THINK INDIA (Quarterly Journal)

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

A Survey On Spatiotemporal Co-Relation Analysis Using Data Mining Approach

R.Sarala[1], Dr.V.Saravanan[2]
[1]Ph.D. Scholar, Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India.
[2]Professor & Head, Department of IT, Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India.
Email: rsarala3011@gmail.com, drvsaravanan2017@gmail.com
Contact No: 9994552286

Abstract: Spatiotemporal data mining is a fastest growing technology for analysing the space and time object. It studies the method of discovering interesting and previously unknown, however doubtless helpful patterns from giant spatiotemporal databases. Spatiotemporal data processing techniques area unit crucial to organizations that create choices supported massive spatial and spatiotemporal datasets. the event of Spatiotemporal knowledge analysis ways will uncover probably fascinating and helpful info. because of the quality of Spatiotemporal knowledge and also the diversity of objectives, variety of Spatiotemporal co-occurring analysis ways exist in terribly low observation. Spatiotemporal co-occurring patterns represent subsets of event sorts that occur along in each area and time. For the correlation of the object this paper presents a survey on Spatiotemporal data processing approach.

Keywords: *Spatiotemporal Data Mining, Correlation analysis, Apriori, Pearson, Expectation Maximization and Survey*

I. INTRODUCTION

Spatiotemporal statistical models are essential tools for acting reasoning and prediction for processes within the physical, environmental, and biological sciences. Such processes are typically sophisticated therein the dependence structure across area and time is non-trivial, typically non-separable and non-stationary in area or time. additionally, it's typically the case that the quantity of spatial locations at that reasoning is desired is kind of giant. what is more, knowledge is typically collected with substantial experimental uncertainty and it's not uncommon to own missing observations at numerous spatial and temporal locations. A spatiotemporal object is outlined as an object that has a minimum of one spatial and one property. The spatial properties are location and geometry of the thing [1]. The property is timestamp or quantity that the thing is valid. The spatiotemporal object sometimes contains spatial, temporal and thematic or non-spatial attributes. samples of such objects area unit moving automotive, fire, and earth quake. Spatiotemporal knowledge sets primarily capture ever-changing values of spatial and thematic attributes over a amount of your time. a happening during a spatiotemporal dataset describes a spatial and temporal development that will happens at an exact time t and placement x. samples of event sorts are earth quake, hurricanes, road traffic congestion and road accidents. In world several of those events move with one another and exhibit spatial and temporal patterns which can facilitate to know the natural phenomenon behind them. Therefore, it's important to spot with efficiency the spatial and temporal options of those events and their relationships from giant spatiotemporal datasets of a given application domain [2]. whereas spatial data processing is superficially thought of because the multi-dimensional equivalent of temporal data processing, in follow the scaling of dimensions yields solely a restricted range of helpful techniques. Thus, the spatial data processing analysis to this point has typically taken the choice path of embedded unambiguously spatial constructs on prime of static techniques; association rules, cluster and characterisation all have their spatial analogues.

The importance of spatial and spatiotemporal data processing is growing with the increasing incidence and importance of enormous datasets like maps, virtual globes, repositories of remote-sensing pictures, the decennial census and collections of

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

trajectories (e.g. gps-tracks). Applications embody atmosphere and Climate (global amendment, land-use classification), Public Health (e.g. watching and predicting unfold of disease), Public Safety (e.g. crime hot spots), security (e.g. common operational picture), Mobile-commerce (e.g. location-based services), etc [3]. Classical data processing techniques typically perform poorly once applied to spatial and spatiotemporal knowledge sets attributable to the numerous reasons. First, these datasets are embedded in continuous area, whereas classical datasets (e.g. transactions) area unit typically separate. Second, patterns are typically native whereas classical data processing techniques typically specialise in world patterns. Finally, one among the common assumptions in classical applied mathematics analysis is that knowledge samples are severally generated. once it involves the analysis of spatial and spatiotemporal knowledge, however, the belief regarding the independence of samples is mostly false as a result of such knowledge tends to be extremely self-correlated. as an example, folks with similar characteristics, occupation and background tend to cluster along within the same neighbourhoods [4]. In spatial statistics this tendency is named autocorrelation. Ignoring autocorrelation once analysing knowledge with spatial and spatiotemporal characteristics might manufacture hypotheses or models that are inaccurate or inconsistent with the info set. Thus, new ways are required to analyses spatial and spatiotemporal knowledge to fascinating, helpful and non-trivial patterns. This speak surveys a number of the new ways together with those for locating interactions (e.g. co-locations, co-occurrences, tele-connections), detection spatial outliers and placement prediction at the side of rising ideas on spatiotemporal pattern mining. analysis accommodating each spatial and temporal data processing is distributed at present; but, 2 directions is known. Firstly, the embedding of a temporal awareness in spatial systems, and second, the accommodation of area into temporal data processing systems [5]. To date, the previous approach has been a lot of well-liked, partly attributable to the relative maturity of geographic info systems and also the convenience of time-stamped snapshots of geographic/spatial take a look at knowledge. massive volumes of spatiotemporal knowledge are progressively collected and studied in numerous domains together with, climate science, social sciences, neurobiology, medical specialty, transportation, mobile health, and Earth sciences. Spatiotemporal knowledge differs from relative knowledge that procedure approaches are developed within the data processing community for multiple decades, therein each spatial and temporal attributes are on the market additionally to the particular measurements/attributes. The presence of those attributes introduces further challenges that has to be restricted. Approaches for mining spatiotemporal knowledge are studied for over a decade within the data processing community. during this article we gift a survey of spatiotemporal data processing correlations and its approaches.

