

# Location Prediction with Markov Model using Long Term Evolution Datasets

S. S. Daodu

Department of Computer Science  
University of Benin, Benin City

Akinola E.

Department of Computer Science  
Federal University of Technology, Akure

## ABSTRACT

Location prediction is fast becoming a wide field for research and has received great attention from diverse fields. Markov model are of various types but the memoryless property of the Markov model makes it easily applicable in location prediction. Network dataset chosen for the analysis and evaluation of the proposed system is a 4G LTE dataset with channel and context metrics. This dataset is an LTE network known as Beyond Throughput: a 4G LTE Dataset with Channel and Context Metrics dataset developed by Raca et al., (2018). It is a 4G trace dataset which is composed of client-side cellular key performance indicators (KPIs) with 135 traces. This paper focuses on using Markov model to efficiently and successfully predict a user location using the long term evolution network datasets.

## Keywords

Prediction, Cell Id, States ST-RNN, Markov

## 1. INTRODUCTION

Markov model chains according to Chan and Lenard 2012, are mathematical models that use concepts from probability to describe how a system changes from one state to another. The basic ideas of this model were developed by a Russian mathematician known as A. A. Markov (1856-1922). His primary subject of his research later became known as Markov chains and Markov Processes.

Markov models are useful mathematical concepts from probability and matrix algebra. Model chain is a sequence of discrete stochastic processes, events or actions which are purely used for predicting various future states. It is a memoryless system where the future state depends on the current state and is independent of past states. A stochastic event is a series of random elements which does not have a defined pattern (Shakraih 2013). Markov model is a user mobility model which can be modified to cope with user mobility in different wireless and mobile networks such as cellular networks, ad hoc networks etc.

Markov models can be used to compute the probabilities of events occurring by viewing them as states transitioning into other states or transitioning into the same states as before. However, a Markov model can only be used to predict finite values.

According to Szalka et al. (2009), the memory less property of the Markov process makes the Markov models easily applicable. In many models, the Markov processes states are based simply on the physical radio cells, i.e. one state represents one radio cell. In this case, any potentially present additional information in the user movements cannot be included in the model. A model, in which the states of the model are constructed according to a group of cells belonging to typical movement directions is proposed. The model takes

into account the user history during the prediction and merges the neighbor cells into a direction group depending on the user behavior and historical distribution and patterns, yet retaining the memory less property.

Markov chain model can be used for predicting the next state of a mobile device. For prediction of the next state model, discrete input values are usually needed. The input states are determined from the Cell ID/Wireless Fidelity which are discrete in nature. The input data will change according to the location of the mobile device. This will affect the state changes, which causes the Markov chain model to create transition probabilities that are represented in a transition matrix and thus the future state is predicted. The probability of obtaining a new state from the previous states can be observed from the data sets of states from the finite state machine and Markov chains. Over the years there has been a great revolution in the generations of mobile networks. Mobile network has greatly evolved and this has led to an increase in its usage. Its revolution has effectively touched virtually all areas in various sectors of life ranging from health sector, office, entertainment industry, educational sector etc. There has been a considerable increase in the number of mobile devices along with the use of wireless communication and a rapid development of location based social networks which collects a large amount of individual movement data. The ability of exactly predicting user's future location could provide more natural and customized services for Location based applications. Generally individual movement data consists of a series of geographic coordinated points stamped with time (Zheng et al. 2009). Location can be specified in geographic coordinates which can be represented in states also known as cell ids. Cell ids changes from one state to another and this process can be referred to as transition. In Markov Model the future state is independent on the past states and prediction is done using its current state. The predicted movement can then be used to increase the efficiency of location prediction. By using the predicted movement, the system can effectively allocate resources to the most probable to move cells instead of blindly allocating excessive resources in the cell neighborhood of a mobile user.

