Location Prediction Models using Data Mining and Machine Learning

Chetashri Bhadane, Ketan Shah, M. A. Khatkhatay, A. M. Darukhanawalla

Abstract*: A vast availability of location based user data which is generated everyday whether it is GPS data from online cabs, or weather time series data, is essential in many ways to the user and has been applied to many real life applications such as location targeted-advertising, recommendation systems, crime-rate detection, home trajectory analysis etc. In order to analyze this data and use it to fruitfulness a vast majority of prediction models have been proposed and utilized over the years. A next location prediction model is a model that uses this data and can be designed as a combination of two or more models and techniques, but these have their own pros and cons. The aim of this document is to analyze and compare the various machine learning models and related experiments that can be applied for better location prediction algorithms in the near future. The paper is organized in a way so as to give readers insights and other noteworthy points and inferences from the papers surveyed. A summary table has been presented to get a glimpse of the methods in depth and our added inferences along with the data-sets analyzed.

***Keywords*: context, mobility, next-location prediction, trajectory.**

I. INTRODUCTION

Location data can be in the form of individual mobility patterns, trajectory, communication or positional data all of which have a wide range of applications. Location prediction forms the basis for many top business firms, market prediction indicators, advertisers, defense and security firms, tower research, travel demand distribution etc. It can be obtained using GPS, WLAN, mobile data, WIFI etc. is an International reputed journal that published research articles globally.

A business recommendation system can use the hidden factors exhibited by user mobility patterns for its success. According to Shudong Liu and Xiangwu Meng [1] user mobility often evinces short term and long-term factors such as daily activity and social networking ties which can be followed and anchored for business Information recommendation. A region-based location graph was developed and compounded with short distance and long-distance mobility factors which displayed both local and long-distance business information to users. Moreover, the cold start problem faced by traditional recommendation systems was palliated by amalgamating user-based collaborative filtering with item-based collaborative filtering.

Advertising firms and agencies can use specific geographic based data like trajectory to find trends in consumer needs and accordingly predict what products to sell where. According to Anindya Ghose, Beibei Li, and Siyuan Liu [2] trajectory-based mobile advertising leads to a high redemption probability, fast redemption behavior, and high transaction amount from consumers compared to other kinds of advertising, all of which facilitate higher revenues for a particular shop as well as the shopping mall as a whole. Interestingly, however trajectory-based mobile advertising becomes less effective in increasing the revenues of the shopping mall as a whole during the weekend.

II. NEXT PREDICTION MODELS

Although location prediction and location recommendation sound similar they are different in a number of ways. Location prediction tells us where a person or object in consideration might or might not be in the future whereas location recommendation tells a person or object where it would be most beneficial to be in the future. One key difference between Location recommendation and prediction is that location recommendation uses popularity as a factor whereas location prediction does not require popularity.

Location prediction requires user or object mobility patterns which indirectly require intentions for prediction. According to Josh Jia-Ching Ying, Wang-Chien Lee, Vincent S. Tsengin [3] there are three categories of intentions:

- **Geographic-triggered:** These types of intentions are based on the geographic location in consideration. For example, people at a railway station A may have location station B or station C as their next predicted location which are next on the route to station A.
- **Temporal-triggered:** These types of intentions are based on the mundane or daily activities of a person. For example, a person leaves his house at a fixed time in the morning and returns at a fixed time in the evening indicates his workday.
- **Semantic-triggered:** These types of intentions are based on logical or symbolical inferences that can be derived from the behavioral pattern. For example, a person visiting a church can be termed as religious where as a person visiting a theatre can be termed as recreational or entertainment.

Revised Manuscript Received on February 26, 2020.
***** Correspondence Author
Chetashri Bhadane*, Assistant Professor, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India. Email: chetashri@gmail.com
Mohammed Aqid Khatkhatay, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India. Email: mohdaqidkhat98@gmail.com
Dr. Ketan Shah, Professor, SVKM's MPTSME, Mumbai, India. Email: ketanshah@nmims.edu
Aamir Murtuza Darukhanawalla, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India. Email: aamir@darukhanawalla.com

A. K-Means

K-means is a very basic clustering technique which forms circular clusters based on some distance metric like Euclidean distance function or Manhattan distance function. K-means initially selects random points as centroids or centers then through a series of iterations converges to a solution. Before we understand whether k-means is suitable for our location prediction model or no we should first understand the term geodesic distance.

Geodesic distance can be defined as the shortest or the least possible distance on a spatial reference plane or curved surface. Visually, it looks ellipsoidal. A detailed study was carried by Jacopo Grazzini, Pierre Soille and Conrad Bielski [4] on the uses of spatial interpolation. It was found that using generalized geodesic distances instead of Euclidean ones enabled one to change the sample points and weights during Gaussian regression interpolation.

