IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

# Spatio-Temporal Neural Network Approach for Location Prediction: State-of-the-Art, Challenges and Future Directions

**N. Venkata Subba Reddy[1*], D. S. R. Murthy[2]**

[1] Department of Computer Science, Sreenidhi Institute of Sc. & Tech., JNTUH, Hyderabad, India
[2] Department of Computer Science, Geethanjali College of Engg. & Tech., JNTUH, Hyderabad, India

*[*]Corresponding Author: nvsreddy@sreenidhi.edu.in, Tel.: +91-6303029125*

DOI: https://doi.org/10.26438/ijcse/v7i5.10571067 | Available online at: www.ijcseonline.org

Accepted: 16/May/2019, Published: 31/May/2019

***Abstract—*** Nowadays, the huge set of spatio-temporal data (STD)are increasingly collected and utilized in different domains that include social sciences, epidemiology, mobile health, climate science, neuroscience, transportation and Earth sciences. Compared to relational data, the STD is different for that the researcher developed computational techniques in the data mining community. The process of extracting implicit knowledge and unknown information, structures in spatio-temporal (ST)dataset , patterns that are not explicitly stored and spatio-temporal relationships is called Spatio-temporal data mining (STDM). As one of information mining procedures, data prediction approach is widely utilized toward forecast the unknown future on the basis of hidden patterns in the past and current data. To obtain ST forecasting, few of them developed analysis tools like spatial information are prolonged to temporal dimension (TD)as well as the time series extended to the spatial dimension (SD) or joined linearly as a ST combination. But, such kind of linear combination of TD and SD is a generalization of difficult ST relations which is present in complex geographical occurrences. The present study, reviews the traditional STDM approach, tools, pattern mining approach and data analysis. On the basis of data mining issues, this literature has been classified into four main categories: trajectory mining approach, clustering, pattern mining, predictive learning and location prediction. We discourse the different forms of STDM issues in each of these groups.

***Keywords* :** Spatio-Temporal Data, Trajectory Clustering and Mining Approach, Location Prediction, Trajectory Pattern Analysis

## I. INTRODUCTION

At the present scenario, there is a rapid usage of Global Positioning System (GPS) devices,smart-phones, wireless communication technology, Internet of Things (IoTs)(Rathore et al., 2018b), it's conceivable to follow practices of moving items over the world. Due to increase of location-acquisitionappraoch that are occasioneda huge set of STD, particularly in the form of trajectories.This kind of information comprises a excessive deal of information (Zheng et al., 2008) that is applicable to most of the location-based services (LBSs)as well as applicable to different application like vehicle navigation, location-based recommendations, traffic management,etc(Rathore et al., 2018a). Due to the necessity of well-organized system process, there is a need of LBSs that can perform to forecast the user activities at the next location to visit. Hence, the location prediction approach for LBSs aiming on desired mobile users. Recently, the location prediction approach have emerged for mining user behaviour which is collected from mobile usersGPS trajectories towards predict the user next move over the world (Ying et al., 2011;Monreale et al., 2009; Lei et al., 2011).

The location-based prediction method usually determined by pre-collected or real-time-dependent, provided a series of locations will deduce the following area where the object is well on the way to go.Because of the progression of time and space, trajectory isn't reasonable to be straightforwardly imported to a prediction model .Beforehand utilizingforecasting model, entire tr ajectory point is pre-processed toward converting the real continuous values which are linked to latitude and longitude of geospatial coordinates into discrete codes related to precise regions. For this, most of the researcher suggested prediction approach that is simply partitioned trajectories into cells or clustering trajectories into stay points or frequent regions. In order towards finding the recurrent pattern with clusters, the pattern mining approach suggested where the trajectories are clusters with discrete codes or transformed into grids. But, the conventional cell-based or cluster-based approach ignore the trajectorydata amongst the clustered result, that might comprise precarious data for particular applications (Wu et al., 2017).In addition, the conventional data mining approach concentrated on determining cluster from non-temporal data and non-spatial data that is unfeasible for ST information. Also, discover of pattern from ST information is highly complicated (Aryal & Sujing Wang, 2017). Hence the present study reviews the traditional approach, identifies the issues and suggests a unique idea for overcoming the issues of handling STD.

The rest of this paper is pre-arranged into five sub-setions. These are discussed as follows. Section 2 discusses the brief summary of previous studies which includes studies related on data mining, trajectory mining, clustering and pattern mining as well as identified the gap of these researches.Section 3 presents the material and methods of literature papers were performed in databases. Section 4 presents the detailed discussion of analysis, demonstrated results of empirical performance and summary of previous studies in a tabular format. Section 5 discusses our concluding remarks and future scope.

## II. SPATIO AND SPATIOTEMPORAL DATA

One imperative part of spatiotemporal information mining is process of mining input feature set.The present section gives scientific categorization of various spatial and ST information sorts. Additionally condenses their remarkable information characteristics and relationships.

### *A. Types of Spatial and Spatiotemporal Data*

The proce ss and analysis of STDM task are more difficult when compared to the classical data science task since its incorporate discrete depictions of continuous time and space. Additionally, various types of spatial and STD are discussed in table 1. It can be classified into three types are field model, object model and spatial network model (Shashi & Chawla, 2007; Worboys & Duckham, 2004). Based on the temporal data, the spatio-temporal information further classified as temporal change,temporal snapshot and processor event model (Li et al, 2008; Shekar et,al,2015).

