Information Science and Technology
ISSN: 0976-917X & ISSN: 0976-9188, Volume 3, Issue 1, 2014, pp.-051-053.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BIJ0000002

AN EFFICIENT MINING APPROACH FOR TRACKING MOVING OBJECTS

KRISHNAMOORTHY K.[1]* AND GNANAJEYARAMAN R.[2]

[1]Sudharsan Engineering College, Pudukkottai- 622 501, TN, India.
[2]SBM College of Engineering, Dindigul- 624 005, TN, India.
*Corresponding Author: Email- kkr_510@yahoo.co.in

Received: March 13, 2013; Accepted: March 27, 2014

Abstract- Data mining is defined as to extract or mining a large amount of data and data mining techniques are used in many research areas, techniques, genetics and marketing. Data warehousing is defined as a set of databases or some other information. Clustering is a data mining technique used to place data elements into related groups without advance knowledge of the group definitions, popular clustering techniques include k-means clustering and expectation maximation clustering. The existing object tracking applications focus on finding the moving patterns of a single object or all objects. In contrast, propose a distributed mining algorithm that identifies a group of objects with similar movement patterns. This information is important in some biological research domains, such as the study of animals' social behavior and wildlife migration. The proposed algorithm comprises a local mining phase and a cluster ensembling phase. In the local mining phase, the algorithm finds movement patterns based on local trajectories. Then, based on the derived patterns, we propose a new similarity measure to compute the similarity of moving objects and identify the local group relationships. To address the energy conservation issue in resource-constrained environments, the algorithm only transmits the local grouping results to the sink node for further ensembling. In the cluster ensembling phase, our algorithm combines the local grouping results to derive the group relationships from a global view. We further leverage the mining results to track moving objects efficiently. The results of experiments show that the proposed mining algorithm achieves good grouping quality, and the mining technique helps reduce the energy consumption by reducing the amount of data to be transmitted.

Keywords- Data mining, Data warehousing, Clustering, Genetics Algorithm.

Citation: Krishnamoorthy K. and Gnanajeyaraman R. (2014) An Efficient Mining Approach for Tracking Moving Objects, Information Science and Technology, ISSN: 0976-917X & ISSN: 0976-9188, Volume 3, Issue 1, pp.-051-053.

Introduction

The extraction of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer business questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

Most companies already collect and refine massive quantities of data. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources, and can be integrated with new products and systems as they are brought on-line. When implemented on high performance client/server or parallel processing computers, data mining tools can analyze massive databases to deliver answers to questions such as, "Which clients are most likely to respond to my next promotional mailing, and why?"

Movement Pattern Mining

The trajectory data sets of moving objects are often modeled as sequential patterns for use in data mining. It defined the sequential pattern mining problem and proposed an Apriori-like algorithm to mine which is an FP-growth-based algorithm that addresses the sequential pattern mining problem by considering the pattern-projection method. For handling the uncertainty in trajectories of mobile objects and developed a new match measure and proposed Trajectory pattern to mine sequential patterns from imprecise trajectories [4]. A number of research works have been elaborated upon mining traversal patterns for various applications and proposed the FS and SS algorithms for mining path traversal patterns in a Web environment. It proposed an incremental algorithm to mine user moving patterns for data allocation in a mobile computing system. sequential patterns or path traversal patterns do not provide sufficient information for location prediction or clustering. The reasons are as follows: First, for sequential pattern mining or path traversal pattern mining extract frequent patterns of all objects, meaningful movement characteristics of individual objects may be ignored. Second, a sequential pattern or a traversal pattern carries no time information between consecutive items, so they cannot provide accurate in-formation for location prediction when time is concerned. Third, sequential patterns are not full representative to individual trajectories because a sequential pattern does not contain the information about the number of times it occurs in each individual trajectory. The TMP-Mine algorithm for discovering the temporal movement patterns of objects. Apriori-like or FP-growth-based algorithms suffer from computing efficiency or memory problems, which make them unsuitable for use in resource-constrained environments.

Trajectories Clustering

Clustering based on objects' movement behavior has attracted more attention. Moving Micro clusters (MMC) to discover and maintain a cluster of moving objects online. It proposed trajectory clustering to discover popular movement paths. Clustering similar trajectory sequences to discover group relationships is closely related to our problem and it transform the location sequences into a transaction-like data on users and based on which to obtain a valid group [12].

