A comprehensive review of hidden Markov model applications in prediction of human mobility patterns

Neha Rajawat*

Career Point University, Kota,
Rajasthan, 324005, India
Email: nrhada@gmail.com
*Corresponding author

Navneet Gupta

J.L.N.T.T. College, Kota,
Rajasthan, 324008, India
Email: ngnavneetgupta@gmail.com

Soniya Lalwani

Bal Krishna Institute of Technology, Kota,
Rajasthan, 325003, India
Email: slalwani.math@gmail.com

Abstract: Proposed work reviews the research and development of HMM, in finding out the mobility patterns and keeping the main focus on the state-of-the-art. Concepts related to Markov chains are explained and then the ideas are extended to the class of HMMs using several simple examples. The mathematics of the HMM is presented, beginning with the Markov chain and then including the three main constituent algorithms: the Viterbi algorithm, the forward algorithm, and the Baum-Welch. Throughout the text, the description of the theory is intertwined with real-world applications.

Keywords: hidden Markov model; human mobility pattern; Viterbi algorithm; forward algorithm; Baum-Welch algorithm.

Reference to this paper should be made as follows: Rajawat, N., Gupta, N. and Lalwani, S. (2021) ‘A comprehensive review of hidden Markov model applications in prediction of human mobility patterns’, *Int. J. Swarm Intelligence*, Vol. 6, No. 1, pp.24–47.

Biographical notes: Neha Rajawat is a PhD scholar from the Career Point University, Kota, MSc in Mathematics and BSc in Mathematics from University of Kota, Kota. She has a teaching experience of three years.

Navneet Gupta is a Mathematician. For the last 25 years, he has been teaching Mathematics and Pedagogy of Mathematics to various sections of students’ fraternity. The best part about his journey is that he has never finished his learning and always been a student. While being an established academician, he pursued degrees and certificates like MSc, MPhil, GATE, NET (Mathematics), MEd, MAEd and PhD.

Soniya Lalwani holds a Postdoctorate from the Science and Engineering Research Board (SERB), DST during September 2016–September 2018. She implemented her postdoctorate project at the Department of Computer Science, RTU, Kota. She received her PhD from the Department of Mathematics, MNIT, Jaipur. Currently, she is working as an Associate Professor with the Department of Mathematics, BKIT, Kota. She has published more than 35 research papers in reputed journals and 15 research papers in conferences. She has over 14 years' job experience at various research and teaching positions. Her research areas include swarm intelligence: particle swarm optimisation, multi-objective optimisation, bioinformatics: multiple sequence alignment of DNA/RNA sequences, sequence-structure alignment, ABS algorithm and clinical/medical biostatistics.

1 Introduction

These years, a lot of data are observed related to geological locations representing human mobility and have enabled the research community to find patterns related to individual and grouped mobility of humans. Also, to generate such models that can learn and predict spatiotemporal patterns and structures in a path formed by this mobility. This research is needed for applications and services like concluding migratory flows, predicting future locations, forecasting the large-scale disaster, to trace taxicab location, traffic forecasting, urban planning, and prototyping the epidemics. From the term mobility, the change in human location is referred to the time and space, thus implicitly targeting human mobility only. There was a time when the human movement was primarily due to climate change events, non-habitable environment, food, and water crisis, inter or intra national crisis, but in modern era, socio-economic factors like working environment, living conditions and globalisation are playing a vital role. From prehistoric times to the present day, people tend to move for livelihood and this movement is daily. Such trips can be performed for social and recreational activities also.

Above mentioned movements have a comparatively shorter time and have spatial domain than the migratory flow. These shorter trips represent a regular and periodic pattern of human life. In a large population, such a pattern symbolises some important aspects of day to day human life and environmental conditions surrounding them. Researchers have found that 15%–25% of total household expenditure amounts to transportation charges of an average household, thus sharing the second largest chunk after the housing. Increasing fossil fuel consumption causes the big problem of greenhouse gas emissions. Hence, human mobility is impacting both human life and the surrounding environment. To understand this impact completely, it is quite important to learn these mobility patterns (Barbosa et al., 2018). Some dynamic models are needed for learning, understanding and predicting this human mobility. In addition to this, these models should show the required flexibility for changing scenarios like special and temporal changes the data provided. Human beings show different movement scenarios concerning their age (children, working professionals, housewives and elderly), time (day, week and year), surroundings, etc. Also, the dataset may not be revealing all these attributes explicitly but provided as some hidden features. Hidden Markov models are famous for their applications with the dynamic situation-based data such as temporal-latent datasets of human mobility. Some of the examples of the applications are speech recognition, text recognition, natural language processing, DNA sequencing, handwriting identification and plan recognition.

The HMM is a sequence model. A sequence model or sequence classifier works on the principle of direct mapping between the observation and their respective labels. It assigns a definite label to the observation (Zion and Lerner, 2018). HMM is a probabilistic sequence model: for the given observation or entities (letters, words, morphemes, sentences, etc.) model calculates a probabilistic distribution over possible combinations of labels and predicting the most accurate sequence of labels.

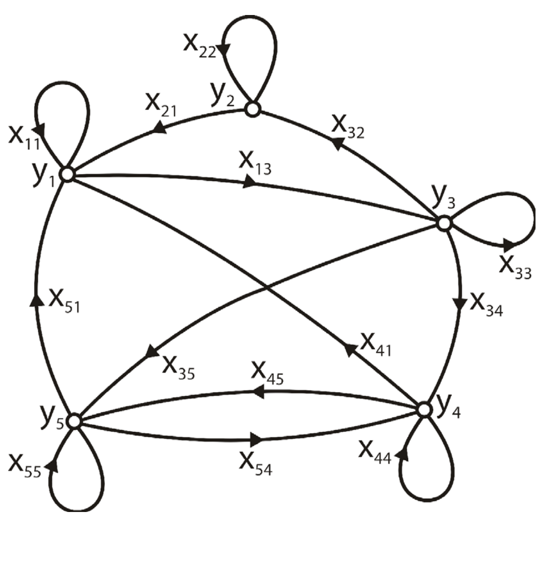
Proposed work presents a comprehensive review of the HMM implemented to find out the mobility patterns. Concepts related to Markov chains are explained and then the ideas are extended to the class of HMMs using several simple examples the mathematics of the HMM is presented, beginning with the Markov chain and then including the three main constituent algorithms: the Viterbi algorithm, the forward algorithm, and the Baum-Welch or EM algorithm for unsupervised (or semi-supervised) learning.

The classification of the manuscript is as follows: Section 2 presents the introduction of the hidden Markov model. Section 3 presents the role of HMM in human mobility prediction. Section 4 contains the literature survey with category-wise classification, followed by the conclusions in Section 5.

2 Hidden Markov model

In the field of speech and language processing, HMM is one of the most significant models. Markov chain is required to be described, sometimes called the observed Markov model. A Markov chain dictates the probability assigned for unambiguous sequences. Assume a system of distinct n states as y_1, y_2, ……, y_n described at any particular time as in Figure 1 (for simplicity $n = 5$).

Figure 1 A Markov chain with five states with selected state transitions



This particular system changes concerning the discrete-time intervals, also the probability associated with the state differs. Assuming the time instances are represented by $t = 1, 2, \ldots$ and at time t actual state is represented by ϕ_t. A general probabilistic description of the above system requires specifying the current state (at time instance t) and all the previous states. To describe a special case of a first-order discrete Markov model needs probabilistic distribution of current and immediate previous state.

$$P[\phi_t \mid \phi_{t-1}, \phi_{t-2}, \ldots] = P[\phi_t \mid \phi_{t-1}] \text{ where } \phi_t = y_j, \phi_{t-1} = y_i, \phi_{t-2} = y_k \tag{1}$$

In addition to this, only those processes are useful which are independent of time. Thus, concluding following representation for state transition probabilities x_{ij} of the form:

$$x_{ij} = P[\phi_t \mid \phi_{t-1}] \text{ where } \phi_t = y_j, \phi_{t-1} = y_i \text{ and } 1 \leq i, j \leq n \tag{2}$$

State transition coefficients follow below properties:

$$x_{ij} \geq 0 \tag{3a}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \tag{3b}$$

As can be seen, they are following standard stochastic properties, given stochastic process could be described as an observable Markov model where outcome of the given process is set of states at a given instance of time, also each state represents a physically (observable) event. Based on this idea, a three-state Markov model representing the weather condition of Manali (Hill Station) can be assumed. Weather condition during a day could lie among the following three states:

State 1 sunny (y_1)

State 2 rainfall (y_2)

State 3 snowfall (y_3).

