International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 30 • 2017

Spatial Co-location Pattern Mining using Delaunay Triangulation

G. Kiran Kumar[a] Ilaiah Kavati[a] and K. Srinivasa Rao[a]

[a]Department of CSE, MLR Institute of Technology, Hyderabad-43

E-mail: ganipalli.kiran@gmail.com, kavati089@gmail.com, hodit@mlrinstitutions.ac.in

Abstract: Spatial data mining, or knowledge discovery in spatial database, is the extraction of hidden information, spatial relations, or various patterns not unequivocally put away in spatial databases. Spatial data mining is the way of retrieving interesting qualities and patterns that may implicitly exist in spatial database. Co-location pattern discovery is the way of finding the subsets of features that are regularly found together in a same location. Spatial co-location patterns relate the co-presence of non-spatial features in a spatial neighbourhood. The Previous strategies for mining co-location patterns, converts neighbourhoods of feature occurrences to item sets and applies mining methods for transactional data to find the patterns, merges the discovery of spatial neighbourhoods with the mining procedure. It is an expansion of a spatial join approach that works on numerous information sources and counts long pattern occasions. Existing systems on finding co-location patterns depends on participation ratio, participation index. In this paper we address the issue of mining co-location patterns with a novel technique called Delaunay triangulation. Delaunay triangulation represents the closest neighbourhood structure of the features exactly which is a major concern in finding the co-location patterns. The results shows that, this approach achieves good performance compared to previous methodologies.

Keywords: *Spatial Data Mining, Association Rules, Co-location rules, Participation Index, Participation ratio, Delaunay triangulation.*

1. INTRODUCTION

The enormous development of spatial data and far reaching utilization of spatial databases accentuate the requirement for the computerized discovery of spatial learning. Spatial data mining [7, 8] is the way toward finding implicit patterns and already obscure, however conceivably helpful patterns from spatial databases. Removing fascinating patterns from spatial datasets is more troublesome than removing the corresponding patterns from customary numeric and clear cut data due to the complexity of spatial data sorts, spatial connections what's more, spatial autocorrelation [10]. A spatial co-location pattern speaks to a subset of spatial components whose cases are much of the time situated in a spatial neighbourhood. For instance, ecologists have found that Nile Crocodiles and Egyptian Plover flying creatures are as often as possible co-located. The co-location manage, *i.e.*, Nile crocodile! Egyptian Plover, predicts the nearness of Egyptian Plover feathered creatures in zones with Nile Crocodiles. Spatial Co-location patterns may yield critical bits of knowledge for some applications including Earth science, general wellbeing, science, transportation, and so on.

Morimoto et al. proposes a space partitioning and non-overlap grouping method for finding neighbouring features for a frequent neighbouring feature set [28]. However, there are number of outliers across the spatial space and false results. In another approach, a join-based algorithm was proposed for co-location pattern mining. With the ever growing co-location patterns and their instance, the join-based approach is considered very computationally expensive. Reducing the number of expensive spatial join operations in finding co-location instances is very important which is proposed in a partial join approach. Join less algorithms have been proposed, but methods including the join less algorithm do not discover both single self-co-location patterns and complex self-co-location patterns [29].

In Tran Van Cahn et al. proposed a novel constraint neighbourhood based method to find co-location patterns. This approach is able to find different co-location patterns such as star and clique, including single and complex self co-locations. Based on the constraint neighbourhood idea, this method neither needs to perform spatial or instance neither joins nor checks for cliques to find co-location instances. [30].

Rest of the paper is organized as follows. Section 2 gives an overview of the basic concepts related to data mining, association rules, etc. In Section 3, we presents the proposed approach for co-location pattern mining. Section 4 presents the experimental evaluation. The conclusion and future work are discussed in Section 5.

2. THEORETICAL BACKGROUND

2.1. Data Mining

Extraction of knowledge from a data set in a human justifiable shape is the objective of data mining. The essential source to perform data mining [1, 2] is where the data lives i.e. database. It is a gigantic collection of truths identified with some space in the outside world in a composed way. One of the essential points or goals of data mining is to break down regularities that empower us for a detailed investigation of the area clarified by the database. Data mining is otherwise called Knowledge Discovery. It leaves plentiful scope to discover the procedure and patterns from an expansive database. It recognizes the substantial, conceivably valuable, reasonable patterns and connections in data. It discovers the new patterns from the huge database [3, 4]. It is advanced as another interdisciplinary field of computer science.