K-means has been used in the past for identifying spatial pattern in storms by Gupta, Upa & Jitkajornwanich, Kulsawasd & Elmasri, Ramez & Fegaras, Leonidas [5] where it was made to adapt to identify different hourly storms based on their sizes and shapes. The paper was successful in identifying different classes of storms which can be used for future geospatial prediction models.

Though k-means can be adapted or used as a very basic medium for location prediction, there are still drawbacks inherent in it. As per Geoff Boeing [6] K-means is not suitable for geo-spatial data mining purposes involving distances, in the form of latitude-longitude as it does not minimize geodesic distances rather minimizes variance. A very significant amount of deformation occurs at latitudes far from equator thereby producing poor results.

B. Bayesian Methods

A Bayesian model is a statistical model that uses probability and uncertainty in both input and output to come to some result. A Bayesian model is heavily dependent on Bayes Theorem which is as follows-

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)} \quad (1)$$

where P (A| B) is the posterior probability or probability of event A occurring provided event, B has already occurred. P(A) and P(B) are the prior probabilities of events A and B. P (B | A) is the conditional probability of event B occurring with respect to event A.

There are numerous Bayesian methods such as Bayesian Belief networks, Bayesian dynamic networks, Naive Bayes etc. As Bayesian Networks take into account uncertainty, they can be used for a large number of prediction models including location and context-based ones. A good choice of prior knowledge can increase the accuracy of the model.

An approach was suggested by Djamel Guessoum, Moeiz Miraoui, and Chakib Tadj [7] for predicting a person's outdoor location based on some context. Algorithms for this were extensively surveyed including Bayesian models. For the prediction step five supervised learning methods (Support Vector Machine, J48 Decision tree, Naive Bayes, Bayesian Networks, and K-Nearest Neighbors) were tested using 10-fold cross validation technique. Prediction accuracy was tested both with and without noise. The results were plotted in the form of a ROC (receiver operating characteristics) curve for visualizing purpose. The curve had sensitivity (True Positive Rate) on the y-axis and False positive rate (specificity) on the x-axis. The results were averaged in terms of overall prediction accuracy as shown in the table below. Though it was concluded that J48 Decision Tree produced best results but for ROC based parameters Naive Bayes was promising.

Table- I: Average Prediction Accuracy as obtained by [7]

Outdoor Location Context	
Algorithm Used	***Average prediction Accuracy***
Naïve Bayes	67.56 %
Bayesian Network	86.54%
J48 Decision Tree	97.09%
SVM	77.48%
KNN	98.09%

Another context awareness application was suggested by Chakkrit Snae Namahoot, Michael Brückner, and Naruepon Panawong [8] in tourism industry in the form of a recommender system called Context Awareness Tourism or CAT. An improved version of Naive Bayes called Naive Bayes with Boundary vales (NBB) was proposed. This approach involved four steps. The first step was the training of a Naive Bayes model with web pages of six categories (Attraction, Dining, Accommodation, Souvenir, Events and One Tambon, One Products (OTOP)). The second step involved improving the efficiency of the Naive Bayes Model. The third step involved testing the NBB on google search results and using F-score as a performance metric. The final step was implementing NBB for CAT. Though NBB exhibited very efficient and precise results, it however failed to consider the user's current location as a parameter for recommendation. Another drawback was its reusability test.

One more location-based application was suggested by Banu Wirawan Yohanes, Samuel Yanuar Rusli, Hartanto Kusuma Wardana in [9] where Naive Bayes probability was considered in received signal strength (RSS) in order to predict indoor location. According to them location prediction was made up of two phases:

- **Offline Phase:** This is the training phase where RSS values are collected and stored in a database.
- **Online Phase:** This is the testing phase where calculations for estimating the location are carried out.

Naive Bayes was used to find the location of the access point once RSS was generated because of its simplicity and good accuracy.

C. Frequent Pattern Mining

Frequent Pattern Mining is the process of finding repeating patterns in the dataset. Different categories of dataset contain different patterns. For example, an NLP dataset may contain a bag of words, similarly a market basket dataset contains itemset. There are three commonly used rules for association in pattern mining:

- **Support:** Support is the measure of how frequent an item is in a transaction. Support tells us which rules are to be considered for further analysis. It can be formulated as follows:

$$Support\{A => B\} = \frac{\text{Transactions having both A \& } B}{\text{Total number of Transactions}} \quad (2)$$

where A is the antecedent and B is the consequent in the transaction.