One of the three states above can describe weather condition on a day and below matrix X provides the probabilities of transition between three states:

$$X = [x_{ij}] = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}$$

Provided the weather on day-1 ($t = 1$) is snowfall (state 3), one can ask 'what will be the probability (according to the given model) of next seven days weather conditions to be 'snow-sun-sun-snow-snow-sun-rain?'' The query can be represented as observable sequence O where $O = \{y_3, y_3, y_1, y_1, y_3, y_3, y_1, y_2\}$ for $t = 1, 2, \ldots, 8$. The probability of transition of the sequence O can be expressed as below:

$$
\begin{aligned}
P(O \mid Model) &= P[y_3, y_3, y_1, y_1, y_3, y_3, y_1, y_2 \mid Model] \\
&= P[y_3].P[y_3 \mid y_3].P[y_1 \mid y_3].P[y_1 \mid y_1].P[y_3 \mid y_1].P[y_3 \mid y_3].P[y_1 \mid y_3].P[y_2 \mid y_1] \\
&= \prod_3 x_{33}.x_{31}.x_{11}.x_{13}.x_{33}.x_{31}.x_{12} \\
&= 1(0.6)(0.2)(0.8)(0.1)(0.6)(0.2)(0.1) \\
&= 1.152 \times 10^{-4}
\end{aligned}
$$

where the notation to denote the initial state probabilities is:

$$\Pi_1 = P[\phi_1 = y_i], 1 \le i \le n \tag{4}$$

One more interesting question can be asked: provided the current state of the model, what is the probability of staying in the same state for next w days? This sequence can be described as… $O = \langle y_{i1}, y_{i2}, y_{i3}, y_{iw}, y_{iw+1} \rangle$ where y_i is the current state. The probability for this model is as follow:

$$P(O \mid Model, \phi_1 = y_i) = (x_{ij})^{w-1}(1 - x_{ij}) = p_i(w). \tag{5}$$

Here, $p_i(w)$ describes the discrete probability density function (PDF) for the duration of w days. This exponential duration density is characteristic of the state duration in a Markov chain. If PDF $p_i(w)$ is given, then one can estimate the number of observation or time duration required for the defined initial state as:

$$\bar{w}_i = \sum_{w=1}^{\infty} w p_i(w) \tag{6a}$$

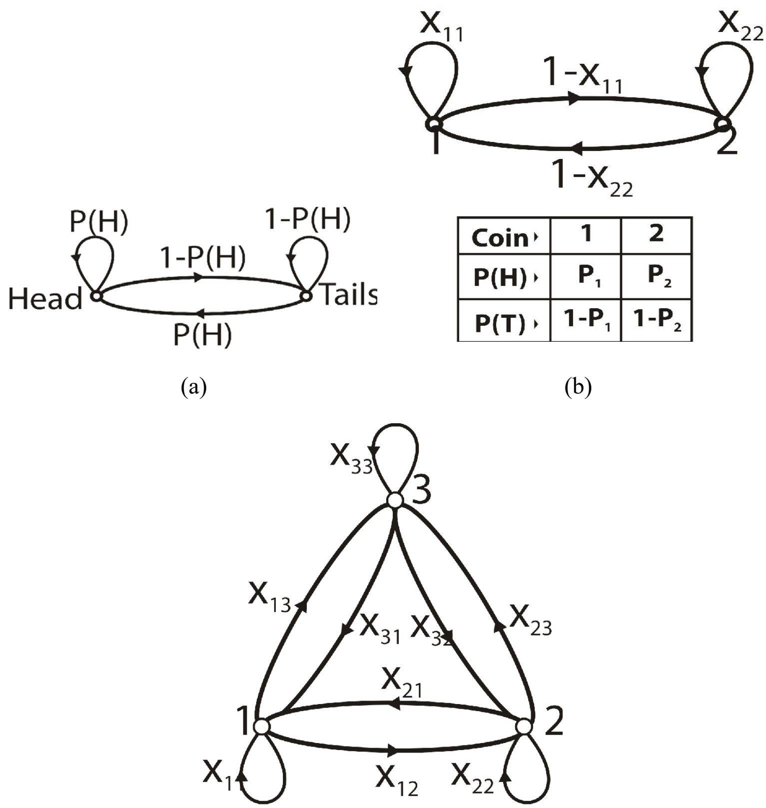
$$= \sum_{w=1}^{\infty} w(x_{ij})^{w-1}(1 - x_{ij}) = \frac{1}{1 - x_{ij}} \tag{6b}$$

Hence, the expected number of continuous day for sunny weather (as per the above model) are 1/0.2 = 5, 2.5 days for snowfall and 1.67 for rainfall. Current model puts too much restriction to apply to many interesting events. The concept of Markov models covers non-observable events which can be described from another set of stochastic processes that can produce a sequence of observable events, finally converted into (hidden Markov model) a doubly embedded stochastic process, where observations are not discrete but probabilistic. Consider the following model of some simple coin-tossing examples. A random sequence of consecutive head and tail is generated by tossing coin represented as:

$$
\begin{aligned}
&H = \text{heads and } T = \text{tails} \\
&O = O_1 O_2 O_3 \ldots\ldots O_T \\
&\quad = HHTTTHTTH \ldots\ldots H
\end{aligned}
$$

The actual problem is to construct an HMM representing the generated observable sequence of possible outcomes are mentioned in Figure 2.

Figure 2 HMM for coin tossing experiment



Coin	1	2
P(H)	P_1	P_2
P(T)	$1\text{-}P_1$	$1\text{-}P_2$

(a) (b)

Coin	1	2	3
P(H)	P_1	P_2	P_3
P(T)	$1\text{-}P_1$	$1\text{-}P_2$	$1\text{-}P_3$

(c)

For coin tossing experiment, three possible hidden Markov models can be counted which are:

a one-coin model

b two-coins model

c three-coins model.

In the one-coin model, one face is declaring about the corresponding states in the model and then declaring about the remaining states should be in the model. If one considers that the single biased coin is tossed, then this situation could be with a two-state model where every state represents either head or tails. This model is delineated in Figure 2(a). It is the case of observable Markov model. An HMM similar to that of Figure 2(a) would be determined to one-state model, where the bias of coin is unknown parameter and state refers to the single biased coin. Figure 2(b) explaining the observed outcomes, which is the second form of HMM (two-state model). This model corresponding to two biased coins. In these two states, every state refers to a different biased coin being tossed. Every state is characterised with the help of probability distribution of outcomes, and conversions between states are characterised by a state conversion matrix. The physical mechanism for selecting the state conversions could be a group of independent tosses or any other probabilistic event. Figure 2(c) explaining the observation outcome which is the third form of HMM (three-state model). This model is related to three biased coins, and the choice will be from these three, after referring some probabilistic event, the three models presented by Figure 2 explaining the observable sequence of heads and tails, an obvious question would be to choose the model which can predict the best possible observation one-coin model of Figure 2(a) contains only one unknown parameter, the two-coin model of Figure 2(b) has 4, and the three-coin model of Figure 2(c) has nine unknown parameters. Thus, with the greater degrees of freedom, the larger HMMs would seem to inherently be more capable of modelling a series of coin tossing experiments than would equivalently smaller models. Although theoretically seems to be true, real scenarios will put some limiting constraints on the considerable model. Furthermore, the choice of the model will depend on the physical event also. Three-coin model will not be a good choice if a single coin is tossed. Such a choice will depict a mismatching between the physical event and the model chosen.