Table 1: Taxonomy of Spatial and Spatiotemporal Data ModelsShekhar et al (2015).

| ***Spatial Data*** | | ***Temporal Snapshots (Time Series)*** | ***Events/Processes*** | ***Temporal Change (Delta/Derivative)*** |
|---|---|---|---|---|
| Field Model | Irregularand regular | Raster time series | Cellular computerization | Variation across raster Snapshots |
| Object Model | Line(s) | Line trajectories | Line procedure | Extension /motion/rotation, Merge / split and deformation |
| | Point(s) | 1. Spatial time series 2.Point trajectories | Spatial/spatiotemporal Point procedure: Poisson, Cox, or Cluster procedure | Motion /displacement (for example., random Walk, Brownian motion), speed/acceleration |
| | Polygon(s) | Polygon trajectories | Flat procedure | Extension /motion/ rotation/ Distortion, split/merge |
| Spatial Network Model | Graph | Spatiotemporal network: 1. Time aggregated graph, time expanded graph, 2. Network flow | 1. Process on spatial network or spatiotemporal event; 2. Random geometric Graph; 3. Percolation theory | Removal or addition of Nodes and edges |

Table 2: Relationship Amongst Spatial and Non-Spatial Attributes

| ***Attributes*** | ***Relationships*** | ***Categories*** |
|---|---|---|

 

| | | |
|---|---|---|
| Spatial | • Often implicit<br>• Set space: intersection ,union,membership, and so on.<br>• Metric space: metric: area ,distance,perimeter.<br>• Others: shape based and visibility<br>• Topological space: within ,meet,overlap, etc.<br>• Directional: above, below, northeastern | • Area<br>• Shape<br>• Location<br>• Perimeter |
| Non-spatial | • Explicit<br>• Ordering<br>• Subclass of<br>• Arithmetic<br>• Instance of | • Ordinal<br>• Ratio<br>• Nominal<br>• Interval |

In STD, the data attributes are further classified into three types namely, spatial attributes, non-ST attributes and temporal attributes. The statial attributes are used to define spatial location (e.g., latitude and longitude),spatial extent (e.g., area, perimeter) (Bolstad, 2005; Ganguly & Steinhaeuser, 2008),defined elevation in a spatial reference frame and shape features. The non-spatial attibutes are utilied toward characterize non-contextual features of objects like population, name, ratio of un-employment in city, etc. The same attributes are utilized for the classical data mining process(Tan et al., 2006).The temporal attributes are comprisesa raster layer, duration of a process, timestamp of a spatial object and spatial network snapshot. The telationsof these data attributes are discussed in table 2.

*B. Issues and Challenges*

With respect to the data analysis, processing and mining of ST information, the general issues and challenges are discussed as follows(Rao et al., 2012):

1. The remarkable features of STdatasets are that they convey remove and topological data that need geometric and transient calculation.

2. The spatio-temporal relations like topology, distance, before and afterare data bearing and direction. In order to analyze these data, there is a need of effectivespatiotemporal data mining and analysis approach.

3. The spatial and temporal data are not explicitly encoded in a database where its implicitly defined. The relational details should be extracted from these information.Before the actual mining process starts, there is a trade-off amongst pre-processing them and computing on the fly as while itsrequired.

4. The distinctive feature of spatio-temporal data needs substantial change of minig approach hence able toabuse annoying temporal and spatial relations as well as embedded the patterns in the given dataset.

5. In the qualitative reasoning process, many rules are required for the analysis of temporal and spatial data that also depents on the domain independent knowledge needs to be consider while generating different patters. There isa issue like how to integrate them and how to express rules with spatiotemporal reasoning mechanism.

6. The scalability of data mining approach, efficiently index spatiotemporal datasets, visualization of ST phenomena and patternwhere as the data structure representation is challenging problem. So needs to be resolved in future.

7. To obtain insight of essentialoccurrenceswhich is signifiedthrough the known information, the development of effective approach for data visualization of STD and interaction facilities are challenging issues. This requires the outcome of ST information mining is installed inside a procedure that deciphers the outcomes for further appropriately organized examination reason behind the outcomes.

8. The implementation of efficient visual interfaces for review and controlling the temporaland geometrical traits of spatiotemporal information is an challenging issues. For solving the different issue most of the researcher suggested algorithm and unique technique. These are discussed in the following section.

## *III. PROBLEMS AND METHODS*

In earlier, most of the researcher addressed the issues of trajectory prediction that comprises the issues of short-term prediction like next location prediction, a long-term prediction like a prediction of future location. These approaches primarlily focused on determiningrecurrentpatternsthrough utilizing different data mining approach. Most of the suggested approach is combined in nature and categorised into different categories such as data mining approach, clustering technique and rule-based learning approach. These are discussed as follows:

*A. Trajectory Mining Approach*

Most of the researcher suggested trajectory mining technique for managing, mining and processing of trajectory information.The mining approaches arelocation-based recommendations(Son et al., 2013; Bao et al., 2013), mining trajectory data for road map inference (Li et al., 2015), parallel pattern mining of massive trajectory data (Qiao et al., 2010), route planning (Kurashima et al., 2010; Chen et al., 2011) and discovering movement patterns from animal trajectory data(Wang et al., 2014; Technitis et al., 2015). Few of them suggested an approach for mining of personal moving object trajectories that comprising destination prediction(Lei et al., 2012), public transit forecasting(Zhang et al., 2016; Ni et al., 2016), personal semantic places discovery(Lv et al., 2012), privacy protection(Zheng, 2015) and trajectory prediction(Liu et al., 2016b). In the field of environment monitoring, smart city and data security, the data mining based trajectories plays a major role and applied to real as well as hug application(Zheng, 2015). In such applications, one of key operation is position and or location prediction. For this, most of the researcher used mining based trajectory patterns along with historical location points for position prediction. These are categorised into two types such as geometry and road network based methods.By the use of route segments and tree-shaped data structure, a fixed road network is separated into atomic roads indexed in a road network based approach that is supported to precise querying future locations and current locations in large-scale moving objects (Qiao et al., 2015; de Almeida & Güting, 2005).At that point of trajectories, an area stayedthrough an object is isolated into cells which is also represented as cell arrangements in geometry-based techniques. Based on the historical directions, the trajectory position are anticipated and built a prediction tree with districts as hubs to represents a trajectory patterns(Lei et al., 2012).In order to realize position predictions, most of the researcher utilized spatial data. But the we know, the trajectory information has both temporal and spatial data. The TDs are closely linked with the trajectories of an object. On the other hand, moving object behaviour depends on the spatio-temporal regularity features. Hence, the trajectory data plays a significant role with respect to time dimension in the position prediction and or location prediction issues(Li et al., 2018). So there is a require to consider three aspects in future. First, at a similar time, the object desires to select a similar trajectory. Second, velocity of data that are used to record the trajectory data which provide precise result for moving object position predition than the previously available logged information. Third, trajectory periodicity denotes the object trajectory incline to a same time period as well as same with each other, for example, weekends, a different trajectory with different time period and weekdays.