- **Confidence:** Confidence is the measure of the conditional probability of occurrence of the consequent given that antecedent has already occurred. It can be formulated as follows:

$$Confidence\{A => B\} = \frac{\text{Transactions having both A \& } B}{\text{Transactions having B}} \quad (3)$$

- **Lift:** Lift is a measure of likeliness keeping prior probability in mind. It can be formulated as follows:

$$Lift\{A => B\} = \frac{\text{Support (X} \cup \text{Y)}}{\text{Support (X)} * \text{Support (Y)}} \quad (4)$$

Pattern Mining can be useful in location prediction in a number of ways. If trajectory data is represented in the form of a graph, mined patterns can serve as good indicators of future locations. An approach for graph-based pattern mining using Apriori was proposed by Akihiro Inokuchi, Takashi Washio and Hiroshi Motoda [10] where subsets in a graphical dataset were mined using an approach known as Apriori based Graph Mining (AGM). It was evaluated on both simulation data and real-world data, both showing powerful performance results.

Sometimes location data points will be in a sequence, for such data points many frequent data mining algorithms cannot be used as they do not consider sequence. We use sequential mining techniques which are somewhat based on string processing algorithms. Sequential mining was first introduced by R. Agrawal and R. Srikant [11] where sequential pattern mining was used to find all the subsequences of a set of sequences.

Another application of pattern mining was the prefix scan approach by Jian Pei, Jiawei Han,,Behzad Mortazavi-Asl, Jianyong Wang, Helen Pinto, Qiming Chen, Umeshwar Dayal and Mei-Chun Hsu [12] where a sequence database was projected recursively into a smaller databases. These smaller databases were checked locally for frequent pattern growths. A further improvement in performance was achieved by using a pseudo projection technique in PrefixSpan. It was concluded that instead of improvising apriori like candidate set generation algorithms, it is better to use divide and conquer based pattern growth approaches which are extensions of FP-growth.

Pattern trajectory, mined from historical data can be used for predicting locations. This was demonstrated by L. D. M. Lam, A. Tang and J. Grundy in [13] where indoor spatial movements were predicted using data mining and movement patterns. An empirical study of a tear-room was carried and a probability tree was constructed. The system architecture proposed consisted of three layers. The first layer was the sensing layer which collected the sensory data and gave it to the next layer.

The second layer was the positioning layer which computed the positional data from the sensory data and gave it to the next layer. The next layer was the contextualizing positioning layer which converted the positional data into points of interests. These were then further stored in a database for mining patterns. For predicting location priority allocation strategy was used.

The usefulness of mined pattern was successfully demonstrated in predicting next location.

Another precise application of pattern mining related location prediction was demonstrated by Z. Zhang and W. Zhu in [14] where an improved Apriori algorithm called AprioriOS (Apriori for Ordered Sequences) was proposed for mining consumer trace in a large shopping mall. Firstly, the paper designed a method to discard outliers and sequence the data into regions using accelerometer of mobile phones or RFID (Radio Frequency Identification) sensor. Next, AprioriOS was proposed which mined frequent patterns, generated association rules and predicted a number of regions that the customer could visit. Then an association rule querying and tree storing structure was designed for the model. Finally, a motion prediction function was added. This method was found have a very high accuracy and an even high operating range.

D. Decision Tree

A decision tree is a tree based supervised learning algorithm. It is constructed in a top-down recursive manner similar to divide and conquer. Every internal node in the tree represents a condition whose outcome decides the branch. Every branch represents the outcome. Every leaf or terminal node represents the class label or parameter of prediction. Every path from the root node to the leaf or terminal node represents a classification rule which is generated after the training phase.

Decision tree can be useful in predicting next location due to its easy implementation and high prediction accuracy. It has been studied a number of times for this purpose. A thorough proposal regarding its use was made by Jae Sung Lee & Eun Sung Lee in [15] where the usefulness of using Decision tree in predicting people's location was studied. The C4.5 decision tree algorithm was used for its effectiveness in approximating discrete valued functions. A contextualized dataset was compiled in Excel and used for the experiment which consisted of 20 attributes. Every activity was coded in the form of numbers. Weka software was used to construct the tree and rules. The rules constructed were very simple and easy for the users to interpret. True positive Rate was used as a performance metric. The method produced very promising results and also had the advantage of handling noisy data.

Another supporting approach for location prediction can be obtained in the form of traffic prediction which can determine a person's next location. Decision tree-based approach was suggested by in Alajali, Walaa, Zhou, Wei, Wen, Sheng, Wang, and Yu [16] for intersection traffic prediction. In this paper an approach was proposed for intersection traffic prediction by using data collected from road accidents and roadwork reports. Batch learning was done using the algorithms: Support Vector Regression (SVR), Gradient Boosting Regression Trees (GBRT), Random Forest (RF) and Extreme Gradient Boosting Trees (XGBoost). Online learning was done using Fast Incremental Model Trees with Drift Detection (FIMT-DD) model.