2.1 Elements of an HMM

The above examples give a pretty good idea of what an HMM is and what are the possible use cases for its applications. The elements of an HMM are formally defined, and the process of model generation of observation sequences is explained.

An HMM is represented by the following:

1 In this model, M is the number of states. States (or set of states) of the HMM are hidden for most of the real-life scenarios and states have some physical reference attached to them. For example, when the coin is tossed in some experiment, every state is associated with a distinctly biased coin. Generally, an interconnection can be observed between these states such that any state can be obtained using any other state (e.g., an argotic model); however, different types of interconnections among the states present different interesting findings. The individual states are denoted as $y = \{y_1, y_2, \ldots\ldots\ldots, y_M\}$, and the state at time t as ϕ_t.

2 N represents the numbers of unique observable symbols per state, i.e., the discrete alphabet size. These observable symbols are physical outputs of the experiment or system to be modelled. For the coin-tossing example, head and tails are these observation symbols. The individual symbols are denoted as $Q = \{q_1, q_2, \ldots\ldots\ldots, q_N\}$

3 The state transition probability distribution $X = \{x_{ij}\}$ where

$$x_{ij} = P\{\phi_{t+1} = y_j \mid \phi_t = y_i], \quad 1 \le i, j \le M \tag{7}$$

For the special cases, where any state can reach any other state in a single step, $x_{ij} > 0$ for all i and j.

For other types of HMMs, $x_{ij} = 0$ for one or more (i, j) pairs.

4 The observation symbol probability distribution in state j, $A = \{a_j(r)\}$, where:

$$a_j(r) = P\left[q_r att \mid \phi_t = y_j\right], \quad 1 \le j \le M, 1 \le r \le N \tag{8}$$

5 The initial state distribution $\Pi = \{\Pi_i\}$ where $\Pi 1$ is determined by equation (4).

Given appropriate values of M, N, X, A, and HMM can work as a generator to provide an observation sequence $O = O_1 O_2 \ldots\ldots\ldots O_T$, where each observation O_t, belongs to Q and duration of the generated sequence will be represented by T. Algorithm for such HMM should work as follows:

a Select the first state $\phi_1 = y_i$ the initial state distribution Π reported in the first state.

b Choose $t = 1$.

c Select $O_t = q_r$ according to the symbol probability distribution in the state y_i, i.e., $a_i(r)$.

d Switch to a new state $\phi_{t+1} = y_j$ as per the state transition probability distribution for state y_i, i.e., x_{ij}.

e Set $t = t + 1$; back to point C if $t < T$; otherwise, close the procedure.

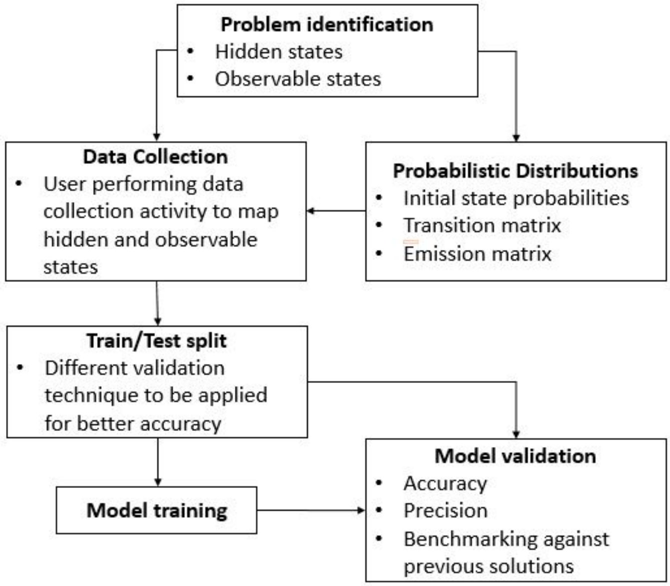
The method described above can be utilised for two purposes, first to generate observations and second being a model which can find out the generation algorithm used by some other HMM, for a given sequence of observations. From the above discussion, it is clear that an HMM can be described using two model parameters (M and N), properties related to observation symbols, and the properties of the three probability measures X, A and Π. For simplicity, a compact description $\lambda = (X, A, \Pi)$ can be used, where λ is a complete set of parameters of the model (Rabiner, 1989).

3 Human mobility prediction using HMM

Movement of human-being provides a data series that varies concerning time. The particular location of a person can be mapped with this function of time. Each individual in this space draws different patterns of period mobility which can be represented as a probabilistic transition among a set of geographical locations. This concept motivates to use HMM for mobility prediction (Ryu and Song, 2013).

The basic requirement to model an HMM is to define sets of hidden and observation states and transition and emission probabilities respectively. Generally speaking, geographical locations are always associated with some physical landmarks. Thus, these geographical locations can be used as observation while physical landmarks will be the associated hidden state. The example of mobile communication can be taken, where mobile devices need to connect with a suitable base station to receive or deliver the message. As the mobile user moves in 3D space communication continuity has to be maintained with a smooth transition between two base stations. Figure 3 provides a general overview to model an HMM for the prediction of human mobility.

Figure 3 General overview of HMM to predict human mobility



Dataset of human positions composed of both geographical locations and respective timestamp thus needs multidimensional processing. As it is known that a multidimensional dataset can be easily processed by HMM (Yang et al., 1997), it can be used as a candidate for modelling human mobility which can represent complex models. Process of predicting human mobility pattern using HMM is explained further:

- *Problem identification:* Task is to design an HMM which can predict the next physical location (hidden state) using the probabilistic transition between geographical location (observation state).
- *Data collection and probabilistic distributions:* According to the task type, different datasets will be prepared which maps the relation between different geographic locations and physical landmarks. These relations are represented as the transition and emission matrices.
- *Train/test splitting of dataset:* This step depends on the implementation of the algorithm and accuracy desired from the algorithms. One can go for 60:40 rule or 80:20 rule, or *K*-fold cross-validation can be done to achieve expected results.
- *Model training and solution benchmarking:* Against previous solutions available for human mobility pattern analysis.

In order to design an efficient HMM probabilistic distribution defined for the hidden and observation states plays a major role. Apart from this, a nicely collected data always plays a significant role.

3.1 Importance of human mobility analysis

From the term mobility, the change in human location is referred to, both individuals and groups, concerning the time and space thus implicitly targeting human mobility only. As per research titled 'New world' by the Europeans, around 70,000 Homo sapiens migrated from Africa, and are living around the world now in present times. Thus, it can be seen that the existence of human beings is always affected by their movement.

Today, it is possible to record every location an object, animal, or human has visited during a period, such as a day, week, or year. From these records, one can try to perceive their (object/animal/human) mobility pattern and seek answers to queries about their type of human or animal. If the answer is human, then, what kind of human are they? Are they students, working or any other? What are the characteristics of the human mobility pattern? Are there any changes amid the mobility pattern of persons or communities in distinct time?

Mobility models of humans are one of the essential knowledge for mobile computing, location-based service, and any other related fields. According to the extension of cities and evolution in lifestyle, the human mobility models progressively being intricate and the comparison finds that the frequency and intensity of natural disasters have substantially increased and its tendency is foreseen to continue. Interfacing this unpredicted disaster correctly forecasting the emergency behaviour of humans and their mobility will become problematic for effective planning of disaster management and societal reconstruction to provide those best facilities. For example, a mobile ad hoc network requires a realistic mobility model.

