**Study Paper of Spatial Mining for Location Recommender System**

1st Sunitha Geedikanti , 2nd Javed A Khan

1st Department of Computer Science & Engineering

2nd Department of Computer Science & Engineering

1st Indur Institue of Engineering & Technology

2nd Indur Institue of Engineering & Technology

Telengana [INDIA]

**Abstract - This Article include the Basic introduction part of Data mining along with its various domain . We are make study of one type of Data mining that is spatial mining for the location recommender system. A various Data mining concept are already existing for the spatial mining for the so many filed and application . In this article we are include the study of LRS .**

***Keyword – Data mining , Spatial mining , KDD***

**1.Introduction –**

Data mining [1] is the powerful new technology to analysis of data and the use of software techniques for finding patterns and regularities in sets of data. Data mining is also help companies to extraction of hidden predictive information from their large database and finds valuable information hidden in large volumes of data that’s why Data mining as also known as “knowledge discovery in database(KDD)[1]”.Data mining tools give answer business questions easily that conventionally method is has to time consuming to resolve. All the Data mining tools also analyze massive databases to deliver answer to question like “which customer are most likely to respond to my upcoming promotional activity & why?”.The data mining include the several field like database system, data statistics and information theory, machine learning etc .Data mining is very demanding topic for the resarcher.

**1.2Application of Data Mining**- There are so many application of data mining , a list of application where we use the concept of data mining shown in the table

|  |  |  |
| --- | --- | --- |
| S.No | Data mining Application | Description of application |
| 1 | Medical Discovery | In medical field there are very big scope for applications relate to data mining like as diesis diagnose, health care, information of patient profiling and patient history generation etc . |
| 2 | Market basket analysis | Data mining is also used in the concept of market basket analysis(MBA).When the customer want to buying some products then MBA technique helps us finding all the items that selected by the customer put those item it in to the their his or her shopping buckets. This technique also used retailer to identify that customer intension and buying pattern. MBA technique mainly used in the profits calculation of the any business, promoting all type of business and also helps to selecting the items . |
| 3 | Education System | The data mining is used in the education system in the whole world .Now a days number of universities are established in the whole world, each and every day a millennium of students are enrolls across the country. |
| 4 | Intelligence Agencies | The intelligent agencies collect information and analyze information to investigate using the Data mining process and track the terrorist activities. it is the difficult task of analyzing large volume of data involve in criminal and terrorist activities. |
| 5 | Engineering Applications prediction | The prediction engineering faces problems like the cost estimation, engineering’s design and concept that involves decisions where some of the parameters, actions, selected components, etc are finalize . This selection is based on prior data, information and knowledge then numerous models and algorithms have been developed for autonomous predictions using data mining . |

**Table 1.1Data Mining Application**

**1.3 Data Classification as per the type of data source -**

* **Spatial data**- in the Spatial databases collection we are store the usual data information , some of the geographical related information in terms of spatial data say for the example map , information about the location these location may be local or may be global. So this type of database is show the some of the new issue for the data mining algorithms.
* **Multimedia data**- Multimedia databases include images, audio, video and text media etc. These type of multimedia data is saved on extended Object relational or some time store as in the OOD , its another format is File system .
* **A Time series database Storage** - In this Time-series databases storage system include the time related data only say for the example stock market related data information or logged/ login activities . database usually have a large in data size, high dimensionality and necessary to update continuously .
* **Text data-** The Text databases is the kind of databases that include the word descriptions of the objects. This description of word are include all the product specification , product related error , some of the technical bug of particular product in other word we can say that in the text data the word description is not a analysis of simple keyword but this is a analysis if all products related information.
* **World wide data –** This is a representation of inter connected data or document in the world wide server .this a process of organizing the inter connected world wide data this is include the the all kinds of data like audio data , txt data of information kind of raw data , and applications. As all know the some type world wide of services we are used these daily basis these services are Google, yahoo, Gmail, worldwide, on line information services etc. The World Wide Web is three major components-
* **Content of the Web**- Web content mining is a methodology of concentrating data from writings, pictures and different substance.
* **Web structure –**The structure of the web is include all the hyperlinks of data those data or information that are available in the internet , this is also shoe the relationships between data information .
* Usage of the web- This is shoe the utilization of wed data. this is also include the information about the how and when the resources are accessed or used . In other words how to utilize sites.