2.2. Spatial Databases

Prior in the research of data mining, just non-spatial data was concentrated. This brought about the immense accomplishment in finance, GIS, business, protection, and so on. Later the concentration of research has been moved to spatial data from non-spatial data. The imperative contrast between non-spatial data and spatial data is that spatial data incorporates the constraints like scales, coordinates and have pictorial data. Attributable to a few constraints, the covert learning inserted in the spatial database can't be recovered utilizing conventional database administration frameworks. It needs propelled administration frameworks to handle with. Utilization of spatial data upgrades the efficiency of data mining. This was recognized by researchers, natural specialists, chairmen and business experts. Spatial data can't be dealt with in an indistinguishable route from it was with alternate types of data. A spatial database administration framework is utilized for this reason.

The essential innovation for SDBMS is Geographic Information System. The contrast between the GIS and SDBMS is that the GIS has a rich arrangement of operations more than few questions and layers, while SDBMS gives less complex operations on an arrangement of items and set of layers. Web gives a spatially touchy web crawler. Questions, for example, “Discover all Udipi Hotels in Hyderabad” can be replied by such web crawlers. Spatial advancements can be utilized to decide the location of the cell phones which helps in giving location-based administrations. The uses of spatial data are those which include choices. For instance, In Geo-Marketing, a store can set up its exchange zone, i.e. the spatial limit of its clients. At that point the profile of those clients is examined on the premise of both their properties and the properties identified with the region where they live.

In Geography, the fundamental premise of data examination is conventional insights and multi-dimensional investigation which doesn't consider spatial data [5]. However the geographic data suggests that the perceptions that are close by in the space tend to have comparable (or correlated) quality qualities. This comprises the principal of particular logical territory that spotlights on between reliance of close by perceptions, dissimilar to conventional measurements. This is known as "spatial insights". The expansion of spatial insights is Multi-dimensional diagnostic methods [6]. The review demonstrates that spatial measurements is an indispensable piece of spatial data mining as it gives data driven investigation. Some of these techniques are presently actualized in operational GIS or examination instrument.

2.3. Spatial Data Mining

The additional components that recognize spatial data from different types of data are spatial co-ordinates, topological separation, and bearing data. By incorporation of many elements, inquiry dialect has become excessively complicated. The procedure to discover learning is not all that simple as prior at this point. The reason is "spatial measurements". In contrast to the mining in social databases, spatial data mining calculations need to consider the items that are close by with a specific end goal to extricate helpful information as there is impact of one question on the neighbouring article. Diverse methodologies have been created for learning discovery from spatial data, for example, Spatial Classification, Spatial association govern mining, Spatial Clustering [7].

2.3.1. Spatial classification

The fundamental target of arrangement is to dole out a question a class from a given arrangement of classes in view of the quality estimations of the protest. In spatial characterization the characteristic benefits of neighbouring items may likewise be pertinent for the participation of articles and in this manner must be considered too. Gathering of data things according to their quality qualities into classes is known as Classification [8]. This is recognized as managed order; unsupervised characterization is called as bunching. Supervised arrangement needs a preparation dataset to prepare or configure the order display, an approval dataset to approve (or advance) the configuration, and a test dataset to test or assess the execution of the prepared model [9,10]. Order strategies incorporate, for instance, choice trees, straight discriminant work, support vector machines (SVM), simulated neural systems, most extreme probability estimation, closest neighbour techniques and case-based thinking.

2.3.2. Spatial association

The first point of association rule mining is to distinguish the regularities between the things in the extensive value-based databases. For the most part the quality is shown by confidence and support means the recurrence of the association rule. Rules that have sensibly huge bolster needs more concentration [11]. Spatial properties and predicates are considered while detailing spatial association rules in spatial database which is like mining of the association rules in value-based or social databases. Spatial co-location pattern mining varies with association rule mining [14] just in the specialized way, where as they are conceptually comparable. A co-location pattern can be acquired by the contribution of a dataset of spatial components and their locations. These patterns speak to subsets of components that are found together more habitually. For instance, certain types of flying creature tend to natural surroundings with certain kind of trees. A location is not an exchange and two elements seldom exist at the very same location. To check the components that co-situate in a similar neighbourhood, a client determined neighbourhood is required. Many measures and calculations have been proposed to mine spatial co-location patterns.