The proposed AGP and VG-growth algorithms are Apriori-like or FP -growth-based algorithms that suffer from high computing cost and memory demand. Apply a density-based clustering algorithm to the trajectory clustering problem based on the average Euclidean distance of two trajectories [8]. The discover group information based on the proportion of the time a group of users stay close together or the average Euclidean distance of the entire trajectories may not reveal the local group relationships, which are required for many applications.

Apriori Algorithm

Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation). Other algorithms are designed for finding association rules in data having no transactions (Winepi and Minepi), or having no timestamps (DNA sequencing). The purpose of the Apriori Algorithm is to find associations between different sets of data. It is sometimes referred to as "Market Basket Analysis" [10]. Each set of data has a number of items and is called a transaction. The output of Apriori is sets of rules that tell us how often items are contained in sets of data

Algorithm 1: Apriori

1. L1 = flarge 1-itemsetsg;
2. C1 = database D;
3. for (k = 2; Lk1 =6 ;; k++) do begin
4. Ck = apriori-gen(Lk1); // New candidates
5. Ck = ;;
6. forall entries t 2 Ck1 do begin
7. / determine candidate itemsets in Ck contained

// in the transaction with identier t.TID

Ct = fc 2 Ck j (c c[k]) 2 t:set-of-itemsets ^

(c c[k1]) 2 t.set-of-itemsetsg;

8. forall candidates c 2 Ct do
9. c:count++;
10. if (Ct =6 ;) then Ck += < t:TID; Ct >;
11. end
12. Lk = fc 2 Ck j c:count □ minsupg
13. end
14. Answer =SkLk;

Prediction Suffix Tree

The PST building algorithm extracts significant patterns from a data set, prunes unnecessary nodes during tree construction, and then generates a PST. Each node in the tree is labeled by a string, which represents a significant pattern with occurrence probability above the minimal threshold. Each node carries the conditional empirical probabilities and the maximal length of s is specified by maximum. It represents a significant pattern [14]. The PST algorithm has an excellent capacity for extracting structural information from sequences. Its low complexity and it more attractive to be used in streaming or resource-constrained environments. Compared with algorithms that mine all accurate frequent patterns, the compact tree structure and the controllable size of a PST are particularly useful in resource-constrained environments.

Cluster Ensembling

CH collects location data locally and generates group information with the proposed GMPMine algorithm. The objects may not pass through all the clusters and the group relationships of objects may vary in different areas, the local grouping results may be inconsistent, The objects scattered in grassland is hardly identified as a group a group of objects move across the margin of a sensor cluster. The group relationship is difficult to determine [8]. The CE algorithm to combine multiple local grouping results. The algorithm solves the inconsistency problem and improves the grouping quality. The algorithm measures the similarity of each pair of objects to construct a similarity matrix based on the local grouping results The sink node uses the CE algorithm to combine the local grouping results [1]. It then assigns a global group ID to each group and sends the group information to the CHs for subsequent collection of the location data.

System Analysis and Design

Tracking an single or all objects. It contains a large amount of location data. The results of the object are inconsistent and suitable pattern for our application is not adequate. To find the similarity measure and group the object. The energy consumption is high and time loss.

Problem Analysis

In the proposed system tracking all the moving objects and determine similarity by clustering ensembling algorithm. We propose a technique that identifies a group of moving objects between the client in a network. It reduces energy consumption. A distributed mining frame work to discover group relationships as well as group movement patterns. A new pairwise measure based on pattern similarity to compute the similarity of moving objects. The discovered information to track moving objects efficiently. The detail of the drifting object can be stored and display in server as chart view.

The GMP Mine algorithm is used to identifies the group of objects and determine the movement patterns. The GMP Mine Algorithm generates The grouping results and associate grouping movement pattern.

