,respectively. When two sequences are totally different, their similarity score is 0.

As we are considering two additional parameters profession and age statically, we need to generate profession score and age score to increase similarity score.

When two mobile transactions are aligned, their time penalty and service reward is measured. TP focuses on their time distance. The farther the time distances between them, the larger their time

penalty. TP that is generated to decrease their similarity score is defined as :
(|u1 time - u2 time|)/TS, where TS indicates the total no. of time slots.TS is calculated from genetic algorithm concept. SR focuses on the similarity of the service requests. The more similar their service requests, the larger their service reward. SR that is generated to increase their similarity score is defined as [|(u1.services) INTERSECT (u2.services)|]/[|(u1.services)UNION(u2.services)|].

Fig.1 shows the procedures of an LBS-Alignment measure.Input data include mobile transaction sequences of two users (line 1).Output data are the similarity between two user's mobile transaction sequences(line 2). The base similarity score is set as 0.5 and location penalty is calculated (line 3). Some parameters are initialized (line 4 to line 6).Programming to calculate M(i,j) (line 7 to line27) is used. In this procedure, profession and age are considered statically. If the profession of two transactions are same then set static profession score (line 9 to line 11).For age we are assuming following age groups: ag1(0-17), ag2(18-27),ag3(28-39),ag4(40-50),ag5(50above). If age groups are same or their difference is one then we set static age score (line 12-line18).If the locations of two transactions are the same (line 19), both the time penalty (line 20) and the service reward (line 21) are calculated to measure the similarity score (line 22). Otherwise, the location penalty is generated to decrease the similarity score (line 24). Finally, u1.length and u2.length matrix is returned and final similarity score of the two mobile transaction sequences is
M(u1.length,u2.length) +profs +ages. (line 28).

Genetic Algorithm

Genetic algorithms(GAs) are search method based on principles of natural selection and genetics. There are mainly following operators in genetic algorithm.1.selection 2.crossover 3.mutation. Selection operator selects the chromosomes in the population to create the next population. The chromosome which fits to the population is selected to reproduce. Crossover operator randomly chooses a locus and subsequences are swapped here and there resulting into children. Mutation operator randomly flips some bits present in the chromosomes. It can occur in some probability at any bit.

In a mobile transaction database, similar mobile behaviors exist under some certain time segments. Hence, it is important to make suitable settings for time segmentation so as to discriminate the characteristics of mobile behaviors under different time segments. A GA-based method to automatically obtain the most suitable time slots with common mobile behaviors is proposed.In this way we get the total no. of time slots.
Algorithm for obtaining TS is as follows :

1.consider the total no of transactions.

Tcnt=total no. of transactions

2. In order to calculate the fittest points we are using the formula

Fitness= Tcnt /24

Here we get the fittest point to select the appropriate time interval.

3. The transactions are swapped and fitted into those time intervals.

4.Finally, the best time intervals of the d/b are found.

5Thus it returns total no of time slots.

Example

Let u1 and u2 be two mobile transaction sequences.

u1={ (4,4,25,2,2), (16,8,7,2,2), (8,12,9,2,2), (1,1,1,2,2) }

u2={ (16,3,8,2,2), (3,11,5,2,2), (20,8,14,2,2), (1,4,28,2,2) }

Total no. of Time Slots is Ts ,computed from genetic algorithm concept. The location penalty is 0.5/(|u1.length|+|u2.length|)=0.0625,where both |u1.length| and |u2.length| are 4. Fig.2 shows the detailed process of u1 and u2 according to algorithm. The M(u1.length,u2.length) value of matrix is 0.2142,profession score is 0.05 and age score is 0.1 . Therefore similarity of u1 and u2 is (0.2142+0.05 + 0.1) = 0.3642. Fig.2 shows the LBS-Alignment result of u1 and u2.Where: l-location, t-time, s-service, a-age,
p-profession.

(l,t,s,a,p)		(4,4,25,2,2)	(16,8,7,2,2)	(8,12,9,2,2)	(1,1,1,2,2)
	0.5	0.4375	0.375	0.3125	0.25
(16,3,8,2,2)	0.4375	0.375	0.3125	0.25	0.1875
(3,11,5,2,2)	0.3750	0.4151	0.3125	0.25	0.1875
(20,8,14,2,2)	0.3125	0.3526	0.2901	0.2276	0.1651
(1,4,28,2,2)	0.2500	0.2901	0.2276	0.1651	0.2142

Fig:2. Similarity Matrix.

Future Scope

In future we will cluster the obtained similarity matrix as well as go for segmenting time intervals inorder to get a proper grouping. Then after applying this technique to real data sets we will mine those groups using Collaborative technique. After mining, we will predict the services that should be opted by the user in future by developing a precise and efficient system for their convenience.

References

[1].Eric Hsueh-Chan Lu, Vincent S. Tseng and Philip S. Yu, "Mining Cluster-Based Temporal Mobile Sequential Patterns in Location-Based Service Environments", *IEEE transactions on knowledge and data engineering*, vol. 23, no. 6, june 2011

[2].Ms.V.C.Belokar and Mr. P.S. Kulkarni, "Mobile Sequential Pattern Mining in Location based Service Environment", *International Journal of Electronics and Computer Science Engineering*,ISSN 2277-1956/V1N3-1328-1339

[3].Eric Hsueh-Chan Lu, Wang-Chien Lee and Vincent S. Tseng, "A Framework for Personal Mobile Commerce Pattern Mining and Prediction",*IEEE transactions on knowledge and data engineering*, vol.24,2012

[4].Ming-Syan Chen,Jong Soo Park and Philip S. Yu, " Efficient Data Mining for Path Traversal Patterns" *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, No. 2, March/April 1998

[5].Sergio Mascetti Claudio Bettini DICo, "A Comparison of Spatial Generalization Algorithms for LBS Privacy Preservation", *International Journal of Engineering Research & Technology (IJERT)* vol. 2 Issue 1, January-2013 ISSN: 2278-0181

[6].E. Rama Kalaivani1, E. Ramesh Marivendan, "Improvement in Performance of Mobile User Behavior Prediction Using Mining Techniques", *International Journal of Engineering Research & Technology (IJERT)*,vol. 2 Issue 1, January-2013 ISSN: 2278-018