Human mobility models can be used for various scenarios such as virus attack on mobile phone devices, spreading of infectious disease, the movement pattern of adolescents, movement due to harmful environment, etc. The explanation of these mobility patterns needs several parameters. The human mobility pattern is individual as well as aggregated where an individual pattern mostly depends on personality. For example, people prefer highways instead of local roads. Social factors also play an important role in human mobility, such as students try to get a home near to their study centre mostly preferring a 1 km radius. A military battalion presents an example to a group or collective mobility (Song et al., 2010).

4 Literature review

Below explained work illustrates selected applications of human mobility modelling, demonstrating the insights the field can bring to the solution of real-world problems.

4.1 Human mobility: models and applications

A detonation of the vast geo-located dataset associated with human mobility has been observed from the last few years. It empowered researchers to quantitatively explore the mobility pattern of the individuals and collective to produce techniques that can accept and regenerate the spatiotemporal structure and regularities in human trajectories. The study of human mobility is notably crucial for an application like estimation migration flows, traffic statements, urban planning and pandemic modelling. Barbosa et al. (2018) analyse the approaches regarding development to regenerate distinct movability methods with the most popular one. This study can be used for knowing the basic modelling theory of human mobility and the accumulation of technical strategies applicable to particular mobility repeated issues. The review stables the topic by differentiating between a person and populace mobility, and also, between a lower limit and upper limit mobility. Ubiquitously, the text detail of the theory is interwoven with real-world applications. Mobility has various meanings. According to them, the movement of persons in place and time undoubtedly adduces to human mobility. The presence of persons has always been intertwined with their movement.

At the current time, socio-economic factors are playing a huge role. Environmental conditions and human life is widely affected by the daily movement of such a large and increasing population. The first work (Olsson, 1965) was conducted to show the effect on the mobility between two cities which is proportional to their population size and inversely proportional to the distance separating the cities. After a few decades, Ravenstein (1885) evolved and published the idea in a seminal work, presenting his laws governing the migration. After this, Stouffer (1940) and Bright and Thomas (1941) refined his theme and published his law of intervening opportunities. The laws are:

- Law of migration: Ravenstein (1889) was the first investigator who attempted an interpretation and forecasting of migration patterns, contributed to making a rough calculation of a maximal global population depending on resource consumption. With the help of this, he analysed the socio-economic factors and distance constraints, which are the main ingredients for the modern population movement. He developed seven laws of migration which stimulate an enormous volume of work (Ravenstein, 1889).
- Law of intervening opportunities: Ravenstein's (1885) law of migration was enhanced by Stouffer (1940). According to his observations, people tend to cover any given distance number for times if the number of opportunities are present and intervening opportunities decrease the number of distance coverage. Based on these observations, he has given below mentioned formula:

  $$\frac{dg(r)}{dr} \propto \frac{1}{f}\frac{df(r)}{dr} \tag{10}$$

  Here, $g(r)$ is the accumulation number of migrates that move a distance r from their original location and $f(r)$ is the accumulation number of intervening opportunities. Hence, f is a continuous function of distance, after integration of equation (10):

  $$g(r) = \log f(r) + K \tag{11}$$

  Here, K is a constant denoting number of opportunities at the original location. So, the relation between mobility distances is indirect and established only through an accessory via intervening opportunities.
- Distance decay and gravity law: In George Kingsley Zipf's (1946) law, the frequency of a word ranked c. In term of usage has a statistical dependence $f_c \approx \frac{1}{c}$ (Cornford, 1936). He found that relation can expand to other regions of society especially the size of cities (Zipf, 1940), where P is a population of the city and c is rank which follows the below relation:

$$P_c \approx \frac{1}{c^{\alpha}} \tag{12}$$

Part of any centre *i* in the dynamism of cargo is proportional to its population P_i, so the dynamism of wares between two centres is proportional to the product of their populations and this dynamism is inversely proportional to the distance between centres. Hence, the relation is:

$$O_{ij} \propto \frac{P_i P_j}{R_{ij}} \tag{13}$$

Here, O_{ij} is the dynamism of wares between two population centres i, j and R_{ij} is the distance between two centres. They concentrated on the latest work started with the type of information used in a lot of research projects, followed by a description of metrics and models of individual mobility and population mobility. They illustrated a list of selected applications of human modelling and concluded that the internal area can bring to the solution of real-world issues. Hence, study, understanding, and modelling of human mobility are precedence to many fields of society.

4.2 Using photographic data

Barchiesi et al. (2015) present a study, in which individuals move forward around their zone is a result of a large extent of issues as well as planning of adequate transportation channel and that includes devising of ophidian and structure of adept transportation channel. They collected information about the location and time of approximately 16,000 people who uploaded the geo-tagged image to the flicker photo-sharing website from inside places of the UK. Influenced by the principal of Levy's flight, which has been earlier used to represent the statistical qualities of human mobility, they probably prepare a machine learning algorithm for estimating the probability of speed in an interaction between individuals in geographical locations and pairs of locations. Their conclusions are in the flow of traveller between the pair of major cities and official figures in general agreement in the UK. Origin can be used to measure human mobility on a large-scale. Human mobility is administered by people's judgement, nature, and events of life. So far, whenever the population was evaluated in a large area, mobility has displayed a notably statistically regular figure. These have been surveys in various references such as urban mobility, anthropology (Raichlen et al., 2014), crime modelling (Chaturapruek et al., 2013), planning (Rozenfeld et al., 2008; Berlingerio et al., 2013; Botta et al., 2015), advertising (Aalto et al., 2004) and epidemic dissemination (Hufnagel et al., 2004; Colizza et al., 2007; Gonçalves et al., 2013). In a few years, acceptance of new interaction channels like mobile phones and online social media has progressed sharply. It has unwound the favourable circumstances to generate upgraded models of human mobility.

It permits human to better perceive the effect of the social network on how to move forward (Wang et al., 2011; Musolesi and Mascolo, 2007) and where people can move in the future (De Domenico et al., 2013; Hawelka et al., 2014). Current information sources have empowered research on human mobility and common human nature (Lazer et al., 2009). Preceding work on information from mobile logs and recorded mobility of banknotes has informed that the tour length of tourist follows a power-law-distribution (Hufnagel et al., 2006; Gonzalez et al., 2008).

According to this, it leads to a large number of small movements, followed by large ones; sometimes, it is related to Levy flight. Here, they scoop information from the photo-sharing flicker's website and start the examination, the movement is grouped into geographic groups. They argue that it is conformity with the disappearance of groups or references that justify this need for Levy flight (Jeung et al., 2007). Their exemplary is dependent on the contemplation of trajectories of 16,000 photographers and they permit to estimate the general pattern from the nature of many human beings. Their strategy is induced by the search. Human mobility follows the global statistical pattern, mathematically explained by Levy's flight. This pattern can be modelled using techniques of machine learning.

In the UK, the common pattern of human mobility is estimated to the flow of travel only from information and necessarily knowing the map, while their assessment is sometimes problematic. Comprehensive shortage in the official survey of mobility, their available findings appear at the level of the country in the general agreement with the evidence. Arranging a novel statistical tool for the investigation of online data sources and mobilising evidence that can use online data to estimate the human journey. The investigation represented here can also be expanded by assessing different static parameters.

4.3 Next place prediction using MMC

Several authors (Asahara et al., 2011; Ashbrook and Starner, 2002; Ying et al., 2011) discussed the problem of forecasting the future location of a person supported by the contemplation of his mobility nature over the little period and present location that they visited. In their work, many possible scopes like an assessment of geo privacy solitude system, the evolution of area-based services and the design of location-aware proactive resource migration are discussed. They enlarged a mobility model entitled mobility Markov chain. In order to include n places visited in the past, they propose a novel algorithm based on mobility to predict the next place, i.e., n-MMC. The application of three different data shows that the precision of the forecast is 70% to 95%. Also when the precision of the prophecy increases with n, choosing $n > 2$ does not seem to significantly improve (Gambs et al., 2012).