Qiao et al. (2015), carried out a research work on "User location prediction with energy efficiency model in the Long Term-Evolution network". The motivation of the research was to predict users' next location or place using an energy efficient technique which also provides the best tradeoff between prediction accuracy and resource consumption. The Objective of the research is to predict a users' location using two domain-independent algorithms, Markov Model and LZ algorithm. The methodology involved a comparative analysis been done between the two algorithms after it has been applied to various categories of users and the algorithm with highest level of accuracy is gotten. For the Markov Family, The order-k Markov predictor is independent of time, and it assumes that

the current location depends only on the previous  $k$  movements and for the LZ family, the algorithms of LZ family are often used for text compression; they are able to make real time predictions and do not need many resources. The limitation in this research is the inability to effectively group users based on their mobility pattern and spatio-temporal characteristics.

Herder et al.(2014), presented a research title “Predicting User Locations and Trajectories. The motivation of this work is gotten from the need of predicting locations that people will revisit as part of their routine patterns as they observed that Human mobility patterns contain strong regularities, people typically spend most of their time at and between a small number of locations and most followed trajectories (e.g the daily commute) also serve as starting points for visits to several other locations that form the long tail of a person’s whereabouts. The objective of this research is to predict user locations from data collected through GPS. The methodology involved analyzing, visualizing and discussing patterns found in a dataset of GPS trajectories. Further, the authors compare and analyze the performance of common prediction techniques that exploit the locations’ popularity, regularity, distance and connections with other locations. The limitation in this work was that trajectories considered were only those on the user’s current location.

Liu et al. (2016), presented a research on Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts. The objective of this research was to extend Recurrent Neural Network (RNN) and propose a novel method called Spatial Temporal Recurrent Neural Networks (ST-RNN) which can model local temporal and spatial contexts in each layer with time-specific transition matrices for different time intervals and distance-specific transition matrices for different geographical distances. The motivation of the research came from the fact that although so many methods have achieved satisfactory results in some applications, they are unable to handle continuous geographical distances between locations and time intervals between nearby behaviors in modeling sequential data. The methodology used in this research was that geographical distance and time between locations was modeled into the conventional RNN. Considering distance information is an essential factor for location prediction, it is necessary to involve it into our model. Similar to time-specific transition matrices, we incorporate distance-specific transition matrices for different geographical distances between locations. Distance-specific transition matrices capture geographical properties that affect human behavior. The limitation of this research is that the internal representation of space and time features was not extracted automatically it relied on handcraft representation.

Markov model uses a discrete stochastic process which are used for future state prediction. It can also be described as a memory less system that the future state depends on the current state and it is also independent of the past state.

## 2. METHODOLOGY

Network dataset chosen for the analysis and evaluation of this research is a 4G Long Term Evolution dataset with channel and context metrics.

**Channel:** This is measured throughput which is a combination of the eNodeB environment (load, scheduler policy), wireless channel characteristics and mobile device receiver capabilities. Additional information about the channel environment in addition to throughput values increases

accuracy and granularity which helps to give the model a more accurate prediction (Raca et al. 2018).

**Context:** The dataset provides additional context information such as GPS positions (longitude, latitude and velocity)

The dataset has device features:

- a) **Timestamp:** timestamp of sample
- b) **Longitude and Latitude:** GPS coordinates of mobile device
- c) **Velocity:** velocity in kph of mobile device
- d) **Operatorname:** cellular operator name (anonymized)
- e) **CellId:** Serving cell for mobile device
- f) **Network Mode:** mobile communication standard (2G/3G/4G)
- g) **RSRQ:** This represents a ratio between received signal ratio power (RSRP) and Received Signal Strength Indicator (RSSI). Signal strength (signal quality) is measured across all resource elements (RE), including interference from all sources (dB).
- h) **RSRP:** This represents an average power over cell-specific reference symbols carried inside distinct RE. It is used for measuring cell signal strength/coverage which helps in cell selection (dBm).

**RSSI:** This represents a received power (wideband) including a serving cell and interference and noise from other sources. RSRQ, RSRP and RSSI are used for measuring cell strength/coverage.

### Cells

Cells are also known as Mobility Access Points (MAP). These entities are able to communicate with the mobile nodes through the radio interface. These cells or Mobility Access Points have their own geographical area. Users can connect to Mobility Access Points with handovers from neighboring MAPs whereby each user is connected to only one MAP at a time. Mobility Access Points Neighbors could use different access technology than the current Mobility Access Point and can also be located at the very same geographical place as well. The Cell ID is a unique number that is used to identify the Base Transceiver Station (BTS) mainly the cell tower in which the mobile device is connected to the cellular networks such as GSM, GPRS, UMTS/HSDPA( Skariah, 2013). A cell ID is the basic method that is used to identify the location of a mobile of a mobile device.

