Incremental Clustering on Hotspot Data as Forest and Land Fires Indicator in Sumatra

I S Sitanggang1*, A A N Risal2, L Syaufina3
1,2Department of Computer Science, Bogor Agricultural University, Bogor 16680, Indonesia
3Department of Silviculture, Bogor Agricultural University, Bogor 16680, Indonesia
*)Email: imas.sitanggang@apps.ipb.ac.id

Abstract
Hotspot as one of forest and land fires indicators is regularly monitored for early detection of the fires. Large fires are difficult to detect based on a single occurrence of hotspot. However, hotspots that are occurred in high density clusters can be used as indicator of large forest and land fires. Our previous studies have applied several clustering algorithms on hotspot datasets in Sumatra and Kalimantan. However, those studies have not included new hotspot data in the clusters of hotspots. This study aims to implement incremental clustering on hotspot datasets in Sumatra. Clustering was performed using the modified spatio-temporal density based clustering algorithm namely ST-DBSCAN. The main process of the incremental clustering in hotspot datasets is to update the initial clusters based on the existence of new hotspot data without doing clustering from the beginning. The incremental clustering algorithm has three parameters namely Eps1, Eps2 and MinPts. Eps1 is the distance parameter for spatial attributes whereas Eps2 is the distance parameter for non-spatial attributes. Eps represents maximum radius of the neighbourhood. MinPts is the minimum number of objects within Eps1 and Eps2 distance of an object. The algorithm was applied on the hotspot dataset in Sumatra in the period of September to October 2017. As many 2659 hotspots were used in initial clustering and 1023 new hotspots were used for updating the initial clusters. The initial clustering at the parameter eps1 of 0.1, eps of 7, and MinPts of 5 results 16 clusters and 23 outliers. The incremental clustering at those parameters results 7 new clusters and 6 new outliers.

Keywords: hotspot, incremental clustering, peatland fire, ST-DBSCAN

1. Introduction
Forests are natural resources that provide and control various needs of human life. However, negative impacts of exploitative forest management cause many problems so that destruction of forest function is very worrying. One of the causes of forest degradation is forest fire. Forest and land fires have a major impact on the survival of forest ecosystem and decrease economic value of forests and soil productivity. In addition forest and land fires often cause air pollution which can disrupt human activities and health. Early detection of forest and land fires is an effort to protect forests and supervision of forest and land fires. Ministry of Forestry and Environment Indonesia has developed the web-based application Sipongi which visualizes hotspot occurrences in Indonesia. According to Ministry of Forestry and Environment Indonesia (2009) a hotspot is an indicator of forest fires detecting locations that have relatively higher temperatures compared to those of surrounding
locations. Moreover hotspots which are located in clusters are considered as strong indicators of forest and land fires.

Clustering is a method of data analysis in data mining that aims to group data with the same characteristics into the same group and data with different characteristics to different groups. Rao et al.[1] describes spatio-temporal clustering as a grouping process based on spatial and temporal aspects of the data. A hotspot as one of indicators for forest and land fires has spatial features representing the location where the hotspots are occurred and the temporal feature indicating when the hotspot data are acquired. Therefore clustering on hotspot datasets should be performed based on the spatial and temporal features to get accurate results as indicator for land and forest fires. This study aims to apply spatio-temporal clustering algorithm on hotspot dataset in Sumatra in 2015 by considering availability of new hotspot data. The algorithm used for clustering is ST-DBSCAN as the extension of the DBSCAN (Density-Based Spatial Clustering Algorithm with Noise) algorithm. The modified ST-DBSCAN was applied to update hotspot clusters because of the existing new hotspot data.

2. Literature review

Previous studies have successfully identified clusters of hotspot in Sumatra and Kalimantan Islands. Wijaya et al. [2] applied the density based clustering method on hotspot data with road and river as physical obstacles. The data used in this study were hotspot in 2014 that were occurred in peatland in Riau province Indonesia. The algorithm used is CPO-WCC (Clustering in Presence of Obstacles with Computed number of Cells) which produces three clusters of hotspot. The area of dense cluster is 10 202.10 km$^2$ with number of hotspots per km$^2$ is 0.985. This study shows that hotspots were mostly occurred in peatland with type of Hemists/Saprist (60/40) and depth greater than 400 cm.