Few of them suggested feature extraction approach where the trajectory data has been partitioned into self-defined grids or cells (Mathew et al., 2012; Ying et al., 2014). In the global area, the cell-based approach is problematic towardpull a grid and straightforwardlypretentiousvia granularity of the grid network.Some of the researchers presented a model for location prediction. A research by Lei et al.(2011) presented a ST trajectory framework based on probabilistic suffix tree captured the movement behaviour objects. With respect to the temporal and spatial patterns of movement behaviours, the accurateness of suggested method is high. But the study utilized the synthetic dataset that is generated from real dataset.

A study by Monreale et al.(2009) presented a Trajectory pattern (T-pattern) Tree based on decision tree approach for extracting the frequent movement patterns and based on the best matching functions, the next location of a new trajectory is predicted. But it's computationally expensive on the basis of mining of frequent trajectory patterns.A research by Li et al. (2010)presented two kinds of trajectory patterns namely swarm pattern and periodic behaviour pattern approach. Furthermore, they used Fourier- algorithm is utilized to detect period and trajectories are clustered into reference spots. By hierarchical clustering method, they mined the periodic patterns. In order to cluster the frequent places, they utilized core step of pattern-based prediction approach. But it limited to the position devices, track points that lost when satellite signal is low.

A study byLi et al. (2018)suggested a Spatio-Temporal Regularitybased Prediction (STRP) algorithm towards analysing the temporal and spatial regularity of object mobility. In this regards, they have obtained the candidate next positions by extraction of personal trajectory patterns from moving object ancient trajectory information. According to the current time and time components of patterns as well as scored entire candidate next positions using given set of data. On the basis of two different public trajectory input data set, the performance has been validated and its seems better accuracy than the traditional method. In specific to this, the Markov-based algorithm (43.9%), the demonstrated results give a better result with respect to predictionaccuracy is 86.8%. However, the study needs to consider, minimize the execution or retrieval time ofonline prediction and needs to extract the trajectory patterns of moving objects. Also, to improvise the prediction accuracywith respect tomore temporal characters, this is concerned in trajectory data.

*B. Datamining Clustering Algorithm*

 

From the review of literature, a research byCao et al. (2007)presented a data mining based clustering approach for deterimining the dense clusters as valid regions(Ester et al., 1996) and to speed up the algorithm by hash-based approach. A research by Li et al.(2012) utilized Kernel approach (Worton, 1989) towards computingdensities toward discovering dense regions and/or reference spots which is regularly visited through moving objects. But, they did not assume the inherently hierarchical nature of spatio-temporal occurrences. Addition, they found reference spots by trajectory nodesfor periodic patterns whereas disregarding objects sequence in trajectory clustering. The sequence of objects plays a vital role where the trajectory data are spatio-temporal(Zhang et al., 2018).A study by Liu et al. (2017)enhanced the conventional RNN method along with temporal and geographicalcontexts toward handling the prediction issues of spatio-temporal information. The suggested method performs better than the short location sequences like stay points and clusters. But, while assuming the entire trajectory data, the precision of prediction of RNN may decline and the location sequences become longer.

A study bySantipantakis et al. (2018)presented a data analytic task which obtainable a unified depiction of datafor semantic integration of big mobility information. Furthermore, they have extracted the features namely ST link discovery betweenSTobjects in varied data bases, moving entities' trajectories, the transformation of data from huge setinformation as well as heterogeneous sources in RDF. The simulated results (using Apache Kafka and Flink) shows that the suggested method outer performs than the conventional approach with respect to the scalability and efficiency by real-life datasets. In future, need to enhance the performance byconsidereing different data analytic process like trajectory prediction and clustering approach.

A research by Zhang et al.(2018)suggested a trajectory clustering algorithm that deliberates the semantic STD like time-based on Traclus, direction and speed. They had done the experiments with three clustering approach Grid-based, Kernel function and traclus. Furthermore, they have generated a hierarchy of reference spots by considering a single linkage with hierarchical clustering technique. The simulated results show that the presented algorithm outperforms than the conventional reference spot detection method. But the limitation of this approach needs slightly more time than the single-level technique. Then generates periodic patterns and more reference spots which as unable to detect. So there is a need forample experiments with more complex datasets, and there isan extension to multi-level time periods.