The results obtained were checked using 10-cross fold validation. The metrics used were Mean Squared Error (MSE) and Mean Absolute Error (MAE) results which are formulated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \qquad (5)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y}_i|^2 \qquad (6)$$

where y_i is the actua; output & $\bar{y}_i$is the estimated output

Table- II: Results obtained by [16]

Learning Model	*MSE Sensor only*	*MAE Sensor Only*
SVR	0.6474	1.4420
RT	0.06807	1.4383
GBRT	0.6721	1.4281
RF	0.6874	1.4405
XGBOOST	0.6721	1.4300

It was concluded that XGBoost had the best accuracy followed by GBRT. Further when the dataset was increased in volume RF used less time whereas XGBoost required more time.

Decision Tree can also be used for contextualized location prediction. Xia, Linyuan and Huang, Qiumei and Wu, Dongjin in [17] proposed an approach to predict next semantic location based on historical patterns obtained from mobile logs. A rough flow of the approach was in three steps. The first step analyzing GPS trajectories and extracting point of interests using stay point detection method. The second step involved semantic place recognition. The third step involved extracting spatio-temporal features using decision tree. The decision tree used in the third step was an ID3 decision tree and was built in two steps which involved growing and pruning. The approach was tested on two datasets. One was the authors own and the other was Microsoft's Geolife. For validating the decision tree approach, it was compared with the classical Markov Model. Precision, recall rate, and F-measure were the performance metrics used which concluded that the ID3 Decision tree algorithm achieved better contextual location prediction.

E. Mobility Chain

One method to predict next location is the use of mobility chains. Sebastien Gambs, Marc-Olivier Killijian, Miguel Nunez del Prado Cortez [18] made use of a new Mobility Chain with a probabilistic automaton known as Markov Mobility Chain to represent the behavior of an individual. This paper also proposed a method for learning from the mobility patterns of individuals. The mobility pattern consisted of four components. The first component was the identifier which served as the actual identity. Unknown was used in case it was not known. The second component was the spatial component such as latitude, longitude, home, work, or name of some place. Experiments were conducted and compared with actual data results. The third component was the time stamp consisting of the date and time. The fourth component was some additional information speed of the vehicle, direction etc. The mobility chain used consisted of a set of states containing points of interests (POIs) and some transition states represented in the form of a transition matrix or directed graph. The mobility chain when represented on a real map gave accurate trajectories which when subjected to adversary inference attacks gave the semantic label. Thus, it was concluded that mobility chain is a very compact and precise representation of the mobility pattern of a person. The drawback of this model was that it considered only current location to predict the next which can possibly negatively impact accuracy.

In-order to improve the drawbacks to the previous paper an extension was proposed by the same authors in [19] to improve the accuracy of the previous model. The mobility chain was extended and this time included n- previous locations and hence called n-MMC. The components remained the same only extension was in terms of the n-MMC learning algorithm and next place prediction. The algorithm worked in two steps. The first step involved pre-processing static traces then extracting POIs using clustering. The clustering technique used was Density-Joinable cluster (DJ-Cluster). DJ-Cluster has three input variables. The minimum number of points needed to form a cluster - MinPts, the maximum radius of the cluster circle ε and the cluster merge-distance d_{mer} . Once POIs are formed transition and associated probabilities needed for the Markov chain were computed and calculated in chronological order of mobility patterns. Once mobility patterns were identified, they were labelled in two passes. Any unidentified POIs in the first pass was labelled unknown. These were then removed in the second pass and successive labels were combined into one. To predict the locations a modified form of the transition matrix was used. Most probable locations were derived from the Markov chains. In conclusion accuracy ranged from 70 % -95 % and it seemed to grow with the value of n.

Another application of Markov Chain was use for predicting route in [20]. The paper went on to present a framework for predicting route by using Bayesian classification and stochastic process. Assuming that each pedestrian has a context and a group of similar pedestrians show common behavior. These were clustered accordingly. A model was trained based on the history of the driver's trips. The history was a set consisted of time. Each trip in the history set belonged to some cluster based on properties of the trip such as journey pattern, destination etc. Journeys were selected as a stochastic process. Prior probabilities are estimated from context of the trip and additional information like weather, public events etc. Once prior probabilities were estimated Bayes law was used to map the likelihood of each point with respect to a cluster. A criterion was suggested for stopping the process. Then finally a cluster prediction algorithm was proposed which used Markov chain based on the assumption that only the current road segment determines the current trip and the probability distribution of the square of next road segment. Two clusters were formed during experimentation one was based on route and the other on source or destination. The results showed the model achieved 90 % accuracy for a test of around 780 trips. It was concluded that modelling prediction on Markov Chain and Markov assumption is a strong prediction approach.