The organization of this paper is as follows. In Section 2 holds a literature review section, it tells about various authors approaches advantages and disadvantages, in Section 3 discussed about spatiotemporal approach, in Section 4 correlation analysis of spatiotemporal mining were provided, in Section 5 application of spatiotemporal mining were discussed and finally Section 6 contains the a conclusion about this paper, its end or result.

II. LITERATURE REVIEW

Karthik Ganesan Pillai et al [6], locate the spatiotemporal co-happening examples speak to subsets of occasion types that happen together in both realities. In contrast with past work in this field, we present a general structure to distinguish spatiotemporal cooccurring examples for persistently developing spatiotemporal occasions that have polygon-like portrayals. We likewise propose a lot of measures to recognize spatiotemporal co-happening designs and propose an Apriori-based spatiotemporal cooccurrence mining calculation to discover pervasive spatiotemporal co-happening designs for broadened spatial portrayals that develop after some time. We assess our system on genuine information to exhibit the adequacy of our measures and the calculation. We present outcomes featuring the significance of our measures in distinguishing spatiotemporal co-event designs.

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

Rashid, et al [7] investigation the spatiotemporal database (STDB) has gotten impressive consideration during the previous couple of years, because of the rise of various applications (e.g., flight control frameworks, climate estimate, mobile figuring, and so on.) that demand proficient administration of moving articles. These applications record items' geographical areas (sometimes additionally shapes) at different timestamps and bolster inquiries that investigate their verifiable and future (prescient) practices. The STDB altogether expands the customary spatial database, which manages just stationary information and subsequently is inapplicable to moving articles, whose dynamic conduct requires re-examination of various subjects including information displaying, records, and the related question calculations. In numerous application territories, immense measures of information are produced, unequivocally or verifiably containing spatial or spatiotemporal information. In any case, the capacity to break down these information stays lacking, and the requirement for adjusted information mining instruments turns into a noteworthy test. In this paper, we have exhibited the difficult issues of spatiotemporal information mining.

Roddick et al [8] discover the development in the extent of datasets, information mining has as of late turned into a significant research theme and is getting generous enthusiasm from both scholarly community and industry. In the meantime, a more prominent acknowledgment of the estimation of worldly and spatial information has been clear and the main papers taking a gander at the intersection of these two zones are beginning to develop. This short paper gives a couple of remarks on this examination and gives a list of sources of applicable research papers exploring fleeting, spatial and spatiotemporal information mining.

GowthamAtluri, et al [9] from the huge volumes of spatiotemporal information are progressively gathered and considered in various spaces, including climate science, social sciences, neuroscience, epidemiology, transportation, mobile health, and Earth sciences. Spatiotemporal information contrast from social information for which computational methodologies are created in the information digging network for different decades in that both spatial and fleeting qualities are accessible notwithstanding the real estimations/traits. The nearness of these qualities acquaints extra difficulties that requirements with be managed. Methodologies for mining spatiotemporal information have been examined for over 10 years in the information mining network. In this article, we present an expansive study of this moderately youthful field of spatiotemporal information mining. We talk about various kinds of spatiotemporal information and the applicable information mining addresses that emerge with regards to investigating each of these datasets. In light of the idea of the information mining issue examined, we arrange writing on spatiotemporal information mining into six noteworthy classes: clustering, prescient learning, change discovery, visit example mining, inconsistency location, and relationship mining. We talk about the different types of spatiotemporal information mining issues in every one of these classes