A work byRathore et al. (2018a)suggested a hybrid method which is a combination of Markov chain and scalable clustering approach namely Traj-clusivat-based TP for handling both long-term trajectory prediction and short-term trajectory prediction. Also, handles the overlapping trajectories in dense road networks. By the suggested technique, they have determined the number of clusters that represent various movement activities in trajectory data. The demonstrated results show that the suggested approach works better than the conventional technique with respect to long-term and short term prediction performance, distance error andprediction accuracyusing vehicle trajectory datasets. In future, they planned to do a research toward update clusters in real-time using incremental/decremental VAT technique. Also, will improve the prediction performance by considering factors are time, speed and user data.

Despite many issues, knowledge discovery such as extracting meaningful pattern (unknown and implicit knowledge), structures, or pattern and relationship from large data sets is an emerging field of research (Miller & Han, 2009). Many previous studies have focused on efficient data mining techniques such as Association rule mining (ARM) (Mennis & Liu, 2005), space-time autoregressive integrated moving average model (Cheng & Adepeju, 2014), spatial panel data model (Elhorst, 2003), Bayesian Hierarchical model (Subba Rao, 2012), external drift kriging (Ignaccolo et al., 2014), space-time geostatistics, image mining (Umamaheshwaran et al., 2007). The data forecasting is one of data mining approach toward predicting the different patterns of unknown future for both past and current information.

*C. Spatial-Temporal Pattern Mining Algorithm*

For analysing the patterns in mobile user movement data sets, a research by Lu and Tseng(2009),Monreale et al(2009) suggested mobile sequential pattern mining and general-pattern-based prediction approach. But, they incline toward predicting prevalent locations where most of them visited, which is leading to the imbalanced data issues(Yen et al., 2006). In addition, the suggested prediction approach only predicts the pattern but does not considered the anticipated movement of prefix pattern is completely match with the given data. A research by Shanshan et al. (2015)presented ana-priori pattern mining algorithm for analysing the ST relationships. In each time sequences, they have shifted one of the time sequences for rule generation towards generates high-frequency candidates. As a result, the generated rules are disclosing the presence of pollutants along with delays in various locations. But the suggested approach needs frequently execute the rule generation procedure by various combinations that are a time-consuming process.

Some of the researchers suggested mining spatio-temporal periodic patterns toward analysing the spatio-temporal trajectory information (Li et al., 2012). In order to comprehend the behaviours of moving objects, a research byLi et al.(2012) presented valuable periodic patterns and discovered the hidden pattern. However, due to the unique characteristics of spatio-temporal trajectories, the conventional Periodic Pattern Mining (PPM) approach is not directly applied to discover periodic patterns. Specifically, the hierarchical nature of spatio-temporal phenomena, uncertainty of spatio-temporal trajectories and irregularities of time intervals(Sun et al., 2018). It's highly impossible to visits objects of the same location at each and every time period. For solving this issues, a research by Li et al. (2012) suggested clustering approach where clustered reference spots towards replace the objectsprecise locations. Consequently, not at all like in PPM for time-arrangement information, finding fascinating spots for spatio-temporal occasional examples is of incredible significance in spatio-temporal PPM(Zhang et al., 2018).

*D. Machine Learning Approach*

There are two complications while forming STD arises from their temporal lags and size. Few of the researcher suggested a deep learning approach, particularly for dynamic data. For a sequence to sequence prediction, complex sequences classification and generation, most of the researcher suggested recurrent neural networks (RNN) technique (Bengio, 2010; Zhou et al., 2016). But this suggested approach rarely assumed the spatial structure. To overcome this, few of the researcher suggested video pixel networks(N. Kalchbrenner, A. V. D. Oord, K. Simonyan, I. Danihelka, O. Vinyals, A. Graves, 2017)and convolutional RNN(Xingjian Shi, Zhourong Chen, Hao Wang, 2002)has handled bothtemporality and spatiality for video applications.

Few of the researcher suggested deep learning techniques like long short-term memory (LSTM),inverse reinforcement learning (IRL) and recurrent neural networks (rnns)for modelling vehicle trajectories(Zhu et al., 2018; Altche & de La Fortelle, 2017; Bock et al., 2017). Among this, still, they considered on first-order Markov statement for heterogeneous destinations by modelling the routing decisions. Or else, they are too shallow that creates the modelling pattern diversities that agonize from too few parameters.

## IV. MATERIALS AND METHODS

A search of literature papers was performed in databases such as IEEE Xplore, Google Scholar, and other renowned international journals on STDM Methods. The year range for the acquisition of research articles was limited within 2014-2019. In this research, the review articles by researchers and reference sections of the individual articles were manually searched. Also, some research studies which reported well-defined different mining methods and algorithms for the analysis ofspatio-temporal informationwere included. As seen in this review, around 86 suitable records are retrieved which provides relevant results about location-based applicationsof mining and clustering method. From these, we excluded studies that met any of the following criteria: (1) Review papers; (2) Market research information; and (3) Not written in English wherein the present research may have missed systems that are discussed in other languages.

*A. Inclusion and Exclusion Criteria*

The presence and rejection criteria were applied for all the retrieval studies which are listed in the above database.The criteria used for including/excluding papers are as follows:

- Peer-reviewed articles –included keynotes, editorials, reviews, tutorial summaries, position papers and panel discussions, excluded since they are more presentational in nature
- Studies related to market and data research using various approaches such as experimental research, surveys and case studies are excluded.
- Papers with Algorithms and empirical data to be given preference
- Augmented articles- If two articles from similarexamination on a similar subject were distributed in various data presentation scenarios (e.g., Journal and conference), only the journal article was incorporated, since conference data are not very data specific and more informational in nature.
- All copied reviews initiated from various sources were identified and removed

*B. Search String*

We framed our search string via the rules as shown below. We chose to incorporate a condition for picking exact reviews in our search string amid the survey procedure. Given the assortment of research techniques and search Engines, having that condition could have made search string mind-boggling. Running a pilot check for considering the papers, we were mindful of the fact that, we exploited the final search string as obtainable as follows, which is applicable to any Online Search Engine:

"Trajectory Pattern analysis" AND "trajectory clustering" AND "trajectory mining" OR ("STD" OR "trajectory data" OR "Temporal and Spatial Factor" OR "data mining" OR "location prediction" OR "ST patterns" OR" Data Sets")

But, the search was restricted mainly to IEEE Xplore and SCI-Hub, both of which are recognized worldwide for comprehensive public access to research papers.