F. HMM

Hidden Markov Models are mathematical and statistical Markov models which include an unobserved or hidden state and simplify using the Markov assumption.

As real-life location or trajectory data sets contain many hidden or partially observable states HMM is a good model for these. The approach to apply mobility prediction through HMM was proposed by Pratap S. Prasad and Prathima Agrawal in [21]. This paper proposed a method to extract mobility data, train an HMM based on this data and make predictions. The HMM model was based on detecting movement with respect to access points. The access points were visible and the actual geographical locations were hidden. Probabilities and likelihoods were calculated using the Viterbi algorithm to determine where the location of users was hidden. Experiments were carried on a campus wireless trace data set containing user histories association with access points etc. It was observed that the accuracy decreased with the increase in sequence length due to some biases. Thus, it was concluded that HMM can be used for predicting user movements.

Another mobility related application of HMM was proposed by Xuan Song, Quanshi Zhang, Yoshihide Sekimoto and Ryosuke Shibasaki in [22] for predicting human behavior after natural disaster. This paper proposed HMM as a disaster behavior model where number of hidden states were determined using Bayesian Information Criterion (BIC) and the transition probabilities were determined using Baum-Welch algorithm. The approach worked in 6 key steps. The first step was initialization which involved initializing the initial state probabilities. The second step was re-sampling which involved re-sampling particles according to weights. The third step was prediction of next state from transition probabilities. The fourth step was weighting which involved recalculating the weight of the states. The fifth state was State Estimation which involved estimating people's behavior by calculating expectation of the particle set. The sixth and final step was iteration which involved iterating the steps 2 to 5 till convergence. For evaluating the model GPS data of Japan, months before the earthquake (1st, August 2010 to 11st, March 2011) days after was used. The accuracy of Greater Tokyo region was found to be at-least 60%.

A more precise and accurate model for location prediction using HMM was proposed by Yong-Joong Kim and Sung-Bae Cho in [23] for developing a location prediction model for mobile context aware services. The model consisted of two phases a location recognition phase and a path classification phase. The location recognition phase involved using a decision tree and KNN while the path classification phase involved HMM. HMM was selected as it involved both hidden states and previous
actions. Experiments were carried on data set generated by five students using Samsung Galaxy S3 as the device. The accuracy of the path was found to be greater than 80% whereas the accuracy of the paths for several users was found to be less than 70 %. The average accuracy was found to be higher than 87% throughout the experiment.

G. Neural Networks

Neural Networks are biologically inspired machine learning models which can be applied to solve any machine learning problem using complex mappings.
A Recurrent Neural Network (RNN) is a directed graph containing several artificial neural networks which for a network of several interconnected nodes exhibiting dynamic behavior. An RNN based model for activity and location prediction was suggested by Dongliang Liao, Weiqing Liu, Yuan Zhong, Jing Li, and Guowei Wang in [24]. The proposed method called MACARNN- Multi-task Context Aware Recurrent Neural Network used spatial activity for the prediction part. The model took input from the user consisting of activities, check-in time, activity graphs etc. These were then fed accordingly to Context Aware Recurrent Unit (CARU) via the linear unit embedding. The CARU was responsible for calculating the sequential hidden state using several activation functions. These were then fed to the location and activity prediction task specific layers which used gated RNN and SoftMax activation function.
An accuracy of approximately 60% for location prediction and approximately 71% activity prediction was obtained for the New York City data set thus concluding that MACARNN performs better than earlier RNNs.

Another application of RNN in location prediction was put forth by Qiang Liu, ShuWu, LiangWang and Tieniu Tan in [25] for predicting next location. This paper proposed a method called Spatial Temporal Recurrent Neural Networks (ST-RNN) for holding better spatial and temporal information. As a drawback of generic RNN's inability to handle continuous time intervals this paper proposed time specific transition matrix to take care of this. Similarly, the generic RNN fails to model geographical distances which was taken care by a similar distance specific transition matrix. Experiments were carried on the Gowalla and Global Terrorism Database data sets and evaluated using area under ROC (AUC) and Mean Average Precision (MAP).
A MAP of 0.1 and 0.3 and AUC values of 0.8 were obtained for GTD & Gowalla respectively thus showing ST-RNN can be used for location prediction.

III. SUMMARY

We can summarize the various models and approaches studied in the previous sections in a tabular manner with their observations, experimental results, accuracy and data sets used in the table III.