4.4 Predicting future locations

The online location-based services and geo-positioning technologies are playing a vital role to dissect human location histories that permit users to share their data. Tasks like the prediction of human movement can be labelled by using this information. Mathew et al. (2012) introduce a hybrid method named hierarchical triangular mesh (HTM) for predicting human mobility based on HMMs (Rabiner, 1989). This approach clusters the location history according to mobility characteristics. They represent the analysis of a generative model that includes the sequential relation between places visited by a particular individual in a specific period. It supports the logical judgement of statistical patterns for predicting future locations to be visited. They introduce a hybrid method (HTM) based on HMMs which initially clusters the human location histories in sequence with their characteristic (Lu and Tseng, 2009). By the potential gain on HMMs of modelling the sequence visits, they found that the proposed technique will account for location characteristics as imperceptible parameters and also with the essence of every individual's previous actions. They conducted their experiment using a real-world example dataset of locations, collected from the Geo-Life Project and calculated the prediction accuracy of the proposed method using different configurations. They measured a prediction precision of near about 13.85% with the best performing technique (Farrahi and Gatica-Perez, 2011). In terms of the geospatial distance between the actual location of the user, and the geospatial coordinates that are predicted, the best results correspond to an average distance of 143.506 kilometres, and a median distance of 4.957 kilometres (Mathew et al., 2012).

4.5 From taxicab location

Taxicabs are getting popular with the real-time spot perceive tools. Thus, the place can be traced with the knowledge of the location and can be used for better city arrangement, pricing according to the crowd and the process of taxicabs. One vital issue is to recognise the human mobility pattern by locating the taxicab to entitle these experiments to identify pickup and drop points for special travel. Ganti et al. (2013) are very inefficient in recognising the two-sided journey while the previous outlook is influenced to explore the hotspot using location traces. They offer the use of the concept of graph theory. A stretch factor created for a taxicab to identify travel and it infers that the algorithm based on HMM can identify the journey with accuracy and remembrance of 90%–94%. A statistical refinement is performed over the previous outlook that results in remembrance and accuracy of about 50%–60%.

When the taxicab is engaged by a traveller and the corresponding pickup and drop-off points being able to recognise the endpoint of the trip, the trip can have a simple solution to recognise that taxicab drivers can mark when they pickup and drop-off to the passenger. The main object of this paper is to recognise trips of taxicabs and many existent taxicabs after evaluating (Zhu et al., 2012; Cheng et al., 2012). These trips aggregated across a group of taxicabs determine human mobility patterns. The author proposes some easy methods like spatial clustering-based classification, speed-based classification that is impressive to find the pickup and drop-off points, but not so impressive at recognising the trips. s is a spot of the taxicab, that easy method is influential in deciding the possibility of s being a pickup or drop-off point. Although, they memorise so a trip is a couple of pickup and drop-off points; therefore, they are interested in detecting the possibility of s_p being a pickup point and s_d being corresponding the drop-off point. An important insight is as follows. They assume three spots home, coffee shop and office and suppose that coffee shop arrives on the way from home to office.

A classifier can recognise that home is an outsider and the last point of probable travel and coffee shop and office are most visited places and thus both are probably the last point of the tour. But a spot find which includes home, coffee shop, and office in that order academic views did not decide which is the tour's drop-off point is the coffee shop or office. Therefore, to solve this issue, they introduce the graph-theoretical concept, stretch factor (Cheng et al., 2012), as a solution to the problem discussed here; they present the concept of graph theory, stretch factor. The distance stretch factor is a ratio of the original distance for the given route to the shortest distance possible from starting to the endpoint of the route. Similarly, they define time stretch factor which is a ratio of actual time to the shortest possible time (from start point to endpoint of the trip).

Definition: Let $m(*, *)$ be a metric that satisfies the triangular inequality: $m(x, y) + m(y, z) \geq m(x, z)$, $\forall\, x, y$ and z. Then, the stretch factor of a cab tour from the spot s_0 to s_n through $s_1, s_2, \ldots\ldots\ldots, s_{n-1}$ is given by:

$$\sum_{i=0}^{n-1} \frac{m(s_i, s_{i+1})}{m(s_0, s_n)}$$

Here, $m(x, y)$ are two metrics in which one is based on the geographical shortest distance between points x and y (distance stretch factor) and the second based on the shortest time of tour between points x and y (time stretch factor). The main insight is whenever taxicab is engaged drivers choose to travel the smallest way either based on distance or time and if taxicab is not engaged, drivers choose to travel that way which is related to the shortest way. But that is not enforced to find of own tour where the mobility pattern is much dissimilar from the taxicab (Stenneth et al., 2011). They use that theory to the division of a taxicab detecting into various sub-divisions related to the shortest path routes (engaged state) and non-shortest path routes (idle state) (Zhu et al., 2012). Especially, the combination of stretch factor (in a novel manner) concept with traditional machine learning techniques can improve the detection rate of occupancy. They calculate this idea using two comprehensive datasets, one from a deployment in Stockholm, Sweden and another from Shanghai, China to show that in most cases the stretch factor can be used to identify last points of a tour with 90%–94% precision. In this paper, they evolved various algorithms to detect human mobility patterns from taxicab location traces. They explain that taxicabs are a good podium to know about human mobility patterns. This issue is similar to knowing about the tour which is made by the taxicab.

The evaluation present that simple clustering techniques can exactly analyse if a specific spot is the last point, but in order to analyse last points compatible with a specific tour, it is not sufficient. They evolved algorithms to trace this problem that uses stretch factor in a novel way. The HMM depends on an algorithm, which beat the hierarchical segmentation algorithm, obtains 90%–94% accuracy and memorise to find the tours for both the Stockholm and Shanghai datasets. HMM-based algorithm HSMM exceeds past work. They used the outcome of the HMM-based algorithms to calculate the top ten repeatedly travelled routes in both the cities (Ganti et al., 2013).

4.6 User mobility analysis in LTE network

Areas like network planning, resource preservation, mobile computing, etc. have extensive usage of human mobility models. In long-term evolution (LTE) network, the speedy growth of the number of mobile broadband customers' needs the opportunity of an increase in data services and mobility management. Lv et al. (2014) target on service evolved node B (eNodeB) (Zhang et al., 2012) forecasting in LTE cellular network which depends on human mobility, and introduces academic and actual method advantage HMM. The use of HMM permits them to deal with trajectory specialty of eNodeB, which is in the form of unobservable parameters and as well the impact of the person's historical service eNodeB. They test with actual data that comes from real-world cellular networks, and they examine the effects of distinct parameters on the presentation of the forecasting by using of control variant method. The outcome finds that the accuracy of forecasting is 53%. These conclusions are very important for location forecasting issues. If the actual communication system is accepted then the model shows more competence. Researchers have taken into account the speed of mobility patterns and the movement of individual mobility to improve the position of communication and to provide better service delivery. For better results in communication and supply service, researchers have looked at the forecasting of mobility patterns and individual mobility trajectory. People commonly have an exact purpose or habitual path, human mobility is away from random and affected by its historical tendency. Hence, mobility trajectory which depends upon location forecasts is promising to achieve higher precision for assured source conservation and better service results. This paper exhorts the evolution and assessment of models which endorse the forecast for a next service eNodeB. A pattern in the given period for targeted peoples holds a sequential relation between accessed eNodeB. In this paper for forecasting, they use HMM (Huang et al., 2013; Elnashar and El-Saidny, 2013; Huang et al., 2012). Overall, they introduced an adept system model to forecast the next service eNodeB on the base of HMM in the LTE cellular network. If they use HMM, then it permits them to contemplate the user mobility pattern with the dimension of service eNodeB. They determined the parameters of the forecast and used the control variant method to estimate the impact of every different parameter's forecasting presentation. After analysing around 2,800 users for more than a week belonging to Southern China province, final prediction accuracy is up to 53%. The paper is educative in progressive research and exercises, in which these methods are used.