*C. Study Selection*

Our Search in IEEE Xplore Digital library returned 273 outcomes. We separated the papers by evaluating the abstract, title of the research and accurate data available in the paper.

At the point when there were a few articles may possibly not settle on by perusing the topics and abstract, and hence these articles were held for the following round of assessment. We barred the articles that remained unrelated, or whose full content was not assessable or accessible online. Since we were keen on experimental reviews, we rejected papers that were not populated with observational data. Besides, we included the articles related to group data sharing as both static and dynamic manner. Out of the 273 search results, around 70 were found to be suitable for further research. At that point, the references of these 70 selected articles were checked, keeping in mind the end goal to discover more potential essential research papers, from the references too. We found 37 conceivably significant articles by title from the references of these 70 articles. Also, we have avoided duplicate research articles. At the end of the progression, it was found that 18 articles encountered all the insertionnorms.

## V. DISCUSSION

From the above points, the usage of temporal and spatial data are an excitingand rapidly advancing field like Biology, Forestry, Geophysics, Transportation, Biology, Meteorology and so on.This study aims at carrying together practitioners and researchers of spatial, spatio-temporal and temporal data mining process.For anlyzing, data, some of the researchersuggested feature extraction approach where the trajectory data has been partitioned into self-defined grids or cells (Mathew et al., 2012; Ying et al., 2014). But the cell-based approach is difficult to draw a grid in the global area and easily affected through the granularity of the grid. On the other hand, a research by Monreale et al.(2009) suggested T-patternapproach for extracting the frequent movement patterns and based on the best matching functions, the next location of a new trajectory is predicted. But it's computationally expensive on the basis of mining of frequent trajectory patterns. Similarly, a research by Li et al. (2010)suggested swarm pattern and periodic behaviour pattern approach. But it limited to the position devices, track points that lost when satellite signal is low.

Some of the researcher suggested machine learning technique to process the STD via RNN(Bengio, 2010; Zhou et al., 2016), deep learning techniques(Zhu et al., 2018; Altche & de La Fortelle, 2017; Bock et al., 2017), SVMWang et al. (2007), etc. But this suggested RNNapproach rarely assumed the spatial structure. To overcome this, few of the researcher suggested video pixel networks(N. Kalchbrenner, A. V. D. Oord, K. Simonyan, I. Danihelka, O. Vinyals, A. Graves, 2017)and convolutional RNN(Xingjian Shi, Zhourong Chen, Hao Wang, 2002)has handled bothtemporality and spatiality for video applications. A study by Liu et al. (2017)enhanced the conventional RNN method along with temporal and geographicalcontexts toward handling the prediction issues

mining and process of those of spatio-temporal information. But, while assuming the entire trajectory data, the precision of prediction of RNN may decline and the location sequences become longer.

Among this, most of the researcher consideredfirst-order Markov statement for heterogeneous destinations by modelling the routing decisions. Hence there is a need to enhance the performance for analysis and knowledge extraction process of spatial data. A research by Zhang et al.(2018)suggested a trajectory clustering algorithm that deliberates the semantic STD like time-based on Traclus, direction and speed.But the limitation of this approach needs slightly more time than the single-level technique. Then generates periodic patterns and more reference spots which as unable to detect. So there is a need for ample experiments with more complex datasets, and there isan extension to multi-level time periods.Similarly a research by Shanshan et al. (2015)presented an a-priori pattern mining algorithm for analysing the ST relationships. But the suggested approach needs frequently execute the rule generation procedure by various combinations that are a time-consuming process. For a sequence to sequence prediction, complex sequences classification and generation, most of the researcher suggested recurrent neural networks (RNN) technique (Bengio, 2010; Zhou et al., 2016). But this suggested approach rarely assumed the spatial structure. The summary of previous studies are discussed in table 3 and 4.

Table 3: Comparative study

| ***Author (Year)*** | ***Title*** | ***Research method*** | ***Data set*** | ***Important findings*** | ***Conclusion*** | ***Recommendation*** |
|---|---|---|---|---|---|---|
| Santipantakis et al. (2018) | SPARTAN: Semantic integration of big STD from streaming and archival sources | The spatio-temporal link discovery module. | E two maritime datasets (NARI and IMISG). | Increased the predictability of moving oobjects' trajectories and events | Provided the end-to-end solution to the problem of providing enriched streams of mobility data. | Needs to study in depth how the enriched data can improve the quality of different data analysis tasks, like trajectory clustering and trajectory prediction. |
| Li et al.(2018) | Position Prediction | Spatio-Temporal | GPS trajectory dataset. | Realized position prediction utilizing | Obtain high prediction | 1. To reduce the time consuming of online |