A cell cluster is usually referred to a group of nearby cells where transitions occur within the cluster while Transition can be defined as the changing of one state to another state at a given period of time.

### 2.1 Heterogeneous Data

Dataset used in this model was heterogeneous in nature. Sources of trajectory data are diverse eg train, taxis, buses etc. These moving objects often have different sampling rates and movement patterns.

#### 2.1.1 Trajectory Data

Trajectory of a user can be defined as a set of check-ins because each check-in has a location tag and a timestamp respectively which usually corresponds to where and where the

check-in is made. (Cheng et al. 2015). Trajectory can also be defined as a time ordered sequence of locations. In a trajectory data, the basic characteristics are the locations and times of the moving objects. The following notations used are  $p = (x, y, h, t)$  representing a trajectory data point while  $x$  and  $y$  are the latitude and longitude of a given moving object, respectively,  $h$  is the the altitude and  $t$  is the time stamp (in many real world applications  $h$  is usually left out) thus  $p = (x, y, t)$  is frequently used. Trajectory data is usually composed of a sequence of trajectory data points in chronological order.

### 3. MARKOV MODEL

The Markov model used in this research is the order- $k$  (or “O( $k$ )”). This model assumes that the location can be predicted from the current context, that is, the sequence of the  $k$  most recent symbols in the location history  $(a_{n-k+1}, \dots, a_n)$ .

A user’s location history in this model is denoted by  $L = a_1, a_2, \dots$ , (1)

A substring  $L(i, j) = L(i, j) = a_i, a_{i+1}, \dots, a_j$  for an  $1 \leq i \leq j \leq n$ . is created. The user’s location is denoted as a random variable  $X$ .

If we let  $X(i, j)$  be a string  $X_i X_{i+1} \dots X_j$  representing the sequence of random variables  $X_i, X_{i+1}, \dots, X_j$  for any  $1 \leq i \leq j \leq n$

Then the context  $c = L(n - k + 1, n)$ .

Let  $A$  be the set of all possible locations. The Markov assumption is that  $X$  behaves as follows, for all  $a \in A$  and  $i \in \{1, 2, \dots, n\}$ .

$$P(X_{n+1} = a | X(1, n) = L) \quad (2)$$

$$= P(X_{n+1} = a | X(n - k + 1, n) = c) \quad (3)$$

$$= P(X_{i+k+1} = a | X(i + 1, i + k) = c) \quad (4)$$

The notation  $P(X_i = a_i, \dots)$  denotes the probability that  $X_i$  takes the value  $a_i$ .

#### 3.1 Data Preprocessing

One of the most critical steps in a data mining process is data pre-processing which deals with the preparation and transformation of the initial dataset thereafter relevant features is selected. Raw data is preprocessed in order to increase the ease of data mining and also to improve the quality and efficiency of data. Data preprocessing techniques used in this research are categorized as follows

##### 3.1.1 Data Cleaning

Data cleaning is the first step in data preprocessing which is used to find missing values, smooth noise data, recognize outliers and correct inconsistency (Alasadi and Bhaya 2017)

Real world data tend to be incomplete, noisy and inconsistent data cleaning routines attempts to fill in missing values, smooth out noise while identifying outliers, and correct inconsistencies in the data.

##### 3.1.2 Data Integration

This technique in data preprocessing works in a way such that data from multi and various resources are combined into one consistent data store like in a database or in a data warehouse. Data integration was used in this dataset because the dataset was heterogeneous in nature. In data integration, a number of things are taken into consideration some of which are Schema

Integration, Object Matching and Redundancy which are an important aspects.

##### 3.1.1.2 Data Transformation

This includes transforming the data to a suitable form that would be easy for mining. The process for data transformation is classified as follows:

- i. Smoothing: This is a process that removes noise from data. It involves techniques such as clustering, regression and binning.
- ii. Aggregation: This is a process that involves applying statistical metrics like mean, median and variance which are necessary to summarize the data.
- iii. Generalization: In This process, lower level data replaces higher level using hierarchical concepts.