S. Kisilevich et al [10] Spatiotemporal clustering is a procedure of collection items dependent on their spatial and worldly closeness. It is moderately new subfield of information mining which increased high ubiquity particularly in geographic information sciences because of the inescapability of a wide range of area based or environmental gadgets that record position, time or/and environmental properties of an article or set of items continuously. As a result, various sorts and a lot of spatiotemporal information wound up accessible that acquaint new difficulties with information examination and require novel ways to deal with learning revelation. In this part we focus on the spatiotemporal clustering in geographic space. To start with, we give a characterization of various sorts of spatiotemporal information. At that point, we center around one kind of spatiotemporal clustering - direction clustering, gives a diagram of the cutting-edge methodologies and strategies for spatiotemporal clustering and at last present a few situations in various application areas, for example, development, cell systems and environmental investigations.

III. SPATIOTEMPORAL DATA MINING

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

Time and space are significant parts of all true wonders. Database applications must almost certainly catch time and space shifting in nature of the wonders model. In regular databases, properties containing fleeting or the spatial information are being controlled exclusively by application programs, with little assistance from database the board framework. A spatiotemporal database is a kind of database that supports parts of both time and space. It offers spatial and transient information types in its information model and question language. At the point when the information has relations with spatial information, the term ends up spatial information mining. As it were, spatial information mining is the use of information mining strategy to spatial information. It will track with similar capacities in information mining; with the end target is to discover designs in geology, meteorology and so on. Spatial information mining is the way toward separating certain learning, spatial relations, or different examples that are not expressly put away in spatial databases [11]. As a rule, the spatial relations would not be put away in the database but rather disguised in multi-layer geographic information. To extricate spatial depiction of the geometric highlights, some spatial task must be performed which were engaged with the procedure of spatial information mining. For example, spatial information mining can be utilized in spatial (measurable) examination to enormous datasets, for example, discovering malignant growth groups to find unsafe condition or to discover new spatial and intriguing examples, for example, discover areas that are irregular and so forth. Spatial examples of enthusiasm here may incorporate portrayal of areas of an element (for example wrongdoing) and its relationship with other spatial highlights (for example populace thickness, separation to transportation organize, and so forth).

A spatiotemporal database system manages data whose geometry changes over time.Spatiotemporal data is usually modelled by extending temporal databases or spatial databases. That is, spatiotemporal data is modelled in two ways. First, we can add spatial properties and operations in temporal databases. The second way is to add temporal properties and operations in spatial databases.The field of spatial databases has been a functioning zone of research for more than two decades. There are two normal models of spatial information: field-based (for example raster) and item based (for example vector). The field-based model treats spatial information, for example, height, precipitation and temperature as gathering of spatial capacities changing space segment to a property area. The article-based model treats the information space as though it is populated by discrete, recognizable, spatially referenced elements [12]. A usage of a spatial information model with regards to Object Relational databases comprises of a lot of spatial information types and the tasks on those sorts.

IV. CORRELATION ANALYSIS FOR SPATIOTEMPORAL MINING

Connection is a bivariate examination that estimates the quality of relationship between two factors and the course of the relationship. As far as the quality of relationship, the estimation of the connection coefficient fluctuates among +1 and - 1. An estimation of ± 1 shows an ideal level of relationship between the two factors. As the connection coefficient esteem goes towards 0, the connection between the two factors will be flimsier [13]. The course of the relationship is demonstrated by the indication of the coefficient; a + sign shows a positive relationship and a – sign shows a negative relationship. Here give a portion of the connection examination approach for spatiotemporal information mining.

A. PEARSON METHOD

Connection between the information is a proportion of how well they are connected. The most well-known strategy for connection is Pearson Correlation. Pearson connection coefficient(r) is a proportion of the quality and the straight connection between the two arrangements of information [14]. It is utilized when there are two quantitative factors. The Pearson coefficient esteem lies between the - 1 and +1. It has the positive direct relationship, negative straight connection and no straight connection. For figuring the Pearson connection, we have to pursue the beneath steps:

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

- Make a table with the obtained data for two columns (x, y) and add three columns for (xy), (x2) and (y2).
- Complete the table using manual calculation.
- After completing the table, we need to add all the columns from top to bottom
- Use this formula to find Pearson correlation coefficient

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[\, n\sum x^2 - (\sum x)^2 \,][\, n\sum y^2 - (\sum y)^2 \,]}}$$

Where, n =number of pairs of scores

∑xy =sum of the products of paired scores

∑x =sum of x scores

∑y =sum of y scores

∑x2 =sum of squared x scores

∑y2=sum of squared y scores

For the Pearson r relationship, the two factors ought to be typically appropriated (regularly disseminated factors have a ringer formed bend). Different suppositions incorporate linearity and homoscedasticity. Linearity expect a straight-line connection between every one of the two factors and homoscedasticity accept that information is similarly disseminated about the relapse line.