**1.3 Knowledge Discovery-** Knowledge discovery KDD is the last stage of the process of sentiment mining . with the help of this KDD we easily find the opinion from the of the reviewer /people with respect to any particular occurrence of the data and information , and generate statistical graphs and charts . The knowledge gathered from the web or internet when shown in graphs chart would help the individuals in making decision and show the polarity of the sentiments of the individuals . After all these stages are completed, the process of sentiment mining is successfully executed.

**2 Introduction of Spatial Mining**

Spatial data mining [5] is a part of data mining. It is the process of discovery interesting & previously unknown, but potentially useful patterns of information in large database. spatial data mining are also used in different organizations such as NASA means National Aeronautics and space administration, USDOT means United States Department of Transportation ,NCI means National Cancer Institute, NIMA means National Imagery and Mapping Agency etc. These organizations used spatial data mining for extracting information from geospatial data and also used different application domains like as earth science, public transportation, nation safety etc.

**Example**: Earth observation satellites, which are systematically mapping the entire surface of the earth, collect about one Tera byte of data every day. The main different between data mining & spatial data mining is that in spatial data mining is not only used non spatial attributes, but also spatial attributes while data mining is used non spatial data. Spatial data mining explain a locations, these locations are part of graphic primitives that are usually either points ,lines ,polygons or pixels.**Example:** consider a particular square- It center specifies its location ,it shape is a square . it size define one side of length of square. spatial data are generally multidimensional & auto correlated Non spatial data that relate to specific precisely define location .The data are often statistical but may be text ,images or multimedia. These are linked in the GIS to spatial data that define the location.

**2.1 Usages of Spatial Mining**

Table 2.1 Shows the usages of Spatial Mining in the various field

|  |  |  |
| --- | --- | --- |
| S.No | Usages of Spatial Mining | Description |
| 1. | Remote sensing | Spatial Mining is also used for Remote Sensing. |
| 2. | Computer Cartography | Spatial Mining is also applicable for the Computer Cartography in the various aspect. |
| 3. | Geographical information system (GIS) | GIS application are the domain of Spatial Mining . |

**Table 2.1 Usages of Spatial Mining**

**2.2Advantages of Spatial Mining**

There are many advantages of spatial mining-

**1. Massive data** - We know that massive data not used in many algorithm because it implemented for difficulty and excessive calculating amount, therefore spatial data mining is to create a new computing strategy and develop new efficient algorithms to overcome the technical difficulties caused by massive data

**2**.**Non-linear relationship between the spatial attribute** - It is an important symbol of the complexity of space systems, reflects the complex mechanisms of the system internal function. It is one of the advantage of spatial data mining.

**3**. **Scale characteristic of spatial data -** We know that spatial data is not the same at different observing levels. Scale characteristic of spatial data can be used to explore gradual change law of the characteristic in the process of generalization and refinement of information.

**4.** **Spatial dimension increasing** – In spatial mining spatial dimension increases , as in the field of remote sensing, development of sensor technology ,etc.

**5. The ambiguity of spatial information-** Spatial mining exists all types Ambiguity of spatial information , such as the ambiguity of spatial location, the ambiguity of spatial correlation, the attribute values of fuzzy, etc.

**6. The lack of spatial data-** The lack of spatial data is used in spatial mining ,due to some irresistible external force so that data cannot be lost and also recover lost data.

**2.3 Limitation of Spatial Mining [6]**

I-The efficiency of spatial data mining algorithms is not high and the discovery patterns of spatial data mining algorithms are not refined. When the uncertainty appeared in process of spatial data mining, the possibility of error and the dimension of the problem to be solved are all large then not only increases the algorithm's search space also increases the possibility of blind search.

II- Spatial mining is no generally accepted standardized spatial data mining query language.

III- Interaction of knowledge discovery system of spatial data mining interaction is not strong, Therefore the user can't well control the spatial data mining process and also face difficulty to fully and effectively utilize domain experts’ knowledge in the knowledge discovery process.