2.3.3. Spatial Clustering

Clustering is deciphered as the errand of collecting the objects of a database into important perceptible subclasses (that is bunches) so that the individuals from a group are as comparative as could be expected under the circumstances though the individuals from various groups contrast however much as could be expected from each other. In Data investigation Cluster examination is oftentimes utilized, which sorts out an arrangement of data things into gatherings (or bunches) so that things in a similar gathering are like each other and not the same as those in different gatherings [6,15]. Bunching strategies can be comprehensively characterized into five gatherings they are Partitioning calculations, Density based grouping, Hierarchical Algorithms, Grid-Based Methods and Model-Based Clustering Methods. Illustration calculations of the above grouping are K-Means, K-medoids, DBSCAN [16] GDBSCAN [17] Chameleon. To consider spatial data in grouping, three sorts of bunching examination are existing, they are spatial grouping, regionalization, and point pattern investigation. This paper fundamentally concentrates on spatial data mining, with a particular reference to spatial co-location pattern mining which is conceptually like association mining yet in fact altogether different.

2.4. Co-location Patterns

Spatial co-location pattern mining is similar to association mining. A spatial association rule is a rule of the form "A $\rightarrow$ B" where A and B are sets of predicates and some of which are spatial ones. In a large database many association relationships may exist but some may occur rarely or may not hold in most cases [18, 19].

Spatial and non-spatial connections at a concept level are discovered in spatial data mining [20]. Spatial items are characterized as consolidated spatial locales or bunched spatial focuses. Spatial or non-spatial connections which contain spatial predicates, for example, adjacent_to, occur_together, near_by, inside, close_to and so on can't be recognized by the above strategies. Question/predicate connections are spoken to by Spatial Association rules containing spatial predicates, For instance, the accompanying rules are spatial association rules "is_ a(*k*, service station) close_to(*k*, highway)". To encourage the determination of the primitives for the spatial data mining, a SQL-like spatial data mining inquiry interface, which is planned in light of a spatial SQL proposed in [21], has been indicated for an exploratory spatial data mining framework model, GeoMiner.

2.5. Co-location rule modelling

Co-location pattern discovery is the way toward finding the subsets of elements that are oftentimes found together in a similar locale. Spatial co-location patterns relate the co-presence of non-spatial elements in a spatial neighbourhood. Attributable to the accompanying reasons Co-location pattern discovery is a vigorous assignment [22]. With conventional association rule mining calculations it is hard to discover co-location patterns since there is no concept of customary "exchange" in the majority of spatial datasets. The occurrences of a spatial element are getting dispersed in a spatial structure, these occasions impart complex spatial neighbourhood connections to other spatial examples. Time complexity to create the table examples of co-location pattern is high. A model based approach for discovering co-location rules is talked about [23]. The Previous strategies for mining co-location patterns, converts neighbourhoods of highlight examples to thing sets and applies mining procedures for value-based data to discover the patterns [15]. Some prior techniques combined mining process with the discovery of spatial neighbourhoods and they are augmentations of a spatial join calculation that works on different information sources and counts long pattern instances [14]. A Co-location rule is of the form: L1 $\rightarrow$ L2(p, cp) where L1 and L2 are co-locations, p and cp refers to prevalence measure and conditional probability[24]. The meaning of neighbour connection R is an info and depends on the semantics of use areas. It might be characterized utilizing diagram hypothesis, for example, connected, nearby or metric relationship utilizing separation measures, for example, earth mover remove, Euclidean separation, city square separation and so forth or a combination. The table case of a co-location is the collection of all its column occurrences. Past chips away at discovering co-location patterns depends on investment file and maximal interest proportion.

3. METHODOLOGY

In this section, we propose a framework for mining co-location patterns. For any co-location mining algorithm major challenge is neighbourhood enumeration and identifying row instances of co-location. Our approach use Delaunay triangulation to model the spatial proximity between neighbourhood features.

The inspiration of utilizing Delaunay triangulation as part of this work is that the Delaunay triangulation have certain unique properties compared to other topological structures, including:

1. Delaunay triangulation partitions a whole region into many smaller pieces and exactly describes the closest neighbour structures of the spatial features.
2. Insertion of a new point in a Delaunay triangulation affects only the triangles whose circum-circles contain that point. As a result, noise affects the Delaunay triangulation only locally.
3. The Delaunay triplets are not skinny which is desirable as the skinny triangles lead to instabilities and errors.
4. The Delaunay triangulation creates only O(n) triangles hence the computing cost greatly decreases using Delaunay triangulation.
5. Compared to other topological structures, the Delaunay triangulation is less sensitive to distortion.

The fundamental thought of utilizing Delaunay diagram is to stay away from the need of characterizing a separation edges for determining neighbourhoods as they don't need to repeat the way toward discovering neighbourhoods for different client characterized parameters and discovers neighbourhood powerfully.