The distribution pattern of hotspot clusters in the peatland areas in Sumatera in the year 2014 was studied using Kulldorff’s Scan Statistics (KSS) method with Poisson model by Kirana et al. [3]. This study showed that the method is reliable to detect the clusters of hotspots with the accuracy of 95%. High density of hotspot clusters were found in Riau and South Sumatera province. Hotspot clusters were mostly found in peatland with ‘hemic’ maturity level and very deep thickness.

The contextual outliers on hotspot datasets in the period of 2001 to 2009 were identified based on climate context, i.e. rainfall using the K-means algorithm by Thah and Sitanggang [4]. The study detected 54 objects as contextual outliers, many of them occurred in February, March, June, July, and August. The contextual outliers detected have an average of daily occurrences is 65.76 hotspots with an average of rainfall is 37.15 mm.

Areas that have high density hotspot in the peat land area of Sumatra Island were determined using the DBSCAN algorithm by Usman et al. [5]. The data used are hotspot data in Sumatra Island in the years of 2002 and 2013. The study revealed that that there were changes in the pattern of hotspot distribution and land cover of peat land from 2002 to 2013, where in 2002 the distribution of hotspot are mostly found in moderate depth (100-200 cm) peat land, whereas in 2013, the distribution of hotspot were mostly found in very deep (> 400 cm) peat land area. A web-based application for clustering hotspot data in peatlands in Sumatera was built by Hermawati and Sitanggang [6] using the Shiny framework which is available in programming language R. The main functions of application are to perform clustering on hotspots data in peatland in 2002 and 2013 using the DBSCAN algorithm as done in the study by Usman et al (2015) and to visualize clustering results based on a type of land use, land depth, and peat type.

The DBSCAN algorithm was applied on hotspot data in Riau Province in between year 2001 to 2012 by Sukmasetya and Sitanggang [7]. The objective of this study is to detect outliers on hotspots data indicating hotspots which are located far from the others. The study shows that the highest occurrence of outliers is in 2005 which spread across 11 districts/cities and 136 districts. In 2005 the highest number of outliers is found in Rokan Hulu which was located at 186 points. The highest frequency of hotspot considered as outliers is found in August 2005, with a total of 355 outliers in which as many 97 of these outliers are occurred in Rokan Hulu District.
Global and collective outliers were detected on hotspot data in Riau Province in Sumatera Island for the period 2001-2012 [8]. The data used in this study are 4383 daily hotspots and 144 monthly hotspots. Based on the clustering results using the K-means algorithm, the study obtains 59 collective outliers and 30 global outliers on the hotspot dataset which were mostly occur in February, March, June, July, and August. The highest frequency of outliers occurred in 2005 in which as many 1118 hotspots were found in the northern part of the Riau province on 21 June 2005. This study found that in August 2005 outliers spread on the whole area of Riau Province whereas for the period 2001-2012 there are no outliers occurred in April, November and December. A web-based application was developed by Suci and Sitanggang [9] to detect outliers on hotspot data in the period of 2001 to 2012 and to visualize the outliers based on the time and location. The application enables users to easily identify the global and collective outliers on hotspot dataset instead of manually detected as done in the study by Sitanggang and Baehaki [8]. Shiny framework with the R programming language is used in system implementation.

Hotspot data are daily recorded by some institutions in Indonesia. Therefore new hotspot data should be included in clusters which are created for historical hotspot data in order to provide up to date hotspot clusters. All those studies on hotspot data clustering using DBSCAN and K-Means algorithm do not consider new hotspot data. This study aims to apply the incremental clustering algorithm based on spatial and temporal aspects on the hotspot dataset. The main process on the algorithm is to update the existing clusters due to the availability of new hotspot data.

3. Methods
3.1 Data and Study Area
Incremental spatio-temporal density based clustering was applied on the hotspot dataset in Sumatra in April, May, September, and October 2015. The hotspot datasets were collected from FIRMS NASA http://www.firms.modaps.eosdis.nasa.gov/. The attributes used in clustering are latitude, longitude, and date of hotspot occurrence (Acq_date).

3.2 DBSCAN Algorithm
The DBSCAN algorithm was first introduced by Esther in 1996 [10]. The algorithm has two global parameters namely maximum radius of the neighbourhood (Eps) and minimum number of objects in an Eps-neighbourhood of the objects (MinPts). Using those two parameters, the DBSCAN algorithm identifies three types of objects namely border, core, and outlier, based on density of objects in a dataset as illustrated on Figure 1. A core object is an object with at least MinPts objects within a radius ‘Eps-neighborhood’ whereas a border object is an object that on the border of a radius ‘Eps-neighborhood’. Core and border objects create dense area in the dataset whereas other objects in low dense area are identified as outliers.