4.7 Machine learning for sequential data

Sequential data plays an important role in analysing the problems in many areas of statistical learning. Dietterich (2002) shaped the principle learning tasks and interpreted the method that developed under the machine-learning research community which addresses the subject of the sequential data analysis (Mitchell, 1997). There are several methods like sliding window protocol, conditional random field (CRF), graph transform network, recurrent window protocol, HMMs, etc. The very basic concept which they learn about the supervised machine learning is to build a classifier which can correctly identify the object being tested with the help of model build upon the old data objects. A basic use case of this can be identifying the handwritten alphabet images, which describe a letter among the 26 alphabets of the English language. The problem can be formalised as follows:

u handwritten images of the English alphabet

v respective English alphabet, where v belongs to set = $\{A, B, \ldots, Z\}$.

They define a training example then it will be an ordered set of (u, v). If they define a classifier over a set of training examples, then it will correctly map the image *RROAW* to the English alphabet *ARROW*. Let the classifier is a function – *f*, they defined as *ARROW* = *f*(*RROAW*). This task can be done after searching over some set of the classifier which they have defined and choosing the one which is providing the best results possible. There are several other problems that can be solved using the basic idea of supervised learning, i.e., fraud call detection, spam detection, classification of fruits, etc. A very good example is part of speech tagging, where the classifier will identify the word in a given sentence and will tag them as a part of speech like verb, noun, pronoun, adjective, etc. But some problems they solve using supervised learning to exhibit a correlation between the inputs and outputs. For example, when they detect a phone call is a fraud call or not, then this task can be done by finding a change in the pattern of usage of the phone call. They can say if the phone is stolen, then calls which they going to be received will be fraud calls always. Same way, when they talk about the part of speech tagging, a correct statement of the English language will follow the grammatical rules. So they can say that tagging output as (verb verb pronoun) will not occur. Thus, they define some sequential pattern for the data which is being tested using the classifier. This will help to improve the prediction efficiency of our classifier as some relation will be always true according to the standards which they have identified in the sequential data pattern.

As an example, to detect a fraud call, they analyse the usage pattern of the phone so if the new connection differs from the regular distribution then there are higher chances that it will be a fraud call. They formalise the sequential supervised learning as follows.

A training example = $<u_i, v_i>$ where $i = [1, N]$ and $u_i = <u_{i,1}, u_{i,2}, u_{i,3}, \ldots\ldots\ldots, u_{i,t}>$. In normal supervised learning, u_i will be single term but in the sequential supervised learning $\{(u_i, v_i)\}$ is pair. They define u_i as a sequence of terms which have a strong correlation among them. So if one takes an example of part of speech tagging then u_i = (do you want apple) and v_i = (verb pronoun verb noun). So a classifier will do a job of identifying a sequence of terms to a sequence of a prediction, made for each term using a correlation and previous knowledge. A new classifier defined for sequential supervised learning is $h_{sequential}$.

$$v_i = h_{sequential}(u_i)$$

Sequential supervised learning has a contrast relation with two prediction problems. The first one is the time series analysis and the second is sequence classification. Suppose z is the output parameter such that $z = <z_1, z_2, \ldots\ldots\ldots, z_t>$. In time series analysis, it will give them the true value analysed up to time t and they have to predict the value of z_{t+1}. There is some difference in sequential supervised learning and time series analysis. In the first algorithm, given an entire sequence possible to predict the value of z but, in the second one, a prefix of the input sequence is given. In the sequence classification problem, they have to predict a single label let say 'RUPEES', which can be applied to the entire sequence of input. For example, after identifying the handwritten character images, they have to tell who has written those characters. So the problem is the identification of handwriting and prediction of the name of individuals. All of the problems discussed here can be used to solve in some other problems also. For example, if they can identify individually handwritten alphabets, then they can correctly predict a word formed by the letters. After checking the different application of sequential supervised learning here, they can say before predicting any value, if they analyse the given dataset first, then they can improve the prediction efficiency. This paper has analysed sequential supervised learning tasks and discussed the different algorithms (Dietterich, 2002).

4.8 CRFs for activity recognition

In order to create a brilliant multi-agent environment, activity identification plays a very basic role. The activity identification problem is similar to temporal classification. Here, Vail et al. (2007) analysed and compared the performance of two temporal classification models: HMM and CRFs. CRFs are a kind of specialised models for indexing a sequence. When they apply conditions on data, they are eliminating the steps which will be required to put some independent specifications between the observations. Conditioning is useful as it covers the entire collection of features which provides a facility of not to compromise with the assumptions made for the dataset. They are using data from a simulated robot tag environment as it is from a multi-agent scenario and it will provide better random relations among the observations. By taking this kind of dataset, they are analysing the performance of both CRF (Lafferty et al., 2001) and HMM (Rabiner, 1989). Also, they will be able to see the effect of those features which are likely to go wrong with the assumptions made between observations; these features will be helpful to get high classification efficiency. They find that the discriminative CRFs are either performing better than HMM or at least comparable to these, even when any features not going wrong with the prior assumptions. They have also seen that if features are dependent on the time of the observations that have been taken then, CRFs are more promising for the performance which they want. HMMs have been a traditional way for activity recognition, whereas, CRFs are the new players in this field.

The original area of CRF is text processing but now they are applied to other applications such as image processing, gesture analysis, motion tracking and activity recognition. The main focus of this paper is to analyse CRF in the field of activity recognition. In order to achieve a better classification performance using CRF, domain knowledge is very much required which can help them to build a refined set of features. These features often include information from multiple periods. Such features that have information from multiple periods tend to violate the concept of independent assumptions in case of HMM, but not in CRF.

Here, they analyse both the models and how such features are affecting them. In addition to this, they also examine those features which join the state changing in the model directly to the observations. These kind of features are very much tedious to represent in an HMM as the probabilistic factorisation provided by those. They analyse whether the representation of the features has some kind effect on the classification performance. Aside from the features which are tedious to fit in HMM or going wrong with some prior independent assumptions, they also try to find the performance measure of discriminative and generative models for classification. Especially, they are going to consider the case where the HMM and CRF have been treated with same features. After examining the discriminative generative model with the help of logistic regression and naive Bayes classifiers, results are giving an indication that CRF should give better classification precision than HMM even if they use the same set of features for both. This should happen because CRFs and HMMs can be considered a pair of discriminative generative model. After this, they present a method for the conversion of an HMM to CRF. It is also seen that the discriminative models give a lesser value of an asymptotic error compared to the generative model.

In continuation, they represent the simulated robot tag domain which has been used. By considering the activity recognition as a general problem, they defined the design and dimensions of their domain. Also, they give an introduction of the CRF in the context of the activity recognition problem. They analyse CRF and HMM in the view of a simple activity recognition problem. They can show that CRFs are better with a scalable range of computed features where domain knowledge is also taken into consideration. Also, they proved that as the independent assumptions are inherent in HMMs, such features are helping in improving the performance of the model. All the tests which have been included are depicting that CRFs have a better score than HMMs. This performance difference has been observed in such conditions where the CRF model was a discriminatively trained HMM. The whole analysis will make a sense when they compare the performance of generative and discriminative models, for example when they compare the naive Bayes and logistic regression or CRFs and HMMs.

They also presented a solution for incorporating such features in CRFs which are difficult to fit in HMMs, e.g., the distance threshold features, which completely denotes the dynamics of a process. A refined calculation is needed for the features which have data collection from multiple time spans in order to attain a higher performance for classification in activity recognition and for such features, CRFs are more suitable. CRFs do not use a strategy of an independent assumption made between the observations, they are able to connect a particular observation for state change, as discriminatively trained models, and they offer lower value for error when they compare them with generative models (Vail et al., 2007).