IV-We know that the methods and tasks spatial data mining is single and essentially aimed at a specific problem, so the knowledge can be found is limited.

V-The integration spatial data mining with other systems is not enough and ignoring the role of GIS in spatial knowledge discovery.

VI-Expensive spatial processing operations.

**3 Literature**

This section of my dissertation report are include the literature of various research paper and make a proper compression analysis of location acquaint related information . A wide array of techniques [9] are used in this paper this paper is capable of producing recommendations using non-spatial ratings for non-spatial items . As per the M.H Park et.al [10] used the contextual attribute concept for the location recommender , in this research paper author are make statistical recommendation models are introduce the commercial applications make cursory use of location when proposing interesting items to users and displays a “local favorites” list containing popular movies for a user’s given city. The City Voyager system [11] mines a user’s personal GPS trajectory data to determine her preferred shopping sites, and provides recommendation based on where the system predicts the user is likely to go in the future. LARS, conversely, does not attempt to predict future user movement, as it produces recommendations influenced by user and/or item locations embedded in community ratings. The spatial activity recommendation system is proposed by V.W Zheng et.al [12] mines GPS trajectory data with embedded user-provided tags in order to detect interesting activities located in a city (e.g., art exhibits and dining near downtown). It uses this data to answer two query types: (a) given an activity type, return where in the city this activity is happening, and (b) given an explicit spatial region, provide the activities available in this region. This is a vastly different problem than we study in this paper.

LARS does not mine activities from GPS data for use as suggestions for a given spatial region. Rather, we apply LARS to a more traditional recommendation problem that uses community opinion histories to produce recommendations. Geo-measured friend-based collaborative filtering[13] produces recommendations by using only ratings that are from a uerying user’s social-network friends that live in the same city. This technique only addresses user location embedded in ratings. LARS, on the other hand, addresses three possible types of location-based ratings. More importantly, LARS is a complete system (not just a recommendation technique) that employs efficiency and scalability techniques (e.g., merging, splitting, early query termination) necessary for deployment in actual large-scale applications. Author P.Venetis and H Gonzalez [14] proposed the problem of hyper-local place ranking. Given a user location and query string, hyper-local ranking provides a list of top-k points of interest influenced by previously logged directional queries. Hyper-local ranking does not personalize answers to the querying user, i.e., two users issuing the same search term from the same location will receive exactly the same ranked answer set.

**4 Conclusion –** This article include the introduction part of Data mining Data mining application . we are also collect the some information of spatial mining . In the literature analysis part we are include the some existing research article that are proposed some solution , algorithm, comparing analysis review for LR .

**Reference**

[1] http://www.innovativejournal.in/index.php/ajcsit

[2] http://www.ijaiem.org/volume3issue5/IJAIEM-2014-05-31-124.pdf

[3]http://www.ijarcce.com/upload/2014/july/IJARCCE2M%20s%20angulakshmi%20An%20Analysis%20on.pdf

[4] http://inpressco.com/category/ijcet

[5]http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.215.3645&rep=rep1&type=pdf

[6]58.205.208.126/meeting/UploadPapers/1281513282375.pdf

[7]www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/.../isprsarchive.

[8]http://www.geospatialworldforum.org/2012/gwf\_PDF/S%20Jyothi.pdf

[9]G. Adomavicius and A. Tuzhilin, “Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions,” TKDE, vol. 17, no. 6, pp. 734–749, 2005.

[10] M.-H. Park et al, “Location-Based Recommendation System Using Bayesian User’s Preference Model in Mobile Devices,” in UIC, 2007.

[11] Y. Takeuchi and M. Sugimoto, “An Outdoor Recommendation System based on User Location History,” in UIC, 2006.

[12]V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, “Collaborative Location and Activity Recommendations with GPS History Data,” in WWW, 2010.

[13] M. Ye, P. Yin, and W.-C. Lee, “Location Recommendation for ocationbased Social Networks,” in ACM SIGSPATIAL GIS, 2010.

[14]P. Venetis, H. Gonzalez, C. S. Jensen, and A. Y. Halevy, “Hyper-Local,Directions-Based Ranking of Places,” PVLDB, vol. 4, no. 5, pp. 290–301, 2011.