University of Kasdi Merbah Ouargla

Faculty of New Information and Communication Technologies

Department of Computer Science and Information Technology

Dissertation presented for obtaining the Master's degree

Specialty: Network Administration and Security

Presented by: Farsi Amna & Meissa Ibtihadj.

A Unified Graph-Based framework for Mobility Prediction and Recommendation

Supported in**: 18 /06/**2022

The jury:

Dr.Amirat Hanane	Supervisor	UKMOuargla
Dr.Bouanane k	President	UKMOuargla
Dr.Bachir S	Examiner	UKMOuargla

Academic Year:2021–2022

Acknowledgments

*Praise be to **God** first always and forever. I dedicate my graduation to the light that has lit my way and the lamp that never goes out, who has made efforts over the years to climb the ladders of success my dear father, you have given us a fortune.*

*And to the symbol of love and the healing balm of the white heart, **my mother**, the comfort of my eyes to all **my sisters** and members of **my family** and their loved ones to everyone who shared my success.*

*To Professor **Dr: Hanan Amirat** for all the guidance and valuable information she gave us, which contributed to praising the subject of our study in all its aspects, to all of my teachers and teachers at all levels. To all those whom fates willed to bring me together with the gardens of study and make them brothers. The search locomotive passed many obstacles, however, I tried to steadily overcome them, thanks to God and from Him. I dedicate this research to you, and I am grateful to all those who had merit in my career and helped me, even if only a little.*

Amna Farsi & Meissa Ibtihadj.

Abstract

In ad hoc vehicle networks (VANETs), predicting human mobility is very important for a wide range of location-based applications. Vehicle nodes constantly move and exchange data that contains movement context such as vehicle location, speed, etc. The amount of available data is directly proportional to the accuracy of predicting future behaviour of mobile nodes and their movements. In literature, machine learning (ML) based prediction models such as Markov-based and data mining based models have been leveraged and have proven their efficiency through intelligent and adaptive data-driven methods to boost the prediction accuracy. Many of the existing models are Markov-based or data mining based prediction models. These models perform short-term forecasting to predict the next few potential locations that the user would visit (i.e. location prediction) or may be interested in visiting in the near future (location recommendation). Yet, most of the existing models often investigated location prediction and location recommendation in an isolated manner. Therefore, we have designed a unified framework, called *MyLoc,* for route in VANET and location recommendation for Location-Based Social Networks (LBSN). Our proposed framework adopts a graph-based probabilistic representation to predict mobility in both prediction ranges (short and long term). *MyLoc* exploits vehicle communications to collect mobility data and considers the sequential recommendation influence factor. In our work, two experiments conducted to compare our model with the well-known pre-Markov model have shown the superiority of our proposal in terms of prediction accuracy and coverage.

Key words: Mobility prediction, Point of interest, location recommendation, location prediction, Location-based social networks, VANET.

Résumé

Dans les réseaux de véhicules ad hoc (VANET), la prédiction de la mobilité humaine est très importante pour un large éventail d'applications basées sur la localisation. Les nœuds de véhicules se déplacent et échangent constamment des données qui contiennent un contexte de mouvement tel que l'emplacement du véhicule, la vitesse, etc. La quantité de données disponibles est directement proportionnelle à la précision de la prédiction du comportement futur des nœuds mobiles et de leurs mouvements. Dans la littérature, les modèles de prédiction basés sur l'apprentissage automatique (ML) tels que les modèles basés sur Markov et sur l'exploration de données ont été exploités et ont prouvé leur efficacité grâce à des méthodes intelligentes et adaptatives basées sur les données pour améliorer la précision de la prédiction. De nombreux modèles existants sont des modèles de prédiction basés sur Markov ou sur l'exploration de données. Ces modèles effectuent des prévisions à court terme pour prédire les prochains emplacements potentiels que l'utilisateur visiterait (c'est-à-dire la prédiction d'emplacement) ou pourrait être intéressé à visiter dans un avenir proche (recommandation d'emplacement). Pourtant, la plupart des modèles existants ont souvent étudié la prédiction et la recommandation de localisation de manière isolée. Par conséquent, nous avons conçu un cadre unifié, appelé MyLoc, pour l'itinéraire dans VANET et la recommandation de localisation pour les réseaux sociaux basés sur la localisation (LBSN). Notre cadre proposé adopte une représentation probabiliste basée sur des graphes pour prédire la mobilité dans les deux plages de prédiction (court et long terme). MyLoc exploite les communications du véhicule pour collecter des données de mobilité et considère le facteur d'influence de la recommandation séquentielle. Dans notre travail, deux expériences menées pour comparer notre modèle au modèle de Markov bien connu ont montré la supériorité de notre proposition en termes de précision de prédiction et de couverture.