![Figure 1. Border object, core object and noise](image)
Figure 2(a) shows relation between q as a core object and p as a border object. The object p is called directly density reachable from q. A point p is directly density reachable from q wrt Eps and MinPts if [10]

1. \( p \in N_{Eps}(q) \), where \( N_{Eps}(q) = \{ r \in D | dist(q,r) \leq Eps \} \)
2. \( |N_{Eps}(q)| \geq MinPts \) (Core point condition)

Figure 2(b) shows that if the object p is density-connected to q with respect to eps and MinPts if there is an object o in the dataset such that p and q are density reachable from o.

The main tasks in the DBSCAN algorithm are as follows:
1. Select an object p in the dataset
2. Identify all objects which are density-reachable from p with respect to Eps and MinPts
3. If p is a core object then a cluster is created
4. If p is a border object and no other objects are density-reachable from p then the algorithm visits the next object in the dataset
5. Continue step 2 to step 4 until all of the object in the dataset have been processed

3.3 ST-DBSCAN Algorithm

ST-DBSCAN algorithm is an extension of the DBSCAN algorithm by involving not only spatial aspect but also non-spatial attributes of objects in a dataset [11]. ST-DBSCAN has four parameters: eps1, eps2, MinPts and \( \Delta \epsilon \). Eps1 is the distance parameter for spatial attributes (latitude and longitude). Eps2 is the distance parameter for non-spatial attributes [11]. In this study, the distance metric Euclidean is used for eps1 whereas Manhattan is used for eps2. MinPts is the minimum number of objects within Eps1 and Eps2 distance of an object. The parameter \( \Delta \epsilon \) is used to prevent the discovering of combined clusters because of the little differences in non-spatial values of the neighboring locations [11]. Figure 3 shows the ST_DBSCAN algorithm [11].
3.4 Incremental Spatio-Temporal Density Based Clustering

Spatio-temporal data such as hotspots as indicator of forest and land fires are continually updated as new data are included in the dataset. New objects will influence the formation of objects in existing clusters. This study applied the incremental clustering method on the hotspot dataset to obtain new formation of clusters due to availability of new hotspot data included in the dataset. Chakraborty and Nagwani [12] proposed the incremental clustering algorithm which is applied on the DBSCAN algorithm. The incremental DBSCAN clustering algorithm uses the parameters eps1, Minpts and core

Algorithm ST_DBSCAN (D, Eps1, Eps2, MinPts, Δε)

// Inputs: D= (O1, O2, …, On) set of objects, Eps1: Maximum geographical coordinate (spatial) distance value, Eps2: Maximum non-spatial distance value, MinPts: Minimum number of points within Eps1 and Eps2 distance, Δε: Threshold value to be included in a cluster.
// Output: C= (C1, C2, …, Ck) set of clusters.
// C= (C1, C2, …, Ck) set of clusters

Cluster_Label= 0
for i= 1 to n  // (i)
    if Oi is not in a Cluster then  // (ii)
        X= Retrieve_Neighbors (Oi, Eps1, Eps2)  // (iii)
        if |X| < MinPts then
            Mark Oi as noise  // (iv)
        else  // construct a new cluster (v)
            Cluster_Label = Cluster_Label + 1
            for j=1 to |X|
                Mark all objects in X with current Cluster_Label
            end for
            Push (all objects in X)  // (vi)
        end if
    end if
while not IsEmpty()  // (vii)
    CurrentObj= Pop ()
    Y= Retrieve_Neighbors (CurrentObj, Eps1, Eps2)
    if |X| >= MinPts Then
        forAll objects O in Y  // (vii)
            if (O is not marked as noise or it is not in a cluster) and |Cluster_Avg () - O.Value| <= Δε Then
                Mark O with current Cluster_Label
                push (O)
            end if
        end for;
    end if;
end while;
end if;
end for;
e nd for;
end algorithm;

Figure 3. ST-DBSCAN algorithm [11].
objects of existing clusters. Figure 4 shows how the distance between each new object and core objects of existing clusters is determined. The new object which has minimum distance to a core object with respect to Minpts will be assigned to the cluster where the core object is located. Otherwise the new data are considered as outliers.