4.9 Human emergency behaviour and their mobility following large-scale disaster

Song et al. (2014) build a vast human mobility database and developed a human behavioural model that depends on human emergency behaviour and their mobility at the time of large-scale devastation (Cho et al., 2011). This model helps in forecasting human behaviour and their mobility for the effective planning of management of the disaster. On the basis of the empirical analysis through database, they detect that human behaviour and their mobility backing large-scale devastation correlates with their mobility pattern occasionally and also closely connected with their social relationship, damage level, the intensity of disaster, facilities provided by government, etc. These factors are considered in the model of human behaviour for exactly forecasting human emergency behaviour and their mobility backing large-scale devastation. Human mobility and its behaviour patterns keep high levels of independence and diversity, but this mobility is important for effective humanitarian relief, long-term social reconstruction and prediction of disaster management. They also perform structural patterns because of geographical and social limitations.

According to Lu et al. (2012) after the 2010 Haitian earthquake disaster, the population dynamics pattern was much correlated with its daily mobility prior to the event and then concluded that after wide-ranging disasters, population mobility can be undoubtedly estimated more than the assumed. As per Song et al. (2013), in the immediate aftermath of the earthquake and Fukushima nuclear accident in eastern Japan, people demanded asylum and government shelter in a large number of nearby cities. These areas were very affected by the emergence and removal of atomic materials. After the disasters on large-scale, there are some questions and hypotheses about human behaviour and mobility, whose answer is much unknown and whose biggest factor is the lack of supporting data and powerful human behaviour data, which is fully capable of describing the effect on the population dynamics pattern.

In this paper, they created a large population dynamic database that collects and manages a daily GPS record of nearly 1.6 million people in Japan in a year. In addition, Japan created many different datasets for the person's mobility, analysis, and acquisition of behaviour after the earthquake and nuclear accident. Based on the experimental analysis of population dynamics patterns through these datasets, they find that population behaviour and its mobility are sometimes correlated with their dynamism pattern during the normal time after this unprecedented overall disaster and is strongly affected by its social relations, disaster intensity, damage level, etc.

The basis of these results, they interconnected between human emergency behaviour and influencing factors, and after a large-scale disaster, a model is developed regarding human behaviour to forecast population mobility. They divide the prediction problem into two sub-assemblies:

- The current state of disaster, the other impact factor and the human mobility observed, predict possible behaviour on the next step.
- Given the estimated distribution of behaviours, then the prediction of its potential mobility.

They developed a human behaviour model that accurately monitors different disaster factors predicting their behaviour and mobility. Experimental outcome and verification behaviours show the efficiency of the model and much better performance than the previous one. They considered some boundaries in this study. The population dynamics database used is created with the help of mobile devices and they do not include that data where peoples do not use mobile devices and GPS services in some population areas. Data provided here may have a little slope for the young population as this age group is a bigger user of GPS-based devices compared to the elder people. Anyway, they are pretty much sure the data, which includes a pattern of movement for nearly 1.6 million people, can show a generic movement scenario in case of composite disaster. Another case where this predictive model may not fit is the extrapolation of the patterns identified, as it can be useful for the location outside Japan and the locations where such disaster has not come (Song et al., 2014).

5 Conclusions

This literature review discusses the overall approach to design HMM-based algorithm to analyse human mobility pattern. The paper starts with an overview of HMM where various components and their significance is discussed with some examples. Paper also discusses various scenarios where analysis of human mobility is important. Various real-world examples have been discussed like LTE network data, photographic data of user movement, taxicab locations, etc. which can be used to model the observable and hidden states of HMM. It can be easily concluded that human mobility data varies with time and space and the future location of a human being always depends on the past movement. This conclusion is a good remark to model HMM to analyse the problem of Human mobility analysis.

6 Future remarks

An interesting research frontier in human mobility in the near future will likely concern the shift from traditional vehicles to autonomous, self-driving, vehicles. The dispersal of autonomous vehicles will transform individual and local transportation, with extensive outcomes for society, the economy, and the environment. How accurately will all these modifications happen and how will mobility habits change in reaction to these new technologies? How can such vehicles be controlled and how does one plan routes in order to optimise fuel and time consumption, reducing congestion and pollution? Answers to these important questions may well arise from a fuller understanding of human mobility and behaviour at the personal and population (collective) levels. Interdisciplinary and cross-fertilisation is a cause-in-fact to the success of the field. Computer scientists, physicists, social scientists, environmentalists, engineers, government officials, and many other professions need to work together to develop and implement solutions to lots of the issues society faces today that could benefit from understanding mobility (and dynamics) such as crime, urban planning, energy consumption, social integration, to name but a few.

References

Aalto, L., Göthlin, N., Korhonen, J. and Ojala, T. (2004) 'Bluetooth and WAP push based location-aware mobile advertising system', in *Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services*, June, pp.49–58.

Asahara, A., Maruyama, K., Sato, A. and Seto, K. (2011) 'Pedestrian-movement prediction based on mixed Markov-chain model', in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, November, pp.25–33.

Ashbrook, D. and Starner, T. (2002) 'Learning significant locations and predicting user movement with GPS', in *Proceedings. Sixth International Symposium on Wearable Computers*, IEEE, October, pp.101–108.

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C.R., Lenormand, M., Louail, T. and Tomasini, M. (2018) 'Human mobility: models and applications', *Physics Reports*, Vol. 734, pp.1–74,

Barchiesi, D., Preis, T., Bishop, S. and Moat, H.S. (2015) 'Modelling human mobility patterns using photographic data shared online', *Royal Society Open Science*, Vol. 2, No. 8, p.150046.

Berlingerio, M., Calabrese, F., Di Lorenzo, G., Nair, R., Pinelli, F. and Sbodio, M.L. (2013) 'AllAboard: a system for exploring urban mobility and optimizing public transport using cellphone data', in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, Berlin, Heidelberg, September, pp.663–666.

Botta, F., Moat, H.S. and Preis, T. (2015) 'Quantifying crowd size with mobile phone and Twitter data', *Royal Society Open Science*, Vol. 2, No. 5, p.150162.

Bright, M.L. and Thomas, D.S. (1941) 'Interstate migration and intervening opportunities', *American Sociological Review*, Vol. 6, No. 6, pp.773–783.

Chaturapruek, S., Breslau, J., Yazdi, D., Kolokolnikov, T. and McCalla, S.G. (2013) 'Crime modeling with Lévy flights', *SIAM Journal on Applied Mathematics*, Vol. 73, No. 4, pp.1703–1720.

Cheng, S.W., Knauer, C., Langerman, S. and Smid, M.H. (2012) 'Approximating the average stretch factor of geometric graphs', *JoCG*, Vol. 3, No. 1, pp.132–153.

Cho, E., Myers, S.A. and Leskovec, J. (2011) 'Friendship and mobility: user movement in location-based social networks', in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August, pp.1082–1090.

Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.J. and Vespignani, A. (2007) 'Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions', *PLoS Medicine*, Vol. 4, No. 1, p.e13.

Cornford, F.M. (1936) *Studies in Philosophy*, by GC Field MA, D. Litt. Professor of Philosophy in the University of Bristol, Bristol: JW Arrowsmith, Ltd. 1935, p.250, Price 10s. 6d., Philosophy 11, No. 42, pp.249–249.

De Domenico, M., Lima, A. and Musolesi, M. (2013) 'Interdependence and predictability of human mobility and social interactions', *Pervasive and Mobile Computing*, Vol. 9, No. 6, pp.798–807.

Dietterich, T.G. (2002) 'Machine learning for sequential data: a review', in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, Springer, Berlin, Heidelberg, August, pp.15–30.