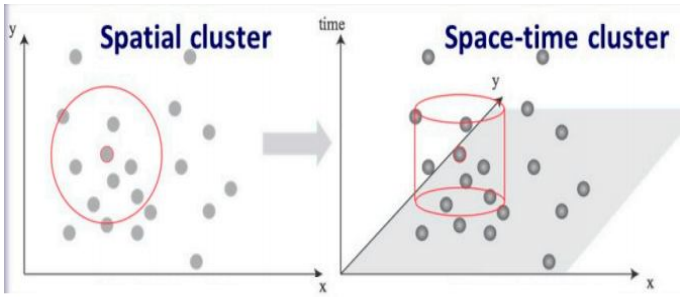

Figure 1: Space and space–time scan window

This technique gives the Space–time check measurements was reached out from space filter insights to recognize correlate with the most noteworthy probability proportion by moving a chamber as an output window to examine Spatiotemporal information. Figure 1 demonstrates the distinction between the two strategies [15].The left graph uses space scan statistics to detect clusters, the red centre is the core point and the larger circle is the scan window for detection. The right graph uses space–time scan statistics to find clusters, it adopts a red cylinder as the scan window.

B. APRIORI

It is applied to a dataset containing a large number of transactions. The algorithm learns associate rules. In data mining, associate rules are techniques for learning relations and correlations among variables in database. Example for Apriori algorithm: Consider the dataset of a supermarket transaction to be a giant spreadsheet. Each row of thespreadsheet is a customer transaction and every column represent a different grocery item. By using Apriori algorithm, we can analyse the items that are purchased together. We can also find the items that are frequently purchased than the other items. Together, these items are called item sets.Apriori firstly introduced by Agrawal, Mannila et al. [16] is an association rule mining algorithm that finds frequent itemset of L in database D using support and confidence. This search is level-wise and it goes up the lattice of item sets. In the first step, the algorithm scans the whole database and eliminates items that are lower than the defined minimum support and then all items are combined together to produce subsets of items for the next scan. For association rule mining using Apriori algorithm we considered spatial and temporal attributes exported form spatiotemporal analysis step as items of the rule mining database. Thus, we used spatial features such as location name and location category. Moreover, temporal features such as day and part of day were also added to the rule mining database.It is a pity that Apriori algorithm is not widely used in data sets involving time and/or space dependent features. However, it is possible to embed time and space features into the data sets and make Apriori algorithm a suitable data mining technique for learning spatiotemporal association rules.

C. EXPECTATION MAXIMIZATION

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

The Expectation Maximization (EM) calculation is an apparatus used to register a most extreme probability estimation of deficient information, with obscure or inactive factors. There are a couple of varieties of this calculation as it is utilized for various unmistakable applications. The EM calculation is a famous instrument in AI and then again is utilized by analysts to tackle issues where certain parameters of information must be assessed by accepting that obscure information [17]. where the issue includes two arrangements of random factors of which one, X, is recognizable and other, Z, is covered up. In more straightforward words calculation works in following two stages:

E-step Estimates the desire for the missing worth for example unlabelled class information. This progression compares to performing characterization of each unlabelled report. Likelihood circulation is determined utilizing current parameter.

M-step Maximizes the probability of the model parameter utilizing the recently registered desire for the missing qualities as though were the genuine ones.

Step 1:Given: X - Labeled data, Z- Missing values, θ - Unknown parameter

$L(\theta;X,Z) = p(X,Z|\theta)$ - likelihood function (probability).

$L(\theta|X) \epsilon \{\alpha\, p(X|\theta): \alpha > 0\}$, $p(X|\theta) * p(Z|\theta)$

Step 2: With the given variables X, Z and θ, the maximum likelihood estimation of the unknown parameters is calculated by the marginal likelihood of the observed data. The value obtained is not tractable.