3.1. Case Study

A Spatial Data Set with spatial feature set F = {A, B, C, D, E} is shown in Fig.1. Objects with various shapes represent different spatial features. The notations used to represent different features are given in Table 1:

Table 1
Notations used for various features

Feature	*Symbol*
A	Circle
B	Triangle
C	Hexagon
D	Square
E	Diamond

Feature A has 4 instances with the instance *ids* {$a1$, $a4$, $a10$, $a14$} which are represented with , Object B has 5 instances with the instance ids {$b2$, $b3$, $b15$, $b16$, $b19$}, Object C has 4 instances with the instance *ids* {$c7$, $c11$, $c13$, $c18$} Object D has 3 instances with the instance ids {$d6$, $d9$, $d17$}, Object E has 4 instances with the instance ids {$e5$, $e8$, $e12$, $e20$}. In Fig. 2, the neighbourhood relation R is defined using Delaunay triangulation [25, 27] instead of using Euclidean distance. Neighbouring instances are connected by edges. For instance {$a1$, $b3$, $c11$, $d6$, $e5$,}, {$a4$, $b15$, $c18$, $d17$, $e8$} and {$a10$, $b2$, $c7$, $d9$, $e12$} are all neighbour-sets because each set forms a clique. Here, we use the instance-id to refer to feature in figure 2. Additional neighbour-sets include {$b16$, $c13$, $d9$} and {$a14$, $b3$, $d6$, $e20$}. {A, B, C, D, E} is a co-location pattern. The neighbourhood-set {$a10$, $b2$, $c7$, $d9$, $e12$} is a row instance of pattern {A, B, C, D, E} but the neighbourhood-set {$a10$, $b2$, $c7$, $d9$, $e12$, $c13$} is not a row instance of co-location {A, B, C, D, E} because it has a proper subset {$a10$, $b2$, $c7$, $d9$, $e12$} which contains all the features in {A, B, C, D, E}. Finally row set ({A, B, C, D, E}) include only {$a10$, $b2$, $c7$, $d9$, $e12$} but not {$a10$, $b2$, $c7$, $d9$, $e12$, $c13$}.Once the outcomes of the above issues have been obtained, a case would be formulated with suggestions to enhance teaching and learning of software engineering.