Elnashar, A. and El-Saidny, M.A. (2013) 'Looking at LTE in practice: a performance analysis of the LTE system based on field test results', *IEEE Vehicular Technology Magazine*, Vol. 8, No. 3, pp.81–92.

Farrahi, K. and Gatica-Perez, D. (2011) 'Discovering routines from large-scale human locations using probabilistic topic models', *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 2, No. 1, pp.1–27.

Gambs, S., Killijian, M.O. and del Prado Cortez, M.N. (2012) 'Next place prediction using mobility Markov chains', in *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, April, pp.1–6.

Ganti, R., Srivatsa, M., Ranganathan, A. and Han, J. (2013) 'Inferring human mobility patterns from taxicab location traces', in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, September, pp.459–468.

Gonçalves, B., Balcan, D. and Vespignani, A. (2013) 'Human mobility and the worldwide impact of intentional localized highly pathogenic virus release', *Scientific Reports*, Vol. 3, No. 1, p.810.

Gonzalez, M.C., Hidalgo, C.A. and Barabasi, A.L. (2008) 'Understanding individual human mobility patterns', *Nature*, Vol. 453, No. 7196, pp.779–782.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P. and Ratti, C. (2014) 'Geo-located Twitter as proxy for global mobility patterns', *Cartography and Geographic Information Science*, Vol. 41, No. 3, pp.260–271.

Huang, J., Qian, F., Gerber, A., Mao, Z.M., Sen, S. and Spatscheck, O. (2012) 'A close examination of performance and power characteristics of 4G LTE networks', in *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, June, pp.225–238.

Huang, J., Qian, F., Guo, Y., Zhou, Y., Xu, Q., Mao, Z.M., Spatscheck, O. et al. (2013) 'An in-depth study of LTE: effect of network protocol and application behavior on performance', *ACM SIGCOMM Computer Communication Review*, Vol. 43, No. 4, pp.363–374.

Hufnagel, L. Brockmann, D. and Geisel, T. (2004) 'Forecast and control of epidemics in a globalized world', *Proceedings of the National Academy of Sciences*, USA, Vol. 101, pp.15124–15129.

Hufnagel, L., Brockmann, D. and Geisel, T. (2006) 'The scaling laws of human travel', *Nature*, Vol. 439, No. 7075, pp.462–465.

Jeung, H., Shen, H.T. and Zhou, X. (2007) 'Mining trajectory patterns using hidden Markov models', in *International Conference on Data Warehousing and Knowledge Discovery*, Springer, Berlin, Heidelberg, September, pp.470–480.

Lafferty, J., McCallum, A. and Pereira, F.C. (2001) 'Conditional random fields: probabilistic models for segmenting and labeling sequence data', *Proceedings of the Eighteenth International Conference on Machine Learning*, pp.282–289.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.L., Brewer, D. and Jebara, T. (2009) 'Computational social science', *Science*, Vol. 323, No. 5915, pp.721–723.

Lu, E.H.C. and Tseng, V.S. (2009) 'Mining cluster-based mobile sequential patterns in location-based service environments', in *2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*, IEEE, May, pp.273–278.

Lu, X., Bengtsson, L. and Holme, P. (2012) 'Predictability of population displacement after the 2010 Haiti earthquake', *Proceedings of the National Academy of Sciences*, Vol. 109, No. 29, pp.11576–11581.

Lv, Q., Mei, Z., Qiao, Y., Zhong, Y. and Lei, Z. (2014) 'Hidden Markov model based user mobility analysis in LTE network', in *2014 International Symposium on Wireless Personal Multimedia Communications (WPMC)*, IEEE, September, pp.379–384.

Mathew, W., Raposo, R. and Martins, B. (2012) 'Predicting future locations with hidden Markov models', in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, September, pp.911–918.

Mitchell, T.M. (1997) 'Does machine learning really work?', *AI Magazine*, Vol. 18, No. 3, pp.11–11.

Musolesi, M. and Mascolo, C. (2007) 'Designing mobility models based on social network theory', *ACM SIGMOBILE Mobile Computing and Communications Review*, Vol. 11, No. 3, pp.59–70.

Olsson, G. (1965) *Distance and Human Interaction: A Review and Bibliography*, Regional Science Research Institute, Philadelphia.

Rabiner, L.R. (1989) 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceedings of the IEEE*, Vol. 77, No. 2, pp.257–286.

Raichlen, D.A., Wood, B.M., Gordon, A.D., Mabulla, A.Z., Marlowe, F.W. and Pontzer, H. (2014) 'Evidence of Lévy walk foraging patterns in human hunter-gatherers', *Proceedings of the National Academy of Sciences*, Vol. 111, No. 2, pp.728–733.

Ravenstein, E.G. (1885) 'The laws of migration', *Journal of the Statistical Society of London*, Vol. 48, No. 2, pp.167–235.

Ravenstein, E.G. (1889) 'The laws of migration', *Journal of the Royal Statistical Society*, Vol. 52, No. 2, pp.241–305.

Rozenfeld, H.D., Rybski, D., Andrade, J.S., Batty, M., Stanley, H.E. and Makse, H.A. (2008) 'Laws of population growth', *Proceedings of the National Academy of Sciences*, Vol. 105, No. 48, pp.18702–18707.

Ryu, S.H. and Song, H.Y. (2013) 'Human mobility model establishment with hidden Markov model', in *WSEAS International Conference. Proceedings. Recent Advances in Computer Engineering Series*, No. 9, WSEAS.

Song, C., Qu, Z., Blumm, N. and Barabási, A.L. (2010) 'Limits of predictability in human mobility', *Science*, Vol. 327, No. 5968, pp.1018–1021.

Song, X., Zhang, Q., Sekimoto, Y. and Shibasaki, R. (2014) '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*, August, pp.5–14.

Song, X., Zhang, Q., Sekimoto, Y., Horanont, T., Ueyama, S. and Shibasaki, R. (2013) 'Modeling and probabilistic reasoning of population evacuation during large-scale disaster', in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August, pp.1231–1239.

Stenneth, L., Wolfson, O., Yu, P.S. and Xu, B. (2011) 'Transportation mode detection using mobile phones and GIS information', in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, November, pp.54–63.

Stouffer, S.A. (1940) 'Intervening opportunities: a theory relating mobility and distance', *American Sociological Review*, Vol. 5, No. 6, pp.845–867.

Vail, D.L., Veloso, M.M. and Lafferty, J.D. (2007) 'Conditional random fields for activity recognition', in *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, ACM, May, p.235.

Wang, D., Pedreschi, D., Song, C., Giannotti, F. and Barabasi, A.L. (2011) 'Human mobility, social ties, and link prediction', in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August, pp.1100–1108.

Yang, J., Xu, Y. and Chen, C.S. (1997) 'Human action learning via hidden Markov model', *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, Vol. 27, No. 1, pp.34–44.

Ying, J.J.C., Lee, W.C., Weng, T.C. and Tseng, V.S. (2011) 'Semantic trajectory mining for location prediction', in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, November, pp.34–43.

Zhang, L., Okamawari, T. and Fujii, T. (2012) 'Performance evaluation of TCP and UDP during LTE handover', in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, April, pp.1993–1997.

Zhu, Y., Zheng, Y., Zhang, L., Santani, D., Xie, X. and Yang, Q. (2012) *Inferring Taxi Status Using GPS Trajectories*, arXiv preprint arXiv: 1205.4378.

Zion, E.B. and Lerner, B. (2018) 'Identifying and predicting social lifestyles in people's trajectories by neural networks', *EPJ Data Science*, Vol. 7, No. 1, p.45.

Zipf, G.K. (1940) 'The generalized harmonic series as a fundamental principle of social organization', *The Psychological Record*, Vol. 4, p.41.

Zipf, G.K. (1946) 'The P 1 P 2/D hypothesis: on the intercity movement of persons', *American Sociological Review*, Vol. 11, No. 6, pp.677–686.