Mots clés : Prédiction de mobilité, Point d'intérêt, recommandation de localisation, prédiction de localisation, Réseaux sociaux géolocalisés, VANET.

ملخص

في شبكات المركبات المخصصة (VANETs) ، يعد التنبؤ بالتنقل البشري أمرًا مهمًا للغاية لمجموعة واسعة من التطبيقات القائمة على الموقع. تتحرك عُقد السيارة باستمرار وتتبادل البيانات التي تحتوي على سياق الحركة مثل موقع السيارة والسرعة وما إلى ذلك. يتناسب مقدار البيانات المتاحة بشكل مباشر مع دقة التنبؤ بالسلوك المستقبلي للعقد المتنقلة وتحركاتها. في الأدبيات ، تم تعزيز نماذج التنبؤ القائمة على التعلم الآلي (ML) مثل النماذج المستندة إلى Markov والنماذج القائمة على التنقيب في البيانات حيث أثبتت كفاءتها من خلال طرق ذكية وقائمة على البيانات للتكيف وتعزيز دقة التنبؤ. العديد من النماذج الحالية هي نماذج تنبؤ قائمة على أساس ماركوف أو تنقيب في البيانات. تؤدي هذه النماذج تنبؤًا قصير المدى للتنبؤ بالمواقع القليلة المحتملة التالية التي سيزورها المستخدم (أي توقع الموقع) أو قد يكون مهتمًا بزيارتها في المستقبل القريب (توصية الموقع). ومع ذلك ، فإن معظم النماذج الحالية غالبًا ما تحقق في التنبؤ بالموقع وتوصية الموقع بطريقة منعزلة. لذلك ، قمنا بتصميم إطار عمل موحد يسمى MyLoc ، للتوجيه في VANET وتوصية الموقع للشبكات الاجتماعية القائمة على الموقع .(LBSN) يتبنى إطار العمل المقترح تمثيلًا احتماليًا قائمًا على الرسم البياني للتنبؤ بالتنقل في كلا نطاقي التنبؤ (على المدى القصير والطويل). يستغل MyLoc اتصالات السيارة لجمع بيانات التنقل ويأخذ في الاعتبار عامل تأثير التوصية المتسلسل. في عملنا ، أظهرت تجربتان أجريتا لمقارنة نموذجنا مع نموذج ماركوف المعروف تفوق اقتراحنا من حيث دقة التنبؤ والتغطية.

الكلمات المفتاحية: التنبؤ بالتنقل ، نقطة الاهتمام ، التوصية بالموقع ، التنبؤ بالموقع ، الشبكات الاجتماعية القائمة على الموقع ، .VANET

Table of Contents

Chapiter2 : *The MyLoc* model for route prediction and location.

Chapiter 3 :Experimental evalution.

List of figures

List of tables

Notations and Symbols

POI :Point Of Interest.
LBS: Location Based Service.
LBSN: Location Based Social Network.

General Introduction

Context and motivation

Vehicular Ad hoc NETwork (VANET) is an emergent promising technology that applies various Information and Communication Technologies (ICTs) to improve the safety and quality of service in Intelligent Transportation Systems (ITS) through communication among vehicles and the environment. Equipped with computation, communication, and storage resources, vehicular nodes in VANET are able to share information with other vehicles in a cooperative manner. The shared information includes, but not limited to, geographic location, speed, control, and direction through the designated sensors. In addition to the situational awareness information, important security and warning messages are also shared by either vehicular nodes or by the infrastructure with the neighbors. It is however, worth mentioning that location information is of paramount importance in almost all kinds of messages and can help taking important and timely decisions by the individual nodes. Additionally, location information also aids in predicting the future movement (mobility prediction) that will help optimizing the location-based services and applications. Furthermore, mobility prediction is also important for the design of efficient routing protocols and data dissemination in VANET. Such mobility prediction is essential for different domains of VANET applications such as, but not limited to, car sharing, automatic parking, traffic management and so on. Besides, the knowledge of vehicles future movement a priori could also optimize the fuel consumption.

Contribution

To overcome the limitations of these models, we propose, in this dissertation, a location prediction model called *MyLoc* that is based on the DG prediction model (Dependency Graph). *MyLoc* has the advantages of DG as it depends on determining the next route and destination with the help of the most important recommendation (points of interest) in LBSN.