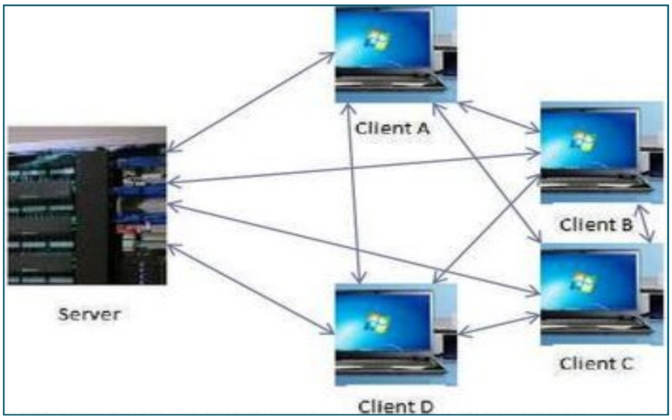

Fig. 1-

Algorithm 2: Gmp Mine

Input:s'={si|o<=i<n}, simmin, pmin, a, ymin, lmax, Σ

Output: G.M, GT

0.G=ϕ

1. m=0
2. *building a pst for each object and puring noise*/
3. for each si in s'
4. ti=pst- build (sin, pmin, a, ymin, f, I max, Σ)
5. for 0<=i<n-1
6. for i+1<=j<n
7. if simple (tin tj)>sim min then
8. add_edge (I, j) to graph (V, E)
9. (G,M)=HCS (graph (V,E)}||G={g,10<=i<=m}
10. for 0<=i<m
11. s'={sj|oi Σ gi, o<=j<n}
12. t'={ti|oj tgi o<=j<n}
13. gti=argmax Σpj(s2) where tiΣT'
14. return G={gj|o<=|i<m}, GT={GT|o<|i<m}

Descriptions

The GMPMine algorithm is comprised of four steps. First, we extract the movement patterns of each object from the location sequence. Second, we construct a similarity graph in which similar objects are connected by an edge. Third, we extract highly connected components to derive the group information. Fourth, we construct a group PST for each group in order to conserve the memory space.

Cluster Ensembling

The Cluster Ensembling algorithm is used to combine the multiple local grouping results. The algorithm solves the inconsistence problems. The algorithm improves the grouping quality. The group relationships of objects may vary in different areas.

Algorithm 3: Cluster Ensembling

Input:o={00,01…….0n},c={G|0<=i<k},d={δ,|0<=i<d}

Output:Gδ

0.init Sum[]

1. init sm[][]
2. idx=0
3. max=0
4. for 0<=i<n-1
5. for i+1<=j<n
6. sm[i,j]=get si(c)
7. for 0<=i<d
8. graph(V,E)=convert 2 graph(sm,δ)
9. 9.Gδ=HCS(graph(v,e))
10. sum[i]=Σnm(gi,gδ)
11. if sum[i]>max then
12. max=sum[i]
13. idx=i
14. gδ=gδn
15. return gδ

Descriptions

The algorithm measures the similarity of each pair of objects to construct a similarity matrix based on the local grouping results.

System Design

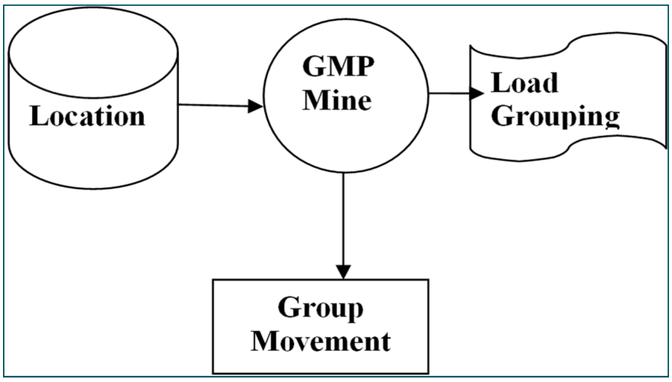

Fig. 2-

System Architecture

In our system architecture the user gives the input to the GMP mine. The GMp mine search in the location sequence dataset. if the given data is present in the location sequence dataset, it moves the data's to the group movement pattern search In the group movement pattern search, the pattern matching operation is performed, then it sends the data to clustering approach.in the clustering approach the data's are clustered, finally it sends to cluster ensembling in that only one best result is produced from the group of results.

Conclusion

The exploitation of group movements to discover the information about groups of moving objects. In contrast to the centralized mining technique, mine the group information in a distributed manner. The novel mining algorithm, which consists of a local GMPMine algorithm and a CE algorithm, to discover group information. An algorithm mines object movement patterns as well as group information and the estimated group dispersion radius. The distributed clustering approach to heterogeneous and distributed sequential data sets, such as web logs or gene sequence. The contribution of our approach is threefold: it reduces energy consumption by allowing CHs to avoid sending the prediction-hit locations, because the locations can be recovered by the sink via the same prediction model; it leverages group information in data aggregation to eliminate redundant update traffic. Mining technique achieves good grouping quality. Data aggregation significantly reduces energy consumption in terms of the transmission cost, especially in the case where moving objects have distinct group relationships.