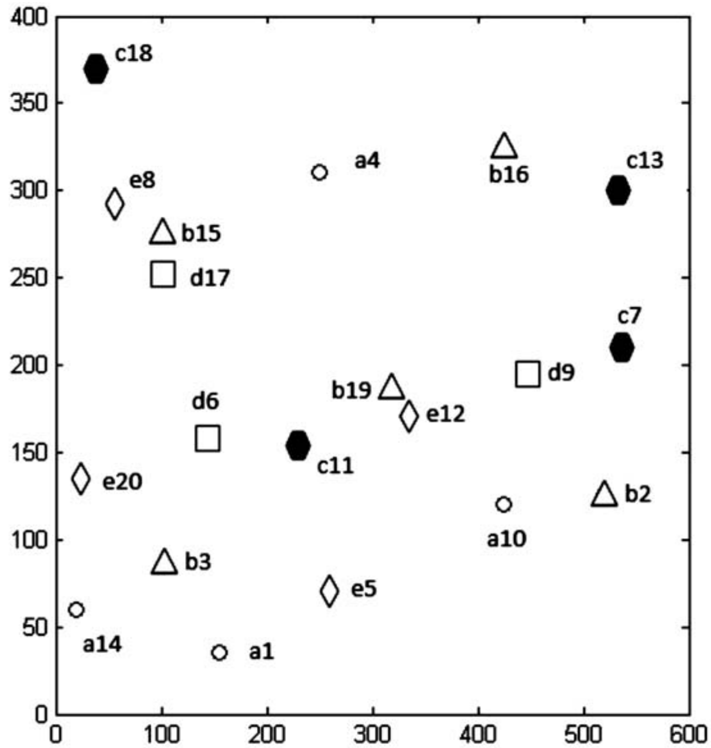

Figure 1: Sample dataset of colocation patterns

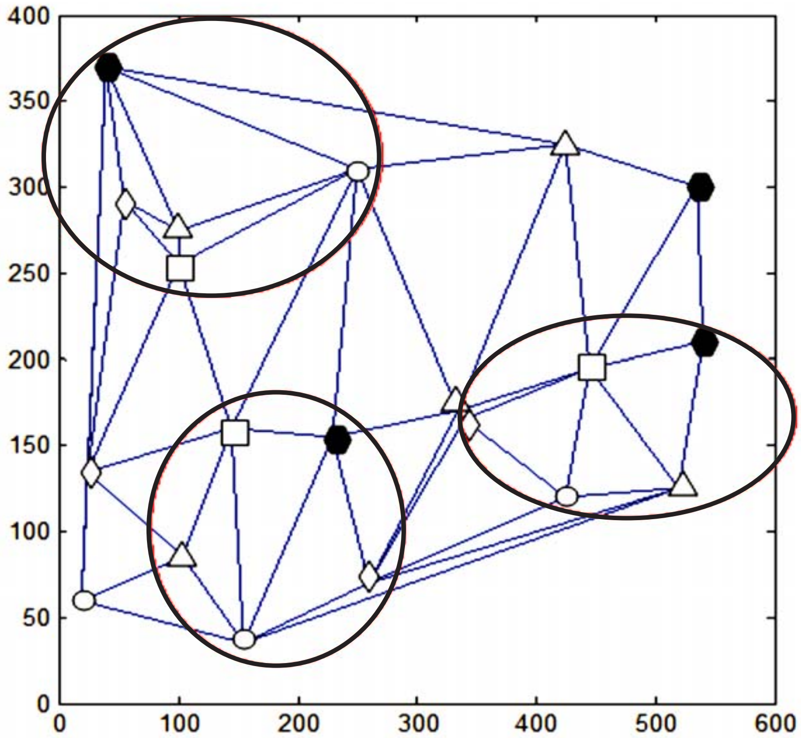

Figure 2: Delaunay triangulation of the patterns

Conditional Probability: For a co-location rule R: A $\rightarrow$ B, the conditional probability cp(R) is defined as

$$\frac{|\{L \in \text{rowest (A)} \, \exists \, L' \text{ such that } (L \subseteq L') \wedge L' \in \text{rowest}(A \cup B))\}|}{|\text{rowest (A)}|}$$

Rowset({A, B, C, D, E}) contains following three patterns {*a*1, *b*3, *c*11, *d*6, *e*5}, {*a*4, *b*15, *c*18, *d*17, *e*8} and {*a*10, *b*2, *c*7, *d*9, *e*12} as shown in figure 2. and rowset({A, B, C}) contains following three patterns {*a*4, *b*15, *c*18},{*a*4, *b*19, *c*11} and {*a*4, *b*16, *c*18}. Since |rowset({A, B, C})| = 3 and only one row of {A, B, C} satisfy the subset condition,{A, B, C, D, E}. The conditional probability *cp*({A, B, C} {D, E}) = 1/3= 33.3%. Let us consider that the occurrence of objects A, B and C represent the occurrence of a traffic jam and the occurrence of the objects D and E represent the occurrence of the Violation of road rules. When all the objects occur simultaneously, it represents the occurrence of a road accident. *cp*({A, B, C} {D, E}) represent the conditional probability of the occurrence of the road accident whenever the traffic is jammed. Traffic is jammed in three cases *i.e*. {*a*4, *b*15, *c*18}, {*a*4, *b*19, *c*11} and {*a*4, *b*16, *c*18}. Among these three, Accident occurs only when the traffic is jammed along with the violation of the road rules *i.e*. {*a*4, *b*15, *c*18}. Hence we conclude that, for every 3 traffic jams, there is a probability of 1 time occurring of an accident *i.e*. 33.3%.

Participation ratio: Given a spatial database S, to measure how a spatial feature f is co-located with other features in co-location pattern C, a participation ratio *pr*(C, *f*) can be defined as

$$\frac{|\{r \,|(r \in S) \wedge (r.\, f = f) \wedge (r \text{ is) in a row instance of C})\}|}{\{|\, r \,|\, (r \in S)\} \wedge (r.\, f = f)\}}$$

We first identify the row-sets of {A, B, C, D, E} that include 3 patterns. Among all the four instances of A *i.e*. {*a*1, *a*4, *a*10, *a*14} three of them, namely {*a*1, *a*10, *a*4} has B, C, D and E in a neighbour-set. So the participation ration *pr*({A, B, C, D, E}, A) = 3/4. Similarly we can have *pr*({A, B, C, D, E}, B) = 3/5, *pr*({A, B, C, D, E}, C) = 3/4 , *pr*({A, B, C, D, E}, D) = 3/3 and *pr*({A, B, C, D, E}, E) = 3/4. Participation ratio of each object in the given event represents the ratio of the distinct instances of the objects participating in the event to the total number of instances of the respective object.