The model *MyLoc* forecasts the driver route and location reyling to regard the sequential nature of human mobility and could be easily extended the influence factor on user:

The contributions of this dissertation can be summarized as follows:

- We introduce *MyLoc* model as a location and route prediction model. With the help of POI recommendation.
- Run two experiments to confirm the effectiveness of *MyLoc* using a real dataset (Gowalla) and compare its performance with the site prediction model and existing methods..

dissertation organization

This dissertation is organized as follows:

- In the first chapter, we will present 1) a set of basic concepts related to movement prediction , and 2) the existing mobility prediction works. These works were categorized by the forecasting technique used and subsequently compared according to certain criteria such as prediction accuracy and the influence factor considered in POI recommendation.
- In Chapter 2, we will provide an overall description of our *MyLoc* model.
- In the third chapter, we present an empirical study of the *MyLoc* performance and its comparison with Markov-Pre (First order Markov) based model in terms of prediction accuracy and coverage.
- Finally, we end the dissertation with a conclusion.

Chapiter 1:

Preliminary notions

1.1 VANETs (Vehicular Ad Hoc Networks)

VANETs (Vehicular Ad Hoc Networks) is a new emerging technology from mobile ad hoc networks (MANETs) as shown in Figure 1, where the nodes mobiles are intelligent vehicles, equipped with very high-tech equipment (Computers, radars, geolocation systems (GPS), different types of sensors
and network devices). VANETs networks allow inter-vehicular communications (V2V) and infrastructure vehicles (V2I). The different nodes can exchange any alerts or useful information to improve road traffic safety. But also data (music, video, advertisements, etc.) to make the time spent on the road more enjoyable and less boring.
VANET networks are based on two types of applications. The apps that constitute the core of an intelligent transport system ITS (Intelligent Transport System), to ensure the improvement of road safety, but also of applications deployed for passenger comfort.

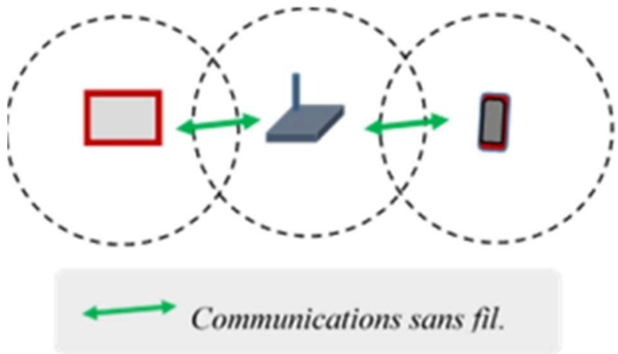

Figure 1 : Mode of communications in mobile ad hoc networks (MANETs).

1.1.1. The VANET network system

➢ Architecture and components :

A VANETs network consists mainly of three entities :

a. TA (Trusted Authority)

Say CA in French (trusted authority). It is a source of authenticity of information. It ensures the management and registration of all entities on the network (RSU and OBU). The TA is supposed to know all the true identities of the vehicles and if necessary disclose them to the police. Also, the TA in some works is responsible for issuing and assigning certificates and pseudonyms for communications.

b. **RSUs (Road Side Unit)**

These entities are the subordinates of the TAs. They are installed along the roads. They can be mainly, traffic lights, streetlights or others. Their main responsibility is to support the TA in traffic and vehicle management. They represent access points to the network and to the various information on the traffic.

c. OBU (On-Board Unit)

These are units embedded in smart vehicles, they bring together a set of high-tech hardware and software components (GPS, radar, cameras, various sensors and others). Their roles are to locate, receive, calculate, store and send data on the network. These are transceivers that ensure the connection of the vehicle to the network.

➢ Modes of communication in VANETs

In VANETs networks, we mainly find the fixed entities that constitute infrastructure (RSU and TA) and mobile entities (vehicles). To can exchange the various information and data related to the safety and comfort of road users, these different entities must establish communications between they. For this reason, there are two types of vehicle-to-vehicle communications. Vehicle (V2V) and infrastructure vehicle As shown in Figure 2.

a) Vehicle-to-vehicle communication (V2V)

This type of communication works using the devices installed in the vehicles called OBU (On-Board Unit), following a decentralized architecture. He is similar to the type of communications between mobile nodes in MANETs.