| | | | | | | |
|---|---|---|---|---|---|---|
| | System Based on Spatio-Temporal Regularity of Object Mobility. | Regularity based Prediction (STRP) algorithm. | | TD information implicated in trajectory data adequately. | accuracy. | prediction.<br><br>2. Needs to improve the prediction accuracy. |
| Rathore et al., (2018a) | A Scalable Framework for Trajectory Prediction. | Scalable clustering and Markov chain based hybrid framework. | Vehicle trajectory datasets. | Improved the prediction performance by o best-matching cluster group. | Obtain better results with respect to prediction accuracy and distance error. | 1. In future planned to focus incremental/decremental VAT method to update clusters in real-time.<br><br>2. Also will incorporate more performance metric to check the validity of proposed method. |
| Ying et al.(2012) | Semantic Trajectory Mining for Location Prediction. | Cluster-based prediction strategy. | Semantic trajectory dataset. | Abel to predict the next location of a mobile user. | 1. Improved the prediction precision value<br>2. Highly efficient for prediction of location. | Need to enhance the quality of location predictions in location-based services. |
| Catlett et al.(2018) | A Data-driven Approach for Spatio-Temporal Crime Predictions in Smart Cities. | Spatial analysis and auto-regressive models. | Crime dataset | Automatically detects high-risk crime regions in urban areas as well as reliably forecast crime trends in each region. | Obtain good accuracy in spatial and temporal crime forecasting over rolling time horizons. | Planned to apply same approach forspatio-temporal prediction of other kind of events, different than crimes. |
| Soh et al. (2017) | ST Pattern Analysis and Prediction of Air Quality in Taiwan. | Dynamic Time Warping approach. | Taiwan dataset (Air Quality) | Analysed the temporal similarity between stations. | Eliminated the global effect on a single station through the performance of multiple stations. | 1. However, needs to focus on shorten the training time and improve accuracy.<br><br>2. Error due to the missing data. |
| Soh et al., (2018) | Adaptive Deep Learning-based Air Quality Prediction Model Using the Most Relevant Spatial-Temporal Relations. | ST deep neural network | Location-based Taiwan dataset andBeijing dataset. | Extracted the temporal delay factor from surrounding target features by learning spatial information. | More useful for longer time frame predictions. | 1. To emphasize air pollution propagation delay effects by<br>2. To detect air pollution .sources, including domestic and trans boundary pollution<br>3. Reduce pollution sources. |
| Wang et al. (2007) | Nonlinear Integration of Spatial and Temporal Forecasting By Support Vector Machines. | Support Vector Machines. | Annual air temperature at 194 meteorological stations provided by national meteorological center of P. R. China. | Better forecasting accuracy<br>Obtained time series model for temporal forecasting. | Obtained better forecasting accuracy than a linear combination. | Needs to address spatio-temporal forecasting involving multiple variables. |
| Wu et al., (2017) | A Spatial-Temporal-Semantic Neural Network Algorithm for Location Prediction on Moving Objects | Semantic neural network algorithm. | Stable and higher prediction accuracy. | Able to transform trajectory into model-friendly sequences.. | Stable and higher prediction accuracy. | 1. In future, algorithm should be robust to trajectory of poor quality.<br>2. More data preprocessing required.<br>3. Enhance the framework toaccept more features and dimensions. |
| Aryal andSujing Wang(2017) | Discovery of Patterns in Spatio-Temporal Data Using Clustering Techniques. | Density-based STclustering algorithm and Shared Nearest Neighbor clustering. | TLC Trip Record Data. | Find clusters of different shapes, sizes, and densities. | Automatically determine the number of clusters. | Their ongoing work focuses on developing a ST data mining framework, for big STD by utilizing high-performance computing resources, such as Hadoop-GIS. |

Table 4:Summary of Location Prediction Approach

| ***TechniqueWu et al., (2017)*** | ***Author (Year)*** | *Model Type* | ***Data Type*** | ***Specialities*** |
|---|---|---|---|---|
| Location Recommender | Xiong et al.(2013) | Recommender/ simple model | Event data | Tensor Factorization |
| | Bahadori et al.(2014) | | Check-in data | STTensor Factorization |

 

| | | | | | |
|---|---|---|---|---|---|
| | | Zheng et al.,(2010) | | Trajectory | Spatial Factorization |
| | | Zhuang et al. (2011) | | Event data | Geo-location, time |
| Movement Pattern Mining | STpatterns | Li et al.(2010) | Pattern recognition/ Unsupervised learning | Trajectory | Period and swarm pattern |
| | | Giannotti et al. (2006) | | Synthetic sequence | Travel time-aware |
| | Mobile patterns | Morzy(2007) | Simple model | Sensor data | Prefixspan |
| | | Jeung(2008),Yavaş et al. (2005) | | Trajectory | Motion functions, a priori algorithm |
| Prediction Model | Neighbourhood-based model | Lathia et al. (2009) | Unsupervised learning | Event data | Time-aware neighbourhood |
| | Semantic-based model | Ying et al. (2014) | Pattern recognition/ Unsupervised learning | Trajectory | ST semantic, pattern tree |
| | | Bogorny et al. (2009) | | Trajectory | Hierarchical cluster with semantic |
| | Markov Chain model | Jeung et al. (2007) | Classification/ Supervised learning | Trajectory | Cell partition, Hidden Markov model |
| | | Mathew et al. (2012) | | Trajectory | Trajectory clustering, Hidden Markov model |
| | Neural Network-based model | Liu et al.(2016a) | Classification/ Neural Networks | Check-in data | Spatial-temporal RNN |

A study byLi et al. (2018)suggested STRP algorithm towards analysing the temporal and spatial regularity of object mobility. However, there is a need to minimize the execution or retrieval time ofonline prediction, to extract the trajectory patterns of moving objects, to improvise the prediction accuracywith respect tomore temporal characters, this is concerned in trajectory data.A research by Zhang et al.(2018)suggested a trajectory clustering algorithm that deliberates the semantic STD like time-based on Traclus, direction and speed. The simulated results show that the presented algorithm outperforms than the conventional reference spot detection method. But the limitation of this approach needs slightly more time than the single-level technique. Then generates periodic patterns and more reference spots which as unable to detect. So there is a need for ample experiments with more complex datasets, and there isan extension to multi-level time periods. Due to the wide spread ofspatiotemporal datasets, the data mining approach plays a major role; thus the rapid growth due to the location aware devices, sensor networks and particular features are linked with the dynamic dataset. However, the STDM faces the many challenging issues toward handling those huge collection of data with different domain hence it required to focus and effective solution required.