Participation index: Participation index (PI) is the minimal of participation ratios of all the features in the event. The significance of PI is that, when the least participated feature occurs, there may high possibility of remaining features also participate in the event. Here for the event {A, B, C, D, E}, the participation ratios of {A, B, C, D, E} are {3/4, 3/5, 3/4, 3/3, 3/4} respectively. Among these, feature B has minimal participation ratio *i.e*. 3/5. Hence, the participation index PI({A,B,C,D,E }) is 3/5 which correspond to feature B.

Max Participation index: Participation index (PI) is the maximal of participation ratios of all the features in the event. This can be utilized to know about the spatial features which impact the most on the event and corresponding action can be take care. The participation proportions of A,B,C,D and E are 3/4, 3/5, 3/4, 3/3, 3/4 respectively for the event {A, B, C, D, E}. Among these, we can see that D has the highest impact of 100%.

Medoid Participation index: We generally utilize the element maxPI to choose the most encouraging component or the one that should be prevented. We will analyse the occurrence of an accident *i.e*. event with an example. Let A is a car, D corresponds to traffic police and B corresponds to traffic signal. In this case, we can conclude that the occurrence of feature D should be prevented. But it is not possible to prevent a traffic police to enter onto the road. For this situation, we can't utilize the variable maxPI to choose the element that must be taken care. Hence, for this situation maxPI has not working and we use another approach called Medoid Approach [26], which takes the medoid of the Participation ratios of the features in the list {A, B, C, D, E}. The participation ratios of {A, B, C, D, E} are 3/4, 3/5, 3/4, 3/3, 3/4 respectively. The PI is 3/5 which corresponds to feature B, maxPI is 3/3 which corresponds to feature D, medoidPI is 3/4 which corresponds to features A, C and E.

4. EXPERIMENTAL RESULTS

In this section we reports the results of the proposed approach. We examined this approach on a synthetic dataset which is given in figure 1. We customized the Colocation pattern mining calculation using Delaunay triangulation approach. We used the following measures to analyse the performance of the proposed approach which are explained in Section 3.

1. Participation Ratio
2. Participation Index
3. Max Participation Index
4. Medoid Participation Index.

It can be seen from the Table 2 that the Delaunay triangulation approach achieves a PR of {0.75, 0.6, 0.75, 1, 0.75} for the colocation patterns {A, B, C, D, E} respectively. The PI, MaxPI, MedPI of the proposed approach are 0.6, 1, 0.75 respectively for the colocation patterns {A, B, C, D, E}. Further, we also analysed this approach with our previous work [26] which is also shown in Table 2. It can be observed that the proposed approach performs better compared our previous work in terms of PR, PI, MaxPI and MedPI. The analysis charts from this work also shown in Fig. 3 and Fig. 4. The Participation Ratios of different features {A, B, C, D, E} with the proposed and the method in [26] are shown in Fig. 3. Further, PI, MaxPI and MedPI of these two approaches are given in Fig. 4.

5. CONCLUSION

In this paper, we proposed an efficient method for finding colocation patterns from the spatial databases using Delaunay triangulation. Our method is based on a case study shown in Fig. 1. The proposed method achieved a better PR, PI, MaxPI and MedPI compared to other approaches. This shows the efficacy of the

Table 1
Outcomes of group-projects coursework

Measure	*Co-Location*	*Rowset*	*PR*	*PI*	*Max PI*	*Med PI*
Euclidean distance [26]	{A, B, C, D, E}	{*e*12, *d*9, *b*2, *a*10, *c*13}, {*e*12, *d*9, *b*16, *a*10, *c*13}, {*e*5, *d*6, *b*3, *a*10, *c*11}, {*e*5, *d*6, *b*3, *a*14, *c*11}, {*e*5, *d*6, *b*3, *a*1, *c*11}, {*e*12, *d*9, *b*2, *a*10, *c*7}, {*e*12, *d*9, *b*16, *a*10, *c*7}, {*e*5, *d*6, *b*3, *a*10, *c*7}	{0.5, 0.67, 0.6, 0.75, 0.75}	0.5	0.7 5	0.67
Delaunay	{A,B,C,D,E}	{*a*1, *b*3, *c*11, *d*6, *e*5}, {*a*4, *b*15,*c*18, *d*17, *e*8}, {*a*10, b2, *c*7, *d*9, *e*12}	{0.75, 0.6, 0.75, 1, 0.75}	0.6	1	0.75

Delaunay triangulation in finding the colocation patterns from the spatial databases. In Future, we will experiment this work on benchmark spatial dataset. Further, we will extend the Delaunay triangulation to finding the colocation patterns.