Communication between two vehicles takes place directly, in Ad hoc inter vehicular mode. They do not need to rely on infrastructure to be able to communicate between them. Provided that each vehicle is within reach of the other (radio area). Otherwise, they call on other vehicles, which will play the role of a bridge (intermediate) for them. This type of transmission is called multi-hop communication.

b) Vehicle-to-infrastructure communication

Vehicle-to-infrastructure (V2I) communication is also called a communication in infrastructure mode. This mode of communication is ensured thanks to the various entities of the VANETs network. Indeed, the OBUs (On-Board Unit) of vehicles, the RSUs (Road Side Units) placed at the edges of the roads and even the Tas (Trusted Authority) all contribute to each other to ensure communications in the vehicle network. This mode of communication provides connectivity relatively strong compared to communication in V2V mode (vehicle to vehicle). Like him ensures better use of network resources.

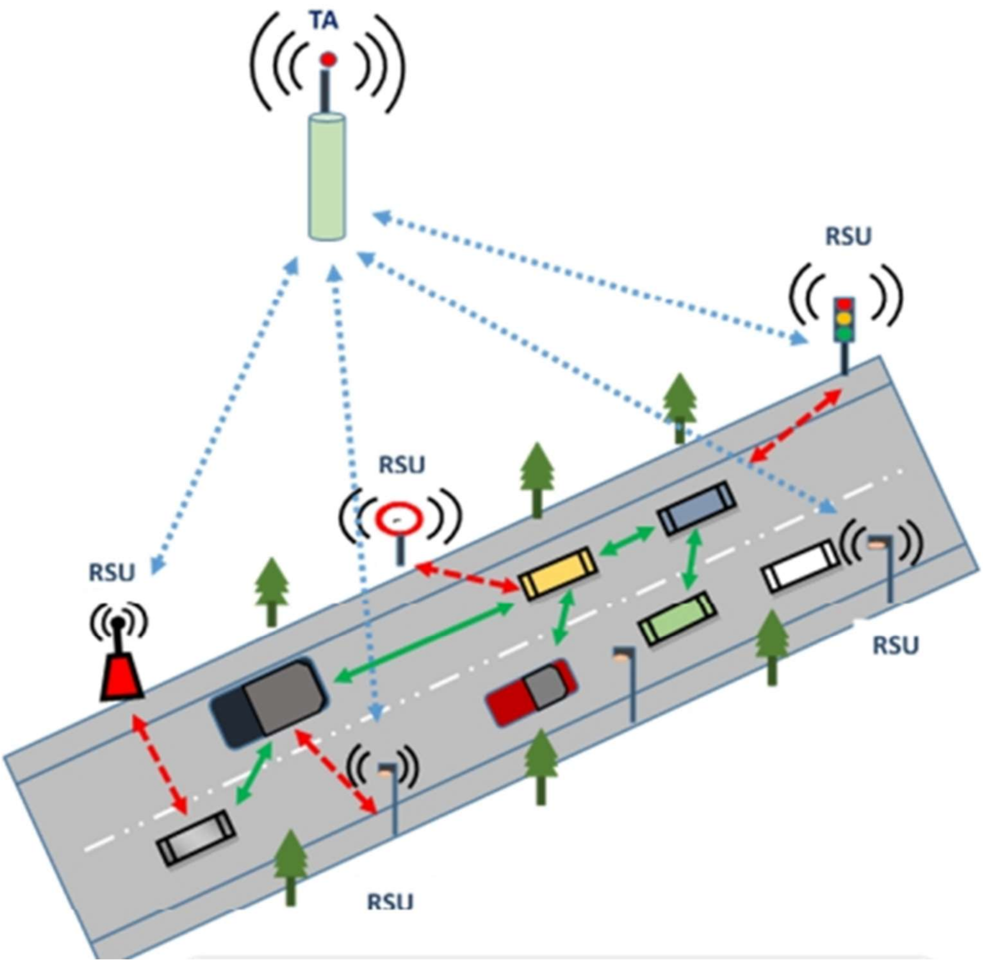

Figure 2 : Model of VANETs networks and its mode of communications.

1.2 Location-based services

The increasing availability of location-based technology (e.g.,GPS) enables people to add a Location dimension to existing online social networks in a variety of ways, such as uploading geotagged content (e.g., photos, recorded videos, and routes), sharing current location, commenting on an event that is taking place (e.g., via Twitter), or leaving notes, tips, opinions about a place (for example, a restaurant). Facebook(, Google, Twitter, and Instagram are examples of mobile LBS applications. End-users can use these applications to share information with their peers, connect with friends and be notified when they are nearby, explore real-world places and events, and receive geolocated adverts. Although these applications are mostly designed by businesses, numerous studies make use of the content generated by users of these applications to make location recommendations.