**Figure 4.** Illustration of spatial distance calculation on the incremental DBSCAN clustering algorithm.

New data which have been considered as outliers are then merged with the existing outliers from the previous clustering to create a new dataset. This new dataset is then grouped using the DBSCAN algorithm. Figure 5 illustrates clustering process on outliers. If the outliers meet the parameters eps1 and MinPts then a new cluster is created. Incremental DBSCAN clustering may results new outliers, new clusters, new members in the existing clusters or merging two clusters to create a new cluster [13].

**Figure 5.** Illustration of creating a new cluster in the incremental DBSCAN clustering
Instead of using the incremental DBSCAN clustering, this study uses the incremental ST-DBSCAN clustering on hotspot datasets. The algorithm has input parameters eps1, eps2 and minpts as in ST-DBSCAN. In addition to those parameters, the incremental ST-DBSCAN clustering requires the hotspot clusters which are determined using the ST-DBSCAN algorithm. When new hotspot data are available, Euclidean distance between those new data to the medoids of existing clusters is calculated. The object which has minimum distance less than or equal to eps1 will be included in the next process. Manhattan distance between new data that meet eps1 and objects in existing clusters is determined on the variable date of hotspot acquired. If minimum Manhattan distance of an object is less than or equal to eps2 and the objects and the object has at least MinPts neighbors then the object will be assigned to the existing cluster as a new member. Otherwise the object is considered as an outlier. The new data which assigned as outliers and outliers from the previous clustering process are grouped using the ST-DBSCAN algorithm to produce new clusters or new outliers. The tasks in incremental ST-DBSCAN clustering will be repeated until all new data have processed.

4. Results and Discussion
The incremental spatio-temporal density based clustering has been applied on the hotspot dataset in Sumatra. The dataset contains as many 2659 hotspots in September and October 2015. The algorithm has been tested to process as 1023 new hotspot points with the ST-DBSCAN parameters eps1 0.1, eps2 3, 7, 30 and MinPts 5. The parameter Eps = 0.1 was selected because according to the study by Usman et al (2015) this parameter value gives small Sum Square Error in hotspot clustering compared to other values of Eps1 [5]. In addition, MinPts of 5 is used to produce few clusters so that analysis of those clusters is easy to be performed. The values of temporal parameter eps2 of 3, 7, 30 are selected to obtain hotspot clusters on the period of three days, a week, and a month respectively. Table 1, Table 2 and Table 3 provides the results of spatio-temporal density based clustering on hotspot data with the parameter eps2 3, 7, and 30 respectively.

Table 1. The results of incremental spatio-temporal density based clustering on the Sumatra hotspot dataset with the parameter eps1 0.1, eps2 3 and MinPts 5.

| Cluster label | Number of member | Cluster label | Number of member | Number of new member |
|---------------|------------------|---------------|------------------|---------------------|
| 0             | 41               | 0             | 59               | 18                  |
| 1             | 82               | 1             | 84               | 2                   |
| 2             | 18               | 2             | 25               | 7                   |
| 3             | 7                | 3             | 32               | 25                  |
| 4             | 20               | 4             | 25               | 5                   |
| 5             | 8                | 5             | 23               | 15                  |
| 6             | 6                | 6             | 6                | 0                   |
| 7             | 254              | 7             | 262              | 8                   |
| 8             | 273              | 8             | 328              | 55                  |
| 9             | 406              | 9             | 407              | 1                   |
| 10            | 23               | 10            | 25               | 2                   |
| 11            | 9                | 11            | 13               | 4                   |
| 12            | 40               | 12            | 40               | 0                   |
Table 2. The results of incremental spatio-temporal density based clustering on the Sumatra hotspot dataset with the parameter eps1 0.1, eps2 7 and MinPts 5.