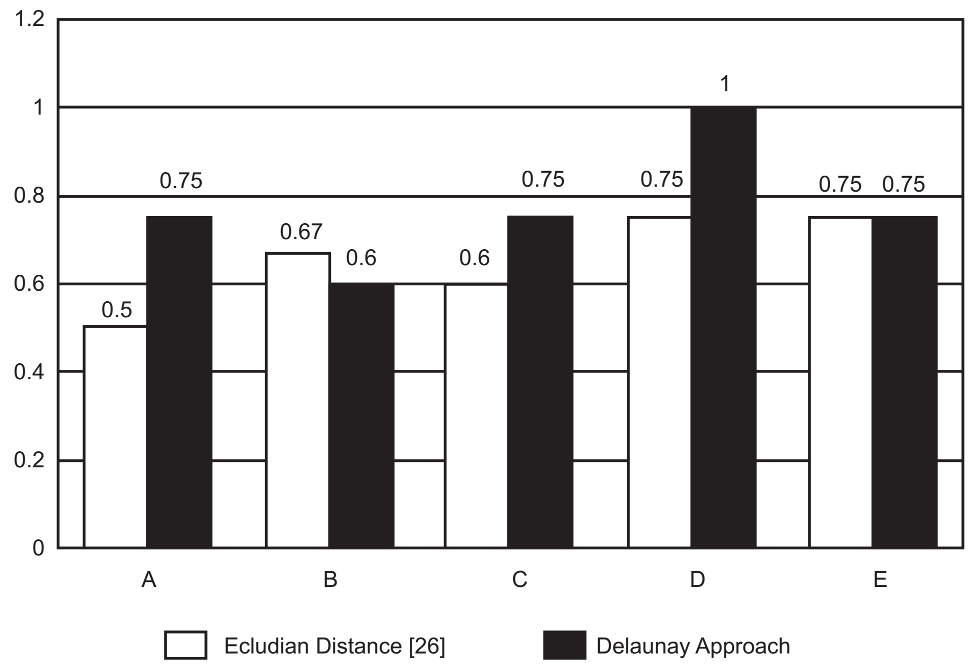

Figure 3: Participation Ratio (PR) comparison for features {A, B, C, D, E}

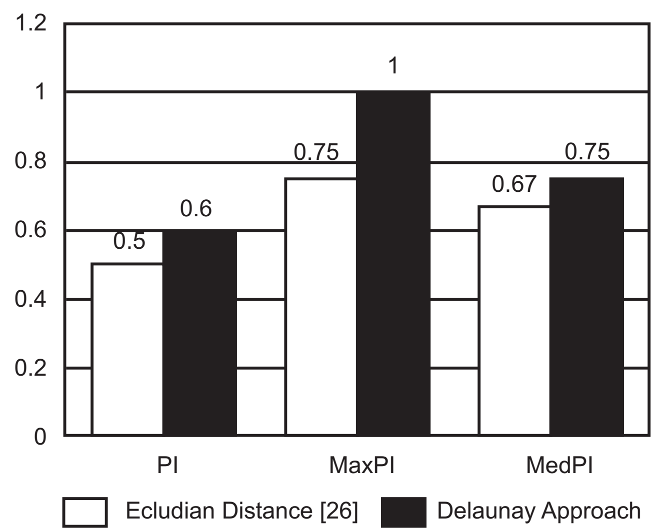

Figure 4: PI, MaxPI and MedPI comparison

REFERENCES

[1] Dasu, T., Johnson, T. Exploratory Data Mining and Data Cleaning, Wiley, 2003.

[2] Hand, D., Mannila H., Smyth, P., Principles of Data Mining, MIT Press, 2001.

[3] Griffith D. "Statistical and Mathematical sources of regional Science theory. Map pattern analysis as an example." Regional science RSAI 1999.

[4] M. S. Chen, J. Han, P. S. Yu. "Data mining, an overview from database perspective", IEEE Transactions on Knowledge and data Engineering, 1997.

[5] K Zeitouni "A survey of spatial data mining methods databases and statistics point of views",

Data warehousing and web engineering, 2002 - books.google.com.

[6] Longley P. A., Goodchild M. F., Maguire D. J., Rhind D. W., Geographical Information Systems - Principles and Technical Issues, John Wiley & Sons, Inc., Second Edition, 1999

[7] Diansheng Guo, Jeremy Mennis ,2009 " Spatial data mining and geographic knowledge discovery—An introduction", Computers, Environment and Urban Systems, Elsevier.

[8] Gordon, A. D. 1996. "Hierarchical classification. Clustering and classification" (pp. 65–122). River Edge, NJ, USA: World Scientific Publisher.