Table- III: Summary of Models

Model/Approach	Observation	Experimental Results	Accuracy	Dataset Used
K means adapted	Convergence depends on the similarity measure or distance metrics. Time complexity to cluster M hourly storms is $O(kMN^2)$	Set Based Similarity Index optimizes result.	60% of 13831 hourly storms were clustered	50-year Raw rainfall data of Texas and surrounding areas
Naive Bayes	Time variant Bayesian networks represent temporal contextual data set properties Bayesian networks account for inherent uncertainty	Prediction accuracy enhances with noise predicted data	AUC value of approximately 0.95 for Bayesian Networks	Mobile Data Challenge database by Idiap Research Institute and owned by Nokia
Naive Bayes with Boundary Values	Province and dates generate suggestions User's current location not needed as a parameter	Classifying web pages into diverse categories possible Imbalanced frequency of words in categories	Average precision 100% Average F-Score 97.03%	1,048 Thai tourism Web pages from Truehits
Naive Bayes Received Signal Strength	Naive Bayes gives the probability of future events Can find a better access point from a combination of two access points	Naive Bayes provides high speed	A 75:25 % Training Testing ratio gave an accuracy of 49.13% More data needed for better accuracy	Netsurveyor Fingerprint Database
Pattern growth PrefixSpan approach	Database scanned one by one	Order of points can be varied	Running time of 3539.78 sec	Gazelle Data-set
Pattern mining movements	Movement patterns represented graphically indicate spatial behaviour	An individual's activity depends on the point of interest	POI prediction accuracy 56-78 %	Manually collected 6 weeks of data using camera footage from a staff tear room
AprioriOS	Accelerometer is an integral part of motion prediction	Customer information is obtained by querying the database Insufficient confidence can lead to less future prediction	Location prediction accuracy of 0.92	No data set used only simulation
Decision Tree C4.5	Discrete valued function approximated Previous activity is considered	Context decided through activity rules	True Positive Rate of 0.51 and ROC area of 0.837	Campus data set collected by 335 students
Decision Tree ensemble	Weak learners can be merged to give a powerful model	Significant handling of sparse data	MAE of 1.4420 for SVR, 14383 for RT, 1.4281 for GBRT, 1.4405 for RF and 1.4300 for XGBoost	Intersection traffic volume data set consisting of sensor data collected in Melbourne, Victoria, Australia
Decision Tree ID3	Spatial and semantic features assist contextual location prediction	High performance on low frequency prediction due to complete utilization of spatial temporal features	Approximately 70% prediction accuracy	Geolife dataset of 178 users in Beijing, China was collected by Microsoft Research Asia and author's own dataset through GPS
Mobility Markov Chain	All important and unimportant POIs are represented as states to assist in finding mobility trace	Inference attacks help to calculate mobility behaviour Semantic labels also affect mobility behaviour	Overall precision better than 50%	Data collected using Nokia 5800 phones
n-Mobility Markov Chain	Previous visited locations are considered	Accuracy of prediction increases with n	location accuracy of 70% to 95% for n=2	Geolife dataset of shanghai, China was collected by Microsoft Research Asia
Markov Chain Bayesian Classifier	Markov chains can be modelled as clusters	Stochastic process release road segments which indicate journey to be predicted	At-least 90% trips predicted correctly	GPS data of Dublin simulated in microscopic traffic simulator SUMO

HMM	Prediction Accuracy decreases with increase in sequence length due to some biases. User Behavioural Issues affect prediction	Prediction model easy to implement in existing network architecture	Prediction accuracy approximately greater than 70 % Prediction accuracy decreases with increase in sequence length	Crawdad: Wireless traces from Dartmouth
HMM-Bayesian Information Criterion & Baum-Welch Algorithm	Mobility three days after was found to be same as mobility before	The three prefectures showed better results than greater Tokyo Patterns biased to younger age groups	At-least 90% visual data matching	GPS data of Japan, months before and after the earthquake (1st, August 2010 to 11st, March 2011)
HMM Context-awareness	Intermediate location path generated using G-means form HMM containing multiple context	Amount of data that was collected was found to be less than actually needed for HMM's learning process Movement patterns were too diverse for prediction	Several path accuracies for users were less than 70 % and path accuracy was higher than 80% Average Prediction accuracy higher than 87\%	Generated by students using Samsung Galaxy S3
MACARNN- Multi-task Context Aware RNN	Time and time-span of the data both affect the RNN hidden state size in the embedded layer Auxiliary representation learning task affects convergence	Size of hidden state determines the capacity of RNN	60% location prediction and 71% activity prediction for NYC 74% location prediction and 87% activity prediction for TKY	New York City (NYC) Foursquare check in data set Tokyo (TKY) Foursquare check in data set
STRNN-Spatial Temporal Recurrent Neural Networks	STRNN requires linear interpolation for training STRNN considers elements in spatio temporal context	STRNN not very sensitive to dimensionality	Mean Average Precision (MAP) of 0.1 and Area under the ROC curve (AUC) of 0.8 for GTD Mean Average Precision (MAP) of 0.3 and Area under the ROC curve (AUC) of 0.8 for Gowalla	Global Terrorism Database (GTD) Gowalla dataset