| Cluster before applying Incremental Spatio-temporal Clustering | Cluster after Incremental Spatio-temporal Clustering |
|---------------------------------------------------------------|-----------------------------------------------------|
| **Cluster label** | **Number of member** | **Cluster label** | **Number of member** | **Number of new member** |
| 13 | 20 | 13 | 20 | 0 |
| 14 | 1023 | 14 | 1112 | 89 |
| 15 | 9 | 15 | 11 | 2 |
| 16 | 179 | 16 | 217 | 38 |
| 17 | 8 | 17 | 10 | 2 |
| 18 | 37 | 18 | 44 | 7 |
| 19 | 12 | 19 | 18 | 6 |
| 20 | 112 | 20 | 113 | 1 |
| 21 | 46 | 21 | 72 | 26 |
| 22 | 7 | 22 | 15 | 8 |
| 23 | 10 | 23 | 16 | 6 |
| 24 | 7 | 24 | 8 | 1 |
| 25 | 360 | 360 | 1 |
| 26 | 24 | 24 | 1 |
| 27 | 8 | 8 | 1 |
| 28 | 250 | 250 | 1 |
| 29 | 13 | 13 | 1 |
| 30 | 7 | 7 | 1 |
| 31 | 15 | 15 | 1 |
| 0 | 23 | 0 | 29 | 6 |
| 1 | 336 | 1 | 337 | 1 |
| 2 | 437 | 2 | 438 | 1 |
| 3 | 7 | 3 | 27 | 20 |
| 4 | 40 | 4 | 64 | 24 |
| 5 | 5 | 5 | 6 | 1 |
| 6 | 273 | 6 | 344 | 71 |
| 7 | 23 | 7 | 64 | 41 |
| 8 | 9 | 8 | 13 | 4 |
| 9 | 46 | 9 | 46 | 0 |
Cluster before applying Incremental Spatio-temporal Clustering | Cluster after Incremental Spatio-temporal Clustering
---|---
Cluster label | Number of member | Cluster label | Number of member | Number of new member
10 | 1024 | 10 | 1139 | 115
11 | 22 | 11 | 30 | 8
12 | 293 | 12 | 339 | 46
13 | 54 | 13 | 58 | 4
14 | 37 | 14 | 46 | 9
15 | 22 | 15 | 27 | 5
16 | 7 | 16 | 23 | 16
17 | 344 | | 344 |
18 | 20 | | 20 |
19 | 24 | | 24 |
20 | 5 | | 5 |
21 | 9 | | 9 |
22 | 18 | | 18 |
23 | 232 | | 232 |

Table 1 and Table 2 show that as many 7 new clusters are created when the algorithm is applied on the dataset at the parameters \( \varepsilon_1 = 0.1 \), \( \varepsilon_2 = 3 \), 7 and \( \text{MinPts} = 5 \). In addition, some new members are included on the existing clusters. New members of clusters are new hotspot data that meet the parameters \( \varepsilon_1 = 0.1 \), \( \varepsilon_2 = 3 \) and \( \text{MinPts} = 5 \) when they are evaluated with the Medoids and member of existing clusters. When the parameters \( \varepsilon_2 \) is set to 30, number of new clusters is increased as shown on Table 3. Number of new hotspot data which are assigned to the existing clusters at the parameter \( \varepsilon_2 = 3 \), 7, and 30 is 33.82%, 36.27% and 22.78% of the all new hotspot data (1023 points), respectively. New outliers and new clusters are created which contain 66.18%, 63.73% and 77.22% data of all new hotspot data.

Table 3. The results of incremental spatio-temporal density based clustering on the Sumatra hotspot dataset with the parameter \( \varepsilon_1 = 0.1 \), \( \varepsilon_2 = 30 \) and \( \text{MinPts} = 5 \).
Cluster before applying Incremental Spatio-temporal Clustering  
Cluster after Incremental Spatio-temporal Clustering

| Cluster label | Number of member | Cluster label | Number of member | Number of new member |
|---------------|------------------|---------------|------------------|---------------------|
| 7             | 20               |               |                  | 20                  |
| 8             | 8                |               |                  | 8                   |
| 9             | 24               |               |                  | 24                  |
| 10            | 6                |               |                  | 6                   |
| 11            | 74               |               |                  | 74                  |
| 12            | 18               |               |                  | 18                  |
| 13            | 5                |               |                  | 5                   |
| 14            | 338              |               |                  | 338                 |

Figure 6, Figure 8, Figure 10 show the clusters of hotspots as the results of ST-DBSCAN algorithm which was applied on the hotspot datasets of Sumatra in October 2015 at the parameter eps 1 0.1, eps2 3, 7, 30, and MinPts 5. The output of the incremental spatio-temporal clustering algorithm on new hotspot data of Sumatra in October 2015 at the parameter eps 1 0.1, eps2 3, 7, 30, and MinPts 5 are provided on Figure 7, Figure 9, and Figure 11.