##### 3.1.1.3 Data Reduction

These techniques can be used to reduce the representation of dataset in smaller volume with respect to maintaining the integrity of the original dataset. Strategies for data reduction include the following.

a. **Dimension reduction:** where irrelevant, weakly relevant, or redundant attributes or dimensions may be detected and removed.

b. **Data compression:** where encoding mechanisms are used to reduce the data set size.

c. **Numerosity reduction:** where the data are replaced or estimated by alternative, smaller data representations such as parametric models, or nonparametric methods such as clustering, sampling, and the use of histogram.

e. **Discretization and concept hierarchy generation:** where raw data values for attributes are replaced by ranges or higher conceptual levels.

Discretization and Concept hierarchies are powerful tools of data mining, in that they allow the mining of data at multiple levels of abstraction.

Preprocessed data is analyzed using Markov for probabilistic predictions, such that, if given a set of possible state  $S(t) = \{s_1, s_2, s_3, \dots, s_t\}$  which corresponds to the locations visited by the user at any given time  $t$ , the probability that the user moves from the current state (location) to a random next location is given by a conditional probability presented in equation (4)

$$p_{ij} = P(s_{t+1} = j | s_t = i) \quad (5)$$

where  $p_{ij}$  represents the transition probability of user’s movement (trajectory) from current location  $s_t = i$  at time  $t$  to next possible location  $s_{t+1} = j$  at time  $t + 1$ .

The user’s next location at  $m$  step transition probability given an initial state  $i$ , is presented as follows:

$$P(s_m = j | s_0 = i) = \sum_{k=1}^n p_{ik} (m - 1) p_{kj}$$

where  $s_m$  represents the multiple step transitions to location  $j$ ,  $s_0$  represents the initial state,  $p_{ik}$  represents transition probability from state  $i$  to the state before the last time step  $(m - 1)$ , and  $p_{kj}$  is the transition probability of trajectories from state  $k$  (state before the last state) to the predicted state  $j$ . The probable outcome from  $i$  to  $j$  obtained from equation above becomes Output of the model.

A **stochastic process** in discrete-time is a family,  $(X(n))_{n \in N_0}$ , of random variables indexed by the numbers  $N_0 = \{$

0,1,2,3.....}. The possible values, S of X (n) variables indexed by the numbers are referred to as the state space of the process (Tolver, 2016)

In Stochastic processes, Markov chains are an important concept. They can be used to greatly simplify processes that satisfy the Markov property, namely that the future state of a stochastic variable is only dependent on its present state. This means that knowing the previous history of the process will not improve the future predictions - which of course significantly reduces the amount of data that needs to be taken into account. Mathematically, Markov chains consist of a state space, which is a vector whose elements are all the possible states of a stochastic variable, the present state of the variable, and the transition matrix. The transition matrix contains all the probabilities that the variable will transition from one state to another, or remain the same. To calculate the probabilities of a variable ending up in certain states after n discrete partitions of time, one simply multiplies the present state vector with the transition matrix raised to the power of n. There are different types of concepts regarding Markov chains depending of the nature of the parameters and application areas. They can be computed over discrete or continuous time. The state space can vary to be finite or countable infinite and depending on which, behave in different ways. A Markov chain with a countable infinite state space can be stationary which means that the process can converge to a steady state.

### 3.2 Clustering

Clustering of cells is a group of nearby cells where most transitions happen within the cluster.

Clustering is one of the most used exploratory data analysis technique across and it is mostly used to get an intuition about the structure of the data, for finding meaningful groups and also feature extraction and summarizing (Trivedi et al., 2011). Given a space, clustering of a location can be thought of as a partitioning of locations of a user into K parts which is  $f: \rightarrow \{1 \dots, K\}$ . This partitioning is done by optimizing some internal clustering criteria such as the intra-cluster distances. The value of Kis found usually by employing a second criterion that measures the robustness of the partitioning. Locations are identifiable in that users can reliably enter and leave them. K-means inbuilt algorithm is used to cluster the coordinates into distinct clusters.