IV. RESULT AND DISCUSSION

The above table III gives valuable insights into our summarization of all models discussed previously. We have compiled a comprehensive analysis so that it is easy for the readers to interpret our survey. As we can clearly infer that selecting the right model for predicting next location based on context, activity, or semantic does not have a clear solution. Every scenario needs its model which may be different for a different scenario. For instance, a person's routine activities in a closed college or university campus can be predicted with a high amount of certainty using Decision tree C4.5 but the route of a cab requires a more complex Bayesian model clubbed with mobility chain. Similarly, special circumstance or not so mundane activities during a disaster can be predicted using HMM. Even small things like a shopping mall or closed indoor require different models like pattern mining and Bayesian networks for their accurate representation and computationally easy implementation.

V. CONCLUSION

To conclude, selecting the right model for predicting next location based on context, activity, or semantic will require a thorough understanding of one's own problem statement, the dataset or initial research that will become the basis of the model. We have presented a complete analysis of each model used, how it progressed over the years to ensure that the readers will proceed with their own model in a similar fashion. Finally, Location prediction with or without context is an important aspect which will dominate both the near and far future.

REFERENCES

1. Shudong Liu and Xiangwu Meng, "A Location-Based Business Information Recommendation Algorithm", Special issue on Advanced Modeling and Services Based Mathematics for Ubiquitous Computing of Mathematical Problems in Engineering, Volume 2015, Article ID 345480, https://doi.org/10.1155/2015/345480, Hindawi publishing Corporation.
2. Ghose A, Li B, Liu S., "Trajectory-Based Mobile Advertising", Working Paper, New York University, New York, 2018
3. Josh Jia-Ching Ying, Wang-Chien Lee, Vincent S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction", ACM Trans. Intell. Syst. Technol. 5, 1, Article 2 (January 2014), 33 pages. DOI: https://doi.org/10.1145/2542182.2542184
4. Jacopo Grazzini, Pierre Soille and Conrad Bielski, "On the use of geodesic distances for spatial interpolation", Proceedings of GeoComputation 2007 p. Paper 7C3, National Centre for Geocomputation, 2007.
5. U. Gupta, K. Jitkajornwanich, R. Elmasri and L. Fegaras, "Adapting K-means clustering to identify spatial patterns in storms," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 2646-2654.
6. Geoff Boeing, "Clustering to Reduce Spatial Data Set Size", arXiv:1803.08101v1 [cs.LG] 21 Mar 2018.
7. Guessoum, D., Miraoui, M. and Tadj, "Contextual location prediction using spatio-temporal clustering", International Journal of Pervasive Computing and Communications, Vol. 12 No. 3, pp. 290-309, 2016 https://doi.org/10.1108/IJPCC-05-2016-0027
8. Chakkrit Snae Namahoot, Michael Brckner, and Naruepon Panawong, "Context-Aware Tourism Recommender System Using Temporal Ontol-ogy and Nave Bayes", Recent Advances in Information and Communication Technology, pp.183-194, 2015
9. Banu Wirawan Yohanes, Samuel Yanuar Rusli, Hartanto Kusuma Wardana, "Location prediction model using Naïve Bayes algorithm in a half-open building", 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE),15-19, 2017