Finding maximum likelihood $L(\theta;X) = p(X|\theta) = (X,Z|\theta)$

Step 3- caculate expected value of log likelihood function. $Q(\theta|\theta^{(t)}) = [\log L(\theta;X,Z)]$

Step 4- find the parameters that maximizes this quantity $\theta^{(t+1)} = \arg\max Q(\theta|\theta^{(t)})$

The observed data points X, X, might be discrete which are limited or taken from countably endless set or it might be continuous which are taken from uncountably interminable set. The EM calculation can be connected to other information models when one of the parameters Z or θ is known. This demonstrates the iterative idea of calculation that keeps on foreseeing the qualities until joins to a specific and explicit esteem. The motivation to go for EM calculation is that it is basic and straight-forward to actualize, it can iteratively make surmises about missing information and can be advanced for model parameters. The confinement are: EM dependably doesn't locate the ideal parameters and stalls out in nearby optima as opposed to worldwide optima. EM is speedy in the early couple of emphasess, yet slower in the last cycles.

Estimation is completed by means of the desire amplification (EM) calculation or general EM calculation. A few parameterization methodologies are proposed and scientific or computational shut structure EM update conditions are inferred for each. This estimation technique broadly utilized for recognizing bunches from spatial information to find high-thickness huge geographic occasions. Contrast and spatial and space–time filter insights that can just identify round or chamber groups with variable size, these two strategies can distinguish noncircular and non-chamber bunches with high exactness.

V. APPLICATION OF SPATIOTEMPORAL

The significance of spatiotemporal information examination and mining is developing with the expanding accessibility and consciousness of tremendous measure of geographic and spatiotemporal datasets in numerous significant application spaces like

- **Meteorology**: a wide range of climate information, moving tempests, tornados, improvements of high-weight zones, development of precipitation territories, changes in solidifying level, dry spells.
- **Biology**: animal movements, mating behavior, species relocation and extinction.
- **Crop sciences**: collecting, soil quality changes, land use the executives, regular grasshopper pervasion.

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

- **Forestry**: forest growth, forest fires, hydrology patterns, canopy development, planning tree cutting, planning tree planting.
- **Medicine**: patients' cancer developments, supervising developments in embryology.
- **Geophysics**: earthquake histories, volcanic activities and prediction.
- **Ecology**: causal relationships in environmental changes, tracking down pollution incidents.
- **Transportation**: traffic monitoring, control, tracking vehicle movement, traffic planning, vehicle navigation, fuel efficient routes.

In addition to these individual zones, blends of marvels are likewise of intrigue. For instance, what changes in woodlands can be connected to which sort of creature conduct, which climate improvements are in charge of grasshopper invasion. In addition, a few blends present specific arranging difficulties. For instance, extraordinary climate occasions require rerouting of vehicles, planes, and ships. Some other definite application is

Environmental Science: Studying the data pertaining to the quality of air, water, and environment is one of the objectives of environmental science. While air quality is measured based on the presence of pollutants such as particles, carbon monoxide, nitrogen dioxide, sulphur dioxide, ozone etc., water quality is measured based on factors such as dissolved oxygen, conductivity, turbidity, and pH. Air quality sensors are typically placed on streets or on top of buildings, and water quality sensors are placed in lakes, rivers, and streams.

Precision Agriculture: Multi-band high-resolution (ranging from 0.25m to 1m) areal or remote sensing images of large farms are being collected at regular intervals (e.g., daily to weekly). One of the purposes of collecting and studying this data is to detect plant diseases and understand the impact of several factors such as misapplication of fertilizer, compaction during planting and weeds on crop yield, as well as their inter-relationships. With the help of this knowledge, steps can be taken in future crop cycles to mitigate the risks due to the factors that adversely affect the crop yield.

Traffic Dynamics: This data contains information about each trip made by customers of the taxi service, including the time and location of pick-up and drop-off, and GPS locations for each second during the taxi ride. This data can be used to understand how the population in a city moves spatially as a function of time and also the influence of extraneous factors such as traffic and weather. In addition, this data can also be studied to explore traffic dynamics based on the collective movement patterns of the taxis

Heliophysics: Heliophysics studies the events that occur in the Sun and their impact on the Solar System. The publicly available Heliophysics Events Knowledgebaseprovides various observations that include solar events and their annotations on a daily basis. Examples of these events include Active region, Emerging flux, Filament, Flare, Sigmoid, and Sunspot. The time and the location of where these events were observed on the Sun are also provided in the knowledgebase. The spatial and temporal information along with the different observations are studied to discover patterns in the solar events [18].