1.3 Location collection devices

Devices in all locations are based on the Global Positioning System (GPS). GPS data, which provides precise information to the Global Positioning System (GPS), provides the highest precision movement routes. A GPS tracker is a device that receives satellite signals. The gadget periodically receives GPS satellite signals and calculates its position. This allows researchers to understand people's movements to a large extent. Accuracy and temporal frequency, result in a large stream of data that can be analysed them and match them to human movement patterns. when the engine is running, each signal is converted into a data point containing information about the position, speed and distance travelled. Admittedly, there are inaccuracies in the data due to variables such as satellite coverage, location accuracy, and high time accuracy, as it is the distance travelled and the instantaneous speed. GPS tracking records location with excellent accuracy and consistency. But there are also certain disadvantages GPS transmitters cannot work indoors, and generally use batteries as power sources therefore, the tracker may not receive a signal for a period of time or may stop prematurely. In addition, the number of GPS individual users in the dataset is usually small. compared to mobile data that can provide navigation information [1].

1.4 Location data type

We begin by summarizing many popular location-based social datasets. Several real location-based social network datasets are available from various web providers.

We outline the basic features of the datasets after briefly introducing these real-world services. Table 1 summarizes the datasets discussed in this subsection.

Name	Type	Statistics
GeoLife	GPS trajectory	trajectory consisting of 13,432 GPS data points
Brightkite	check-ins & friendships	3000 check-ins From 35 000 users
Gowalla	check-ins & friend ships	60000 check-ins From 75 600 users

Twitter	geo- tagged tweets	22,506,721 tweets from 225,098 users
Foursquare 1	check-ins	12,000,000 check-ins from 679,000 users
Foursquare 2	check-ins, friendships, user profiles & venue information	325,606 check-ins from 80,606 users
Foursquare 3	check-ins, friend ships, user profile & venue Information	three sets of check-ins from 33,596 users

Table 1:LBS datasets used in location prediction and recommendation.

1.5 Location recommendation

1.5.1. Definition

The goal of the POI recommendation is to mine users' check-in information and propose POI to users in LBSNs. Consider Foursquare. Figure 3 shows how check-in information is recorded, including user name, POI, check-in time stamp, and geographical information on the map.

Figure 3 : Foursquare check-in information demonstration.

1.5.2. Basic definitions and problem formulation

The problem of POI suggestion by mining users' past position data in LBSNs has been formulated in this section.

Definition 1 (POI)

A POI is a geographical location with specialized functionalities (for example, a hotel, restaurant, museum, or retail) that a user may find useful or interesting.

Definition 2 (User Check-in)

A user check-in is composed of a three-tuple (u, l, t) indicating that the user u visited the POI l at time t. Then, given a dataset consisting of user check-ins, we hope to deliver POI recommendations to the user.

Definition 3 (Check-in sequence)

A check-in sequence is a collection of user u's check-ins, denoted by $S_u = \{< L_1, t_1>, \ldots, <L_n, t_n>\}$, where ti is the check-in time stamp. Su $= l_1, \ldots,$ ln is used for convenience.

Definition 4 (POI recommendation)

Given all users' check-in sequences S, POI recommendation aims to recommend a POI list S_n for each user *u*. Here S is a collected check-in sequence set, that contains all sequences S_u for all users.

1.6 Mobility Prediction applications

The applications are categorized according to the location prediction job, which is broadly classified as person and object, and further classified as next or current location (predicts only a single future location) and series of locations.