10. Inokuchi A., Washio T., Motoda H., “An Apriori-Based Algorithm for Mining Frequent Substructures from Graph Data”, In: Zighed D.A., Komorowski J., Żytkow J. (eds) Principles of Data Mining and Knowledge Discovery. PKDD 2000. Lecture Notes in Computer Science, vol 1910. Springer, Berlin, Heidelberg, 2000
11. R. Agrawal and R. Srikant, “Mining Sequential Patterns”, Proc. 1995 Intl Conf. Data Eng. (ICDE 95), pp. 3-14, Mar. 1995.
12. Jian Pei, Jiawei Han, Behzad Mortazavi-Asl, Jianyong Wang, Helen Pinto, Qiming Chen, Umeshwar Dayal, and Mei-Chun Hsu., “Mining Sequential Patterns by Pattern-Growth: The PrefixSpan Approach”, IEEE Trans. on Knowl. and Data Eng. 16, 11 (November 2004), 1424–1440. DOI: https://doi.org/10.1109/TKDE.2004
13. L. D. M. Lam, A. Tang and J. Grundy, "Predicting indoor spatial movement using data mining and movement patterns," 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, 2017, pp. 223-230.
14. Z. Zhang and W. Zhu, “Location and Motion Prediction of Consumers in a Large Shopping Mall” 2017 Fifth International Conference on Advanced Cloud and Big Data (CBD), Shanghai, 2017, pp. 250-255. doi: 10.1109/CBD.2017.50.
15. Lee, Jae & Lee, Eun., “Exploring the Usefulness of a Decision Tree in Predicting People's Locations”, Procedia - Social and Behavioral Sciences. 140. 10.1016/j.sbspro.2014.04.451, 2014.
16. Walaa Alajali, Wei Zhou,Sheng Wen, Yu Wang, “Intersection Traffic Prediction Using Decision Tree Models”, Symmetry 2018, 10(9), 386; https://doi.org/10.3390/sym10090386, 2018
17. Linyuan Xia, Qiumei Huang, Dongjin Wu, “Decision Tree-Based Contextual Location Prediction from Mobile Device Logs. Mobile Information”, Systems. 2018. 1-11. 10.1155/2018/1852861
18. Sebastien Gambs, Marc-Olivier Killijian, Miguel Nunez del Prado Cortez , “Show Me How You Move and I Will Tell You Who You Are”, SPRINGL '10: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS, November 2010.
19. Gambs, Sebastien´ and Killijian, Marc-Olivier and del Prado Cortez, Miguel Nunez˜, “Next Place Prediction Using Mobility Markov Chains”, Proceedings of the First Workshop on Measurement, Privacy, and Mobility-2012.
20. Epperlein, Jonathan P. et al. “Bayesian classifier for Route prediction with Markov chains.” 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (2018): 677-682.
21. Prasad, Pratap & Agrawal, Prathima., “Movement Prediction in Wireless Networks Using Mobility Traces”, 1 - 5. 10.1109/CCNC.2010.5421613, 2010
22. Xuan Song, Quanshi Zhang, Yoshihide Sekimoto, and Ryosuke Shibasaki, “Prediction of human emergency behavior and their mobility following large-scale disaster”, In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD ’14). Association for Computing Machinery, New York, NY, USA, 5–14, 2014 DOI: https://doi.org/10.1145/2623330.2623628
23. Yong-Joong Kim, Sung-Bae Cho, “A HMM-Based Location Prediction Framework with Location Recognizer Combining k-Nearest Neighbor and Multiple Decision Trees HAIS 2013, LNAI 8073, pp. 618–628, Springer-Verlag Berlin Heidelberg, 2013
24. Dongliang Liao, Weiqing Liu, Yuan Zhong, Jing Li, and Guowei Wang, “Predicting activity and location with multi-task context aware recurrent neural network”, In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI’18). AAAI Press, 3435–3441, 2018
25. Qiang Liu, Shu Wu, Liang Wang, Tieniu Tan, “Predicting the Next Location:A Recurrent Model with Spatial and Temporal Contexts”, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16), 2016

AUTHORS PROFILE

Chetashri Bhadane, currently an Assistant Professor at the Dwarkadas J. Sanghvi College of Engineering, permanently affiliated to Mumbai University. She has been teaching in academia for the more than a decade. She holds a B. E and M.E. both in Computer Engineering from the prestigious NMU and NMIMS Deemed-to-be-University respectively. She is currently pursuing her Ph.D. from the latter. She has numerous publications in reputed journals and conferences both nationally and internationally.

Dr. Ketan Shah, specializes in Distributed Computing, Parallel Processing and Data Mining. He has been teaching in academia for close to two decades. He holds a B.E and M.E both in Electronics from Mumbai University and later was awarded a Ph.D. in Information Technology from NMIMS Deemed-to-be-University. He also has associate dean accreditations and specializes in Data Mining. He has published research papers at reputed national and international journals, conference proceedings as well as chapters of books.

Mohammed Aqid Khatkhatay, currently a final year student at student at the Dwarkadas J. Sanghvi College of Engineering, permanently affiliated to Mumbai University. He aspires to pursue a well-rounded career in Computer Science with a specialization in Data Science. He is academically well versed, conceptually strong and believes in learning through challenges and solving complex real-world problems through thorough research and hands on industrial experience. He has published a paper titled “A Comprehensive Study and Analysis of Semi Supervised Learning Techniques” in the International Journal Engineering Research and Technology, Vol 08 Issue 11, November 2019 as third author.

Aamir Darukhanawalla, currently a final year student at the Dwarkadas J. Sanghvi College of Engineering, permanently affiliated to Mumbai University. His current interests lie in the field of Data Science and Reinforcement Learning. He has displayed his technical skills at various internships and hackathons. As a proficient code her believes in learning through experience. He aspires to pursue a fruitful career in Data Science. He has completed his research internship in Machine Learning at the Indian Institute of Technology, Bombay. He was part of the team that won Capgemini’s TechNext HackIT Hackathon, 2019 at the KJ Somaiya College of Engineering.