Future Enhancement

The characteristics of group movements to discover the information about groups of moving objects in an OTSN. To evaluate the efficiency of the proposed OTSN, we investigate the impacts of the network structure, accuracy error bound (EB), and SG radius (R) as GDR varies. To the best of our knowledge, group relationships in location data aggregation. We only compare our design with a conventional update based OTSN with a naive data aggregation techniques. We assume that there are five groups of objects, each of which contains five objects, walking in a mesh network composed of 65,536 sensors.

Conflicts of Interest: None declared.

References

[1] Abe S. & Lan M.S. (2005) *IEEE Trans. ON Syst. Man Cybern,* 25 (1), 119-129.

[2] Azzag H., Monmarche N., Slimane M. & Venturini G. (2003) *IEEE Congress on Evolutionary Computation*, 4, 2642-2647.

[3] Belacel N., Hansen P. & Mladenovic N. (2002) *Pattern Recognition*, 35(10), 2193-2200.

[4] Baraldi A. & Blonda P. (1999) *IEEE Transactions on Systems, Man and Cybernetics*, 29(6), 778-785.

[5] Ke B.R., Chen M.C. & Lin C.L. (2009) *IEEE Transactions on Intelligent Transportation Systems*, 10(2), 226-235.

[6] Jin B. & Zhang L. (2010) *Asia-Pacific Conference on Wearable Computing Systems,* 311-314.

[7] Li B., Shi L. & Liu J. (2010) *Seventh International Conference on Fuzzy Systems and Knowledge Discovery,* 6, 2905-2908.

[8] Chaharsooghi S.K. & Meimand Kermani A.H. (2008) *IEEE World Congress on Computational Intelligence,* 1195-1202.

[9] Coulibaly Y., Latiff M.S.A. & Selamat A. (2009) *Second International Conference on Communication Theory, Reliability, and Quality of Service,* 108-112.

[10] Juang C.F. & Chang P.H. (2010) *IEEE Transactions on Fuzzy Systems*, 18(1), 138-149.

[11] Dorigo M., Gambardella L.M. (1997) *IEEE Transactions on Evolutionary Computation*, 1(1), 53-66.

[12] Dimitrov S., Sinanovic S. & Haas H. (2011) *IEEE International Conference on Communications*, 1-5.

[13] Dorigo M. (1992) *Optimization, Learning, and Natural Algorithms*, Ph. D. Thesis, Italy.

[14] Bae D.H., Baek J.H., Oh H.K., Song J.W. & Kim S.W. (2009) *IEEE International Conference on Network Infrastructure and Digital Content,* 803-807.

[15] Ganji M.F. & Abadeh M.S. (2010) *IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications,* 573-581.

[16] Wang G., Gong W., DeRenzi B. & Kastner R. (2007) *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* 26(6), 1010-1029.

[17] Handl J. & Meyer B. (2002) *Parallel Problem Solving from Nature,* 913-923.

[18] Handl J., Knowles J., and Dorigo M. (2003) *Frontiers in Artificial intelligence and Applications,* 104, 204-213.

[19] Handl J., Knowles J., Dorigo M. (2003) *Self Organising Applications, Challenges and Trends,* 2977, 90-104.

[20] Holden N. & Freitas A.A. (2008) *Journal of Artificial evolution and Applications*, 2.

[21] Hassini N. & Zouairi S. (2011) *Electronics, Communications and Photonics Conference,* 1-6.

[22] Imani M., Pakizeh E., Pedram M.M. & Arabnia H.R. (2010) 9*th IEEE International Conference on Cognitive Informatics*, 186-193.

[23] Yang J.Q., Yang J.G. & Chen G.L. (2009) *Second International Symposium on Computational Intelligence and Design*, 1, 49-51.

[24] Jiang J., Chen X., Zang M., Wang Z. & Tan Z. (2010) *In Future Computer and Communication (ICFCC), 2nd International Conference on*, 2, 2-45.