VI. CONCLUSION

Spatiotemporal data mining can thus be defined as identifying the interesting and non-trivial knowledge from large amounts of spatiotemporal data. Mining information from such datasets is an important problem. In most cases, these databases changes with time, it is therefore important to capture the corelatingbehaviour of the spatial data points with respect to time as these give insights for predicting future occurrences of events. For the analysis of the corelated object this paper presents a survey on data mining approach of corelating the spatial data over time. From this analysis Expectation maximization algorithm provide a better result in case of flood in future because of climate change, water level, etc, it can provide the predicted value over time analysis. And also, this paper discussed about the applications which were used by the spatiotemporal objects.

VII. REFERENCES

ISSN:0971-1260

Vol-22-Issue-4-October-December-2019

[1]. Hui Yang and SrinivasanParthasarathy. Mining Spatial and Spatiotemporal Patterns in Scientific Data. Proceedings of the 22nd International Conference on Data Engineering Workshops (ICDEW'06) - Volume 00:146, 2006.

[2]. ShashiShekhar, Michael R. Evans, James M. Kang and Pradeep Mohan, "Identifying patterns in spatial information: a survey of methods", Volume 1, June 2011, wires.wiley.com/widm

[3]. Abhinaya Mohan, "A New Spatiotemporal Data Mining Method and its Application to Reservoir System Operation", Spring 4-25-2014

[4]. Abdulsamet HASILOGLU, "The Use of Spatiotemporal Data Mining for Detection and Interpretation of Trajectory Outliers in Health Care Services", ECJSE, Vol: 4, No: 3, 2017

[5]. ShashiShekhar, "Spatiotemporal Data Mining: A Computational Perspective", ISPRS International Journal of Geo-Information, vol 4, 2015.

[6]. Pillai, KarthikGanesan, Rafal A. Angryk, Juan M. Banda, Michael A. Schuh, and Tim Wylie. "Spatiotemporal co-occurrence pattern mining in data sets with evolving regions." In *2012 IEEE 12th International Conference on Data Mining Workshops*, pp. 805-812. IEEE, 2012.

[7]. Rashid, A. N. M. B., and Md Anwar Hossain. "Challenging issues of spatiotemporal data mining." *Computer Engineering and Intelligent Systems* 3, no. 4 (2012): 55-63.

[8]. Roddick, John F., and Myra Spiliopoulou. "A bibliography of temporal, spatial and spatiotemporal data mining research." *ACM SIGKDD Explorations Newsletter* 1, no. 1 (1999): 34-38.

[9]. Atluri, Gowtham, AnujKarpatne, and Vipin Kumar. "Spatiotemporal data mining: A survey of problems and methods." *ACM Computing Surveys (CSUR)* 51, no. 4 (2018): 83.

[10]. S. Kisilevich et al, "Spatiotemporal Clustering: a Survey", Springer September 2015

[11]. K.VenkateswaraRao, "Spatiotemporal Data Mining:Issues, Tasks And Applications", IJCSES, Vol.3, No.1, February 2012

[12]. R. Geetha, N. Sumathi &Dr. S. Sathiyabama, "A Survey Of Spatial, Temporal And Spatiotemporal Data Mining", Journal of Computer Applications, Vol – 1, No.4, Oct – Dec 2008

[13]. KatruAkhila, "Survey on Data Analysis Using Correlation Methods", IRJET,Volume: 04 Issue: 04 | Apr -2017.

[14]. Zhicheng Shi and Lilian S.C. Pun-Cheng, "Spatiotemporal Data Clustering: A Survey of Methods", MDPI, 2018

[15]. https://www.statisticssolutions.com/correlation-pearson-kendall-spearman/

[16]. Agrawal, R., H. Mannila, R. Srikant, H. Toivonen and A. I. Verkamo (1996). "Fast Discovery of Association Rules." Advances in knowledge discovery and data mining 12(1): 307- 328.

[17]. BhawnaNiga, "Document Classification Using Expectation Maximization with Semi Supervised Learning", International Journal on Soft Computing (IJSC) Vol.2, No.4, November 2011

[18]. Abhinaya Mohan and Peter Z. Revesz. A new spatiotemporal data mining method and its application to the north platte river reservoir system. submitted, International Database Engineering and Applications Symposium (IDEAS) 2014.