## VI. CONCLUDING REMARKS AND FUTURE RESEARCH NEEDS

In this study, we attempted to provide a broad overview of the field of STDM. However, the given literature on this topic we have discussed the different mining technique, clustering approach and machine learning approach for process of spatio- temporal data. Hence, the huge set of collected STD might be hided some useful data thus will help us knowledge extraction process. However, the manual examination of these information is unimaginable and information mining may give valuable devices and innovation in this process. STDM is a rising examination area that is devoted to the improvement of novel calculations and computational procedures for the effective investigation of substantial ST databases.

In future, there is a need of effective approach to capture the object movement behaviour in both temporal domain and spatial domain for prediction of location. Also, to transforms the trajectory into location sequences by mapping track points to fixed and discrete points in the road networks.In this regards, given an article ongoing development for location prediction, spatial-temporal data and trajectory data analysis in video applications. A brief summary of on going research is discussed as follows:

To develops a ST trajectory semantic deep neural network (STTS) algorithm along with Long-term short memory (LTSM) model building to capture the moving oject behaviour in both temporal domain and spatial domain for prediction of location. By using discrete points and mapping track points, the proposed technique able to renovates the trajectory data into sequence of location in road network.

Since traditional feature extraction failed to perform with long sequential data, the study, therefore, proposed a LTSM based model toward enhance the performance of data analysis.

 

We propose the STTS -LTSM prediction algorithm that provides accurate prediction and interpolating developing spatial based forecasting

In order to handle missing value in time series, date pre-processing mechanismlike filter of noise, compression of data and data filling information will be carried out using Fuzzy Gaussian Measure (Janaki, 2016) by incorporating into ST trajectory semantic deep neural network (STTS) algorithm. To validate the performance of proposed method with respect to accuracy and effectiveness, a real world trajectory information will be used.

The proposed STTS-LTSM model will be compared with conventional techniques like Hidden Markov model (HMM), RNN and xgboosimplemented in Python.

The performance of the proposed algorithm will be evaluated using precision, recall, and F-score withseveral evaluations.

## REFERENCES


[1] De Almeida, V.T. & Güting, R.H. (2005). Indexing the Trajectories of Moving Objects in Networks*. *Geoinformatica*. [Online]. 9 (1). Pp. 33–60. Available from: http://link.springer.com/10.1007/s10707-004-5621-7.

[2] Altche, F. & de La Fortelle, A. (2017). An LSTM network for highway trajectory prediction. In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. [Online]. October 2017, IEEE, pp. 353–359. Available from: http://ieeexplore.ieee.org/document/8317913/.

[3] Aryal, A.M. & Sujing Wang (2017). Discovery of patterns in spatio-temporal data using clustering techniques. In: *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*. [Online]. June 2017, IEEE, pp. 990–995. Available from: http://ieeexplore.ieee.org/document/7984703/.

[4] Bahadori, M.T., Yu, Q. (Rose) & Liu, Y. (2014). Fast Multivariate Spatio-temporal Analysis via Low Rank Tensor Learning. *Nips*. [Online]. Pp. 1–9. Available from: https://pdfs.semanticscholar.org/82e5/c9a1ccb59b7ca6952d1f2741881c3a15ab57.pdf.

[5] Bao, J., Zheng, Y. & Mokbel, M.F. (2013). *Location-based and preference-aware recommendation using sparse geo-social networking data*. [Online]. Pp. 199. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/locationrecommendation.pdf.

[6] Bengio, Y. (2010). Neural net language models. *Scholarpedia*.

[7] Bock, J., Beemelmanns, T., Klösges, M. & Kotte, J. (2017). *Self-learning Trajectory Prediction with Recurrent Neural Networks at Intelligent Intersections*. [Online]. (Vehits). Pp. 346–351. Available from: https://www.scitepress.org/papers/2017/63740/63740.pdf.

[8] Bogorny, V., Kuijpers, B. & Alvares, L.O. (2009). ST- DMQL: A Semantic Trajectory Data Mining Query Language. *International Journal of Geographical Information Science*. [Online]. 23 (10). Pp. 1245–1276. Available from: http://www.tandfonline.com/doi/abs/10.1080/13658810802231449.

[9] Bolstad, P. (2005). *Gis Fundamentals: A First Text on Geographic Information Systems.* [Online]. Eider Press. Available from: https://books.google.co.in/books?Id=cfheaaaayaaj.

[10] Cao, H., Mamoulis, N. & Cheung, D.W. (2007). Discovery of Periodic Patterns in Spatiotemporal Sequences. *IEEE Transactions on Knowledge and Data Engineering*. 19 (4). Pp. 453–467.

[11] Catlett, C., Cesario, E., Talia, D. & Vinci, A. (2018). A Data-Driven Approach for Spatio-Temporal Crime Predictions in Smart Cities. In: *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*. [Online]. June 2018, IEEE, pp. 17–24. Available from: https://ieeexplore.ieee.org/document/8421327/.

[12] Chen, L., Lv, M., Ye, Q., Chen, G. & Woodward, J. (2011). A personal route prediction system based on trajectory data mining. *Information Sciences*. [Online]. 181 (7). Pp. 1264–1284. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0020025510005888.

[13] Cheng, T. & Adepeju, M. (2014). Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *Plos ONE*.