Figure 6. Clusters of hotspot data as the results of ST-DBSCAN algorithm in Sumatra in October 2015 at the parameter eps 1 0.1, eps2 3, and MinPts 5.
Figure 7. Clusters of hotspot data as the results of Incremental spatio-temporal clustering algorithm in Sumatra in October 2015 at the parameter eps1 0.1, eps2 3, and MinPts 5.

Figure 8. Clusters of hotspot data as the results of ST-DBSCAN algorithm in Sumatra in October 2015 at the parameter eps1 0.1, eps2 7, and MinPts 5.
Figure 9. Clusters of hotspot data as the results of Incremental spatio-temporal clustering algorithm in Sumatra in October 2015 at the parameter eps1 0.1, eps2 7, and MinPts 5.

Figure 10. Clusters of hotspot data as the results of ST-DBSCAN algorithm in Sumatra in October 2015 at the parameter eps 1 0.1, eps2 30, and MinPts 5.
Figure 11. Clusters of hotspot data as the results of Incremental spatio-temporal clustering algorithm in Sumatra in October 2015 at the parameter eps1 0.1, eps2 30, and MinPts 5.

5. Conclusion
This study has successfully applied the incremental spatio-temporal clustering algorithm on hotspot dataset in Sumatra in October 2015. The algorithm was designed to handle new hotspot data included to the existing hotspot datasets. The results of incremental spatio-temporal clustering algorithm implementation are new members assigned to the existing clusters which are generated using the ST-DBSCAN algorithm, new outliers, and new clusters. The implementation of the algorithm on 1023 new hotspot points was conducted on the parameters eps1 of 0.1, eps of 3, 7, 30 and MinPts 5. The results show that about 22% to 36% of new hotspot data are assigned to the existing clusters which were created by the ST-DBSCAN algorithm. Moreover, about 66% to 77% of new hotspot data are labeled as outliers and form new clusters. Further works include validation of the results using the real forest and land fires data and implementation of the algorithm on real time hotspot datasets.

References
[1] Rao K V, Govardhan A and Rao K C 2012 Spatiotemporal data mining: issues, tasks and applications International Journal of Computer Science and Engineering Survey 3 39
[2] Wijaya P T, Sitanggang I S and Syaufina L 2016 Density based clustering of hotspots in peatland with road and river as physical obstacles Indonesian Journal of Electrical Engineering and Computer Science 3 3
[3] Kirana A P, Sitanggang I S and Syaufina L 2015 Poisson clustering process on hotspot in peatland area using Kulldorff’s Scan Statistics method Telecommunication Computing Electronics and Control (TEKOMNIKA) 13 4
[4] Thah P H and Sitanggang I S 2016 Contextual outlier detection on hotspot data in Riau Province using k-means algorithm Procedia Environmental Sciences 33 258–268
[5] Usman M, Sitanggang I S and Syaufina L 2015 Hotspot distribution analyses based on peat characteristics using density-based spatial clustering Procedia Environmental Sciences 24 132–140
[6] Hermawati R and Sitanggang I S 2016 Web-Based clustering application using Shiny framework and DBSCAN algorithm for hotspots data in peatland in Sumatra Procedia Environmental Sciences 33 317–323
[7] Sukmasetya P and Sitanggang I S 2016 Outlier detection on hotspots data in riau province using DBSCAN algorithm IOP Conference Series: Earth and Environmental Science 31 1
[8] Sitanggang I S and Baehaki D A M 2015 Global and collective outliers detection on hotspot data as forest fires indicator in Riau province Indonesia Proc. Second IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services. (IEEE Xplore) p 66-70

[9] Suci A M Y A and Sitanggang I S 2016 Web-based application for outliers detection on hotspot data using K-Means algorithm and Shiny framework IOP Conference Series: Earth and Environmental Science 31 1

[10] Ester M, Kriegel H, Sander J and Xu X 1996 A Density-based algorithm for discovering cluster in large spatial database with noise Proc. 2nd Int Conf. on Knowledge discovery and Data Mining (AAAI Press) p 226-231

[11] Birant D and Kut A 2007 ST-DBSCAN: an algorithm for clustering spatial–temporal data Data & Knowledge Engineering 60 208-221

[12] Chakraborty S and Nagwani N K 2014 Analysis and study of Incremental DBSCAN clustering algorithm International Journal of Enterprise Computing and Business Systems 1 2

[13] Ester M, Kriegel H P, Sander J, Wimmer M and Xu X 1998 Incremental clustering for mining in a data warehousing environment VLDB 98 323-333