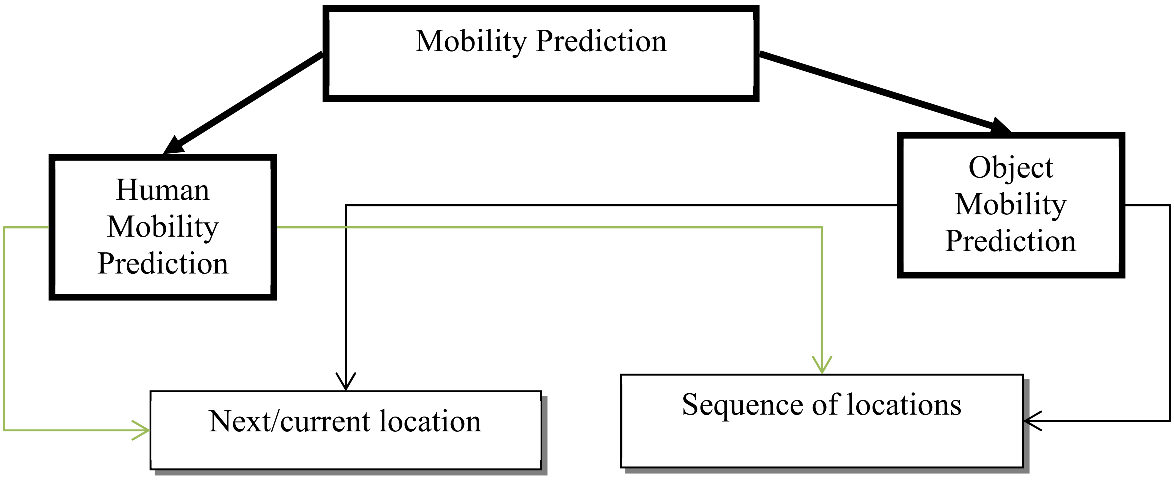

Figure 4 : Categorization of mobility prediction applications

Predicting the next or current location, as well as the sequence of locations, is based on selecting the best one in terms of cost, time, and space.

1.6.1. Human Mobility prediction

Human movement differs significantly from that of objects on a big scale. When compared to objects, the factors considered to mine human movement are massive. Human mobility is influenced by both external and internal forces. This section discusses some of the applications that have been created to mine human movement patterns:

1.6.1.1. Next or current location

Through mobile devices, mobile service systems provide vital movement information. Mobile service systems may successfully mine an intended request from massive data sets based on dynamic user behavior. in Chen et al.[2] A unique mining approach was presented to mine movement patterns utilizing linked trees for this purpose. It also includes a technique for counting the weighted support values from the matching candidate frequent patterns, which increases efficiency and scalability. Also Huang et al. [3]. understands and predicts the destination using mined movement patterns based on commonly visited location data. Because activity transitions are more accurate than location transitions, the Hidden Markov Model (HMM) is used to estimate them.The model proposed in Vukovic et al. [4].

mined user-specific movement data with the assistance of multi-agents, which increases prediction performance even further. All of these models predict only the human's current and next location.

1.6.1.2. Sequence of locations

The model in Farrahi et al[5]. predicts people's routines using data from mobile phones. It is implemented using a novel probabilistic approach based on location, time variance in people's habits, and proximity information. The forecast is based on the unsupervised discovery of the cell phone's location and proximity. Bluetooth knowledge the models presented above are discussed in three different scenarios, moreover Farrahi et al [5]. the first being a general proximity problem applicable to any site, the second in Kolodziej et al [6]. a small university campus, and the third in Lam et al. [7] .a staff tea room. All of the models forecast a person's next sequence of movements.

1.6.2. Object Mobility prediction

Unlike human movement pattern mining, mining objects is simpler because fewer elements are impacting them. Object pattern mining applications are explored in this section:

1.6.2.1. Next or current location

They presented a Heterogeneous Tracking Model (HTM) in Peng et al. [8], which mines object moving patterns while also tracking them. They employed a variable memory Markov model to overcome the dependency difficulties between object movements. HTM is a hierarchical system that enables multi-resolution object moving patterns. This model anticipates object movements and efficiently uses energy while tracking the object.the model presented in Liao et al. [9]. predicts the next location of the object accurately which increases the network lifetime and outperforms in Peng et al. [8], terms of energy efficiency and tracking accuracy.

1.6.2.2. Sequence of locations

The existing object mobility prediction methodologies focus mainly on the geometric properties of trajectories. The geographic information and semantics are not considered for prediction. in O. Alvares et al. [9], they have proposed a reverse engineering framework to mine and model the patterns of semantic trajectory data. the model in Lee et al.[10] .guarantees the correctness of the mined patterns and assures a better performance in terms of memory usage, runtime, and scalability.

1.7 Route Prediction

1.7.1. Definition

Many drivers now utilize various types of route prediction algorithms to find better driving routes. Route prediction systems are frequently utilized, and they provide a wide range of services for both the driver and his vehicle, such as intelligent transportation systems. It can be used to display traffic alerts as well as location-based services (e.g. systems that determine vehicles paths, recommendation route systems that play an important role in many applications like the display of targeted advertisements about points of interest and shops that a user is approaching). Furthermore, it contributed to lower fuel usage, providing significant economic and environmental benefits. Route prediction is the process of forecasting a driver's next route based on his present path and his previous so that it can be viewed as an instance of the problem of sequence prediction

1.7.2. Basic definitions and problem formulation

Definition 1 (Road network)

A road network is a directed graph (N, RS), with N representing the road junctions (nodes) in a particular area and RS representing the edges.

Definition 2 (Trajectory)

A trajectory is a series of consecutive GPS points that define a person's mobility. As a result, Traj=$p_1 \rightarrow p_2 \rightarrow \cdots \rightarrow p_n$, where pi (longitude, latitude, timestamp)P is the GPS points set, and P is the GPS points set.

Definition 3 (Stay point)

A stay point is a geographical area where a person stays for a duration longer than a temporal threshold Tthr and the distance between GPS trajectory points is less than a spatial threshold Dthr.

Definition 4 (Trip)

Stay points are trip points because a trip is a GPS trajectory that begins and ends with a stay point.

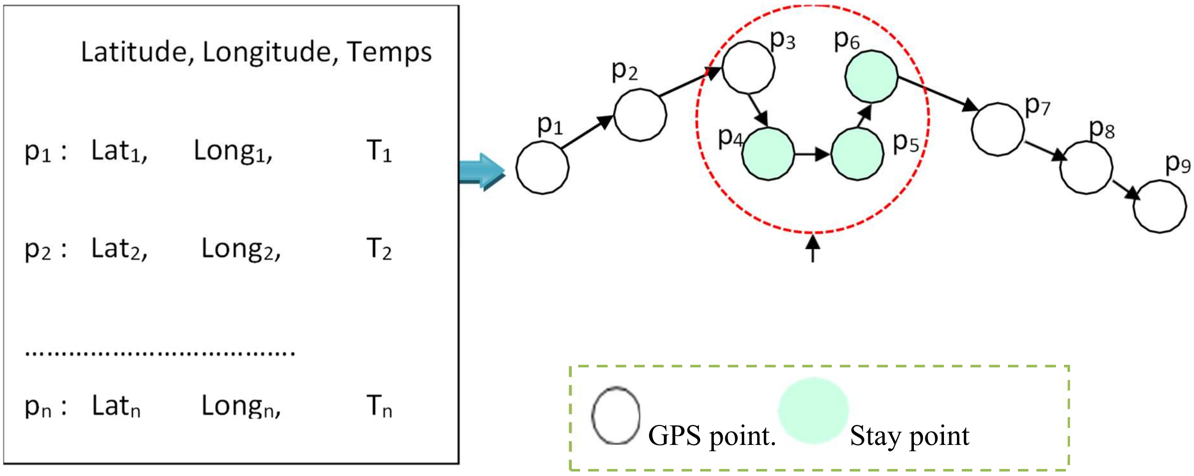

Figure 5 :GPS Trajectories and stay point.

1.8 Related work

1.8.1. Taxonomy of mobility prediction models

In this section, we aim to present the most important existing models for route and destination prediction. To do so, the taxonomy depicted in Figure 7 classify them according to technique used: 1)Trip Matching models, 2) Data Mining models, and 3) probabilistic models.

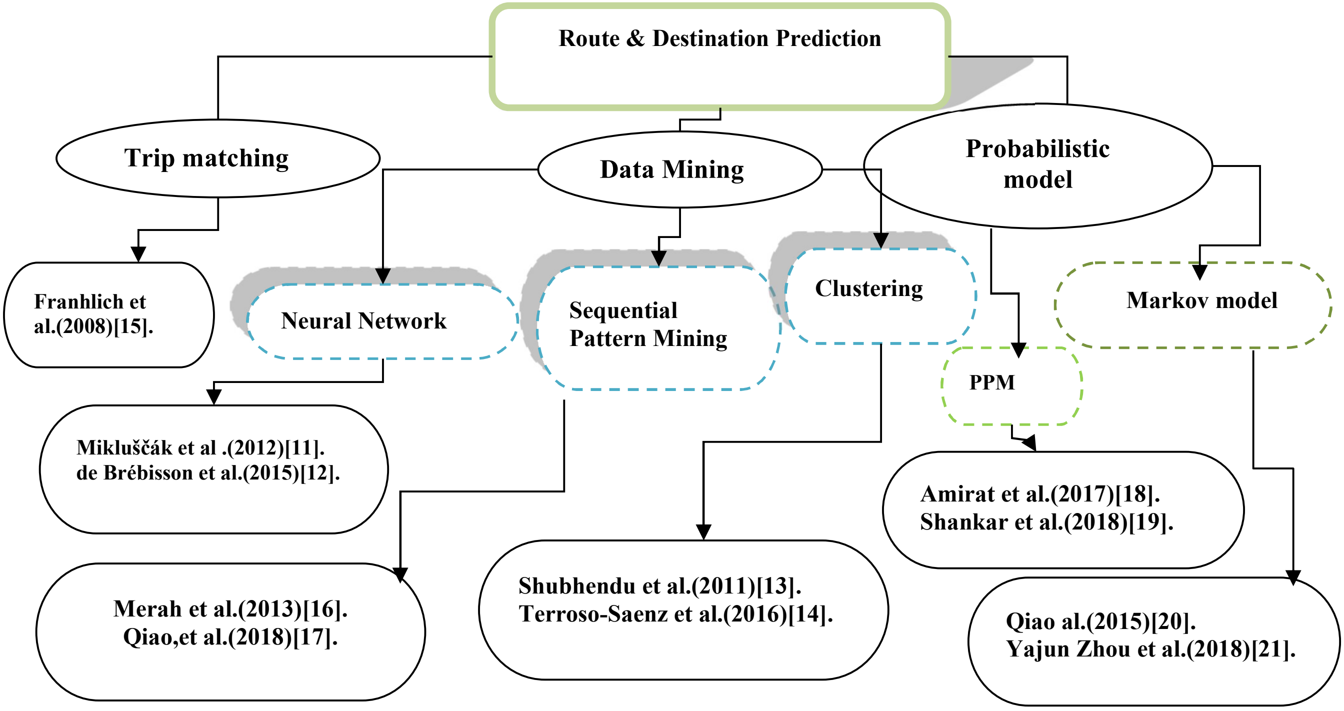

Figure 6 : Taxonomy of route prediction.

➢ **Data mining models:** Data mining techniques have been broadly applied to perform various types of predictions using techniques such as neural networks and sequential pattern mining**.**

A) Neural Network

Artificial neural networks(ANN) have been used for route prediction . For instance, in the work of, the model proposed by Mikluščák et al.[1] aims at 1) examining Artificial Neural Network Applications (ANN) for prediction, 2) selecting the best architecture for the ANN, and 3) determining the best data representation. In addition, in de Brébisson et al. [2] a neural network approach to predicting the destination based on the prefix is proposed. Various embeddings to encode the information and destination clusters to construct the output is suggested.

B) Clustering

Clustering was utilized to forecast destination and route. In the work Shubhendu et al[3]. They observed that this use of a predictor in clustering improved the prediction accuracy in most datasets.In recent work for personal route prediction (Terroso-Saenz et al. [4], an online density-based clustering algorithm called landmark discovery algorithm (LDM) has been applied for Meaningful Velocity Areas identifcation where user velocity (speed and direction) remarkably changes .

C) Sequential pattern mining

The term "sequential pattern mining" refers to the process of identifying and exploiting In the work .Merah et al. [16] .have presented several communication schemes that could be used to collect historical vehicular paths. Movement patterns are then extracted as the most frequent traveled paths and used afterward to generate movement rules that could be used to forecast vehicle future routes.,.Moreover in Qiao et al.[17]. have offered an incremental trajectory matching strategy with three matching strategies for mining frequent sequential patterns across postfix sequences iteratively.

- **Trip Matching models**

In Froehlich and Krumm , the authors have proposed a long-term route predictor that combines several techniques, including hierarchical clustering for trip clustering and the Harsdorf metric for calculating the similarity between pairs of trips.

- **Probabilistic models**

Probabilistic models such as Markov model has been used in Qiao et al.[10]. The authors They provide a Hidden Markov Model-based driving route prediction system (HMM). Furthermore, in Yajun Zhou et al.[11].the authors have constructed a prediction system based on a Markov chain model of location demand that outperformed previously developed systems. In the work Vishnu Shankar et al .[17]. Scalable prediction by partial match (PPM) and its application to route prediction for GPS location was decomposed into smaller units called user trips. User trips were map matched to the road network to convert the data into a set of edges. This step is part of data preparation, which is a one-time activity. Additionally, in Amirat et al.[8]., MyRoute framework has good performance and scalability compared to PPM and LZ predictors.

1.8.2. Taxonomy of POI Recommendation

In three areas: influencing factors, probabilistic models .In particular, we explore four sorts of influential factors: geographical impact, social relationship influence, temporal influence, and sequential influence. Furthermore, we classify probabilistic models for POI recommendation as Markov chain and Collaborative filtering (CF). classify them in the taxonomy depicted in Figuer 8