[14] Elhorst, J.P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*.

[15] Ester, M., Kriegel, H.-P., Jorg, S. & Xu, X. (1996). A Density-Based Clustering Algorithms for Discovering Clusters. *Kdd*. [Online]. 96 (34). Pp. 226–231. Available from: https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.

[16] Ganguly, A.R. & Steinhaeuser, K. (2008). Data Mining for Climate Change and Impacts. In: *2008 IEEE International Conference on Data Mining Workshops*. [Online]. December 2008, IEEE, pp. 385–394. Available from: http://ieeexplore.ieee.org/document/4733959/.

[17] Giannotti, F., Nanni, M. & Pedreschi, D. (2006). Efficient Mining of Temporally Annotated Sequences. In: *Proceedings of the 2006 SIAM International Conference on Data Mining*. [Online]. 20 April 2006, Philadelphia, PA: Society for Industrial and Applied Mathematics, pp. 348–359. Available from: https://epubs.siam.org/doi/10.1137/1.9781611972764.31.

[18] Ignaccolo, R., Mateu, J. & Giraldo, R. (2014). Kriging with external drift for functional data for air quality monitoring. *Stochastic Environmental Research and Risk Assessment*.

[19] Janaki, V.R.; P.V.K.; V. (2016). Mining of Outlier Temporal Patterns. In: *2016 International Conference on Engineering & MIS (ICEMIS)*. 2016, Agadir, Morocco: IEEE Comput. Soc. Press.

[20] Jeung, H., Liu, Q., Shen, H.T. & Zhou, X. (2008). A Hybrid Prediction Model for Moving Objects. In: *2008 IEEE 24th International Conference on Data Engineering*. [Online]. April 2008, IEEE, pp. 70–79. Available from: http://ieeexplore.ieee.org/document/4497415/.

[21] Jeung, H., Shen, H.T. & Zhou, X. (2007). Mining Trajectory Patterns Using Hidden Markov Models. In: *Data Warehousing and Knowledge Discovery*. [Online]. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 470–480. Available from: http://link.springer.com/10.1007/978-3-540-74553-2_44.

[22] Kurashima, T., Iwata, T., Irie, G. & Fujimura, K. (2010). *Travel route recommendation using geotags in photo sharing sites*. Pp. 579.

[23] Lathia, N., Hailes, S. & Capra, L. (2009). *Temporal collaborative filtering with adaptive neighbourhoods*. [Online]. University College London. Available from: http://www0.cs.ucl.ac.uk/staff/n.lathia/papers/lathia_sigir09.pdf.

[24] Lei, P.-R., Li, S.-C. & Peng, W.-C. (2012). *QS-STT: quadsection clustering and spatial-temporal trajectory model for location prediction*. [Online]. Available from: https://www.researchgate.net/publication/232703396_QS-STT_quadsection_clustering_and_spatial-temporal_trajectory_model_for_location_prediction.

[25] Lei, P.R., Shen, T.J., Peng, W.C. & Su, I.J. (2011). Exploring spatial-temporal trajectory model for location prediction. In: *Proceedings - IEEE International Conference on Mobile Data Management*. 2011.

[26] Li, J., Qin, Q., Han, J., Tang, L.-A. & Lei, K.H. (2015). Mining Trajectory Data and Geotagged Data in Social Media for Road Map Inference. *Transactions in GIS*. [Online]. 19 (1). Pp. 1–18. Available from: http://doi.wiley.com/10.1111/tgis.12072.

[27] Li, X., Yu, C., Ju, L., Qin, J., Zhang, Y., Dou, L. & Yuqing, S. (2018). Position prediction system based on spatio-temporal regularity of object mobility. *Information Systems*. [Online]. 75. Pp. 43–55. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0306437917304726.

[28] Li, Z., Chen, J. & Baltsavias, E. (2008). *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*. ISPRS Book Series. [Online]. CRC Press. Available from:https://books.google.co.in/books?Id=-fqm3buoejcc.

[29] Li, Z., Ding, B., Han, J., Kays, R. & Nye, P. (2010).*Mining periodic behaviors for moving objects*. [Online]. University of Illinois at Urbana-Champaign. Available https://faculty.ist.psu.edu/jessieli/Publications/KDD10-zli-periodmovobj.pdf.

[30] Li, Z., Han, J., Ding, B. & Kays, R. (2012). Mining periodic behaviors of object movements for animal and biological sustainability studies. *Data Mining and Knowledge Discovery*. [Online]. 24 (2). Pp. 355–386. Available from: http://link.springer.com/10.1007/s10618-011-0227-9.

[31] Liu, Q., Wu, S., Wang, L. & Tan, T. (2017). Predicting the Next Location : A Recurrent Model with Spatial and Temporal Contexts. *Aaai*. [Online]. Pp. 194–200. Available from: https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/11900/11583.

## Authors Profile

**Mr. N. Venkata SubbaReddy** working as Associate Professor in the Dept. of Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Hyderabad, India. He has more than 10 years of teaching experience. He presently pursuing his PhD in the area of Data Mining. His research area of interest is Data Mining and Web Mining.

**Dr. D. S. R. Murthy** is currently working as a Professor & Head of the Dept of CSE in Geethanjali college of Engineering and Technology since June 2014.Earlier he worked as a professor of Information Technology in Sreenidhi Institute of science and Technology from OCT 2004 to JUNE 2011. Prior to that he was in NIT Warangal JNTUCE, Anantapur and ICFAI in different faculty positions of Computer Science. He is a Fellow of IETE, Senior Life Engineer (EI(I) & IETE). He published a text book on C Programming & Data Structures. His research interests are Image Processing, Image Cryptography and Data Mining.

 