[9] Deren LI and Shuliang WANG, "Concepts, principles and applications of spatial data mining and knowledge discovery", in ISSTM 2005, August, 27-29, 2005, Beijing, China.

[10] Manuel Alfredo PECHPALACIO, "Spatial Data Modeling and Mining using a Graph-based Representation", PhD Thesis.

[11] R. Agarwal, T. Imielinski, and A. Swami. "Mining association rules between sets of items in large databases," in Proc. of the ACM SIGMOD Conference on Management of Data, Washington, DC, pp. 207–216, 1993.

[12] Han, J., Kamber, M., 2001, Data Mining: Concepts and Techniques (San Francisco: Academic Press)

[13] S Shekhar, P Zhang. Data Mining and Knowledge Discovery 2010 – Springer.

[14] Shekhar, S., & Huang, Y. (2001). Discovering spatial co-location patterns: A summary of results. In C. Jensen, M. Schneider, B. Seeger, & V. Tsotras (Eds.), Advances in spatial and temporal databases, proceedings, lecture notes in computer science (pp. 236–256). Berlin: Springer-Verlag.

[15] Zhang, N. Mamoulis, D.W.L Cheung, and Y. Shou. "Fast mining of spatial collocations", in Proc. of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, pp. 384–393, 2004.

[16] Ester M., Kriegel H.-P., Sander J., and Xu X. 1996 "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining. Portland, Oregon, AAAI Press, Menlo Park, California, pp. 226-231.

[17] Sander J., Ester M., Kriegel H.-P., and Xu X. 1998 "Density-Based Clustering in Spatial Databases: A New Algorithm and its Applications", Data Mining and Knowledge Discovery, an International Journal, Kluwer Academic Publishers, Vol.2, No. 2.

[18] Li D.R., Wang S.L., Li D.Y. and Wang X.Z., 2002, Theories and technologies of spatial data knowledge discovery. Geomatics and Information Science of Wuhan Univrsity 27(3), 221-233.

[19] Spielman, S. E., & Thill, J. C. , 2008, " Social area analysis, data mining and GIS" . Computers Environment and Urban Systems, 32(2), 110–122.

[20] Ester, M., Kriegel, H. P., & Sander, J. ,1997. "Spatial data mining: A database approach ". In Advances in spatial databases (pp. 47–66). Berlin: Springer-Verlag Berlin

[21] Martin Ester, Hans-Peter Kriegel, Jörg Sander "Algorithms and Applications for Spatial Data Mining", Geographic Data Mining and Knowledge Discovery, Research Monographs in GIS, Taylor and Francis, 2001.

[22] Yan Huang, Jian Pei, Hui Xiong, "Mining Co-Location Patterns with Rare Events from Spatial Data Sets Geoinformatica (2006) 10: 239–260.

[23] G. Kiran Kumar, T.Venu gopal and P.Premchand "A Novel method of modeling Spatial Co-location patterns on spatial Database", 2nd International conference ICFOCS 2011 held at IISc Bangalore,India Aug 7-9, 2011.

[24] S. Shekhar, and Y. Huang. "Co-location rules mining: A summary of results," in Proc. 7th Intl.Symposium on Spatio-temporal Databases, Springer, Berlin Heidelberg New York, p.236, 2001.

[25] Kavati, I., Prasad, M. V., & Bhagvati, C. (2013, August). Vein pattern indexing using texture and hierarchical decomposition of delaunay triangulation. In International Symposium on Security in Computing and Communication (pp. 213-222). Springer Berlin Heidelberg.

[26] Kumar, G. K., Premchand, P., & Gopal, T. V. (2012). Mining of spatial co-location pattern from spatial datasets. Int. J. Comput. Applic, 42, 25-30.

[27] Kavati, I., Chenna, V., Prasad, M. V., & Bhagvati, C. (2014, September). Classification of Extended Delaunay Triangulation for Fingerprint Indexing. In 2014 8th Asia Modelling Symposium (pp. 153-158). IEEE.

[28] Y. Morimoto,"Mining Frequent Neighboring Class Sets in Spatial Databases," In Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2001.

[29] Jin SoungYoo, Douglas Boulware and David Kimmey,"A Parallel Spatial Co-location Mining Algorithm Based on MapReduce," IEEE International Congress on Big Data, Anchorage, AK,pp. 25 – 31, 2014.

[30] Tran Van Canh and Michael Gertz, "A Constraint neighborhood based approach for co-location pattern mining," Fourth International Conference on Knowledge and System Engineering, pp. 128-135, 2012.