## 4. RESULTS

Using the four (4) dataset mobility pattern (Bus, Car, Train and Pedestrian) in this work gave us various ranges of performance in the model. Performance metrics used are Prediction Accuracy, Root Mean Square Error, Mean Absolute Error and Computational Time. Results gotten in this model shows that in the Pedestrian dataset mobility pattern, since the speed of the Pedestrian was slow, the level of accuracy was on the high side. Similarly the Bus dataset Mobility dataset yielded a lower level of prediction accuracy as compared with the pedestrian Mobility dataset and this was due to the fact that the bus has a higher speed level than the pedestrian. Also when the car and train mobility pattern dataset were run in the model, the prediction accuracy gotten from the train mobility pattern was lower than that of the car mobility pattern dataset because the speed at which the train moves is more than that of the car. The error rate also increased tremendously in the train mobility pattern dataset as compared to the other mobility pattern datasets as a result of the speed of the train which is faster than others.

**Table 1: Prediction Evaluation for Bus Dataset**

Model	Comp Time (Secs)	Prediction Accuracy (%)	Root Mean Squared Error	MAE
Markov	54	97.54	0.024522	0.01827

**Table 2: Prediction Evaluation for Car Dataset**

Model	Comp Time (Secs)	Prediction Accuracy (%)s	Root Mean Squared Error	Mean Absolute Error
Markov	313	87.11	0.12808	0.59179

**Table 3: Prediction Evaluation for Train Dataset**

Model	Comp Time (Secs)	Prediction Accuracy (%)	Root Mean Squared Error	Mean Absolute Error
Markov	1005	39.21	0.66100	0.33211

**Table 4: Prediction Evaluation for Pedestrian Dataset**

Model	Comp Time (Secs)	Prediction Accuracy (%)	Root Mean Squared Error	Mean Absolute Error
Markov	48	98.21	0.01808	0.014504

## 5. AREAS OF APPLICATION

This work is useful in prediction and tracking; which can be used in :

- (1) Advert placement firms: It helps to place deals and adverts based on the prediction of users in that location.
- (2) Detective organizations like the Force can also make use of this work as it can assist in tracking down suspects.

## 6. CONCLUSION

This research has been able to predict various user location with Markov Model using the long term evolution network datasets. In Future the scope of the research will be expanded using other dataset and covering a wider scope of the Markov Model.

## 7. REFERENCES

- [1] Alasadi, S. A.,and Bhaya, W. S. (2017). Review of Data Preprocessing Techniques in Data Mining. Journal of Engineering and Applied Sciences, 12(16), 4102-4107.
- [2] Cheng, M., Liu, Y., and Yu, X. (2015). Predicting next locations with object clustering and trajectory clustering. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 344-356). Springer, Cham.
- [3] Herder, E., Siehndel, P., and Kawase, R. (2014, July). Predicting User Locations and Trajectories in

- International Conference on User Modeling, Adaptation, and Personalization (pp.86-97). Springer, Cham
- [4] Liu, Q., Wu, S., Wang, L., and Tan, T. (2016, February). Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts. In AAAI (pp. 194-200).
- [5] Qiao, Y., Yang, J., He, H., Cheng, Y., and Ma, Z. (2016). User location prediction with energy efficiency model in the Long Term Evolution network. *International Journal of Communication Systems*, 29(14), 2169-2187.
- [6] Raca, D., Quinlan, J. J., Zahran, A. H., and Sreenan, C. J. (2018). Beyond throughput: a 4G LTE dataset with channel and context metrics. In *Proceedings of the 9th ACM Multimedia Systems Conference* (pp. 460-465). ACM.
- [7] Skariah, E. (2013). Mobile Phone Context Prediction Using Markov Chains.
- [8] Szalka, T., Szabo, S., and FÜLLÖP, P. (2009). Markov model based location prediction in wireless cellular networks. *Infocommunications Journal*, 40.
- [9] Tolver, A (2016). An Introduction to Markov Chains.
- [10] Trivedi, S., Pardos, Z. A., and Heffernan, N. T. (2015). The utility of clustering in prediction tasks. arXiv preprint arXiv:1509.06163.
- [11] Zheng, Y., Zhang, L., Xie, X., and Ma, W. Y. (2009). Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th international conference on World wide web* (pp. 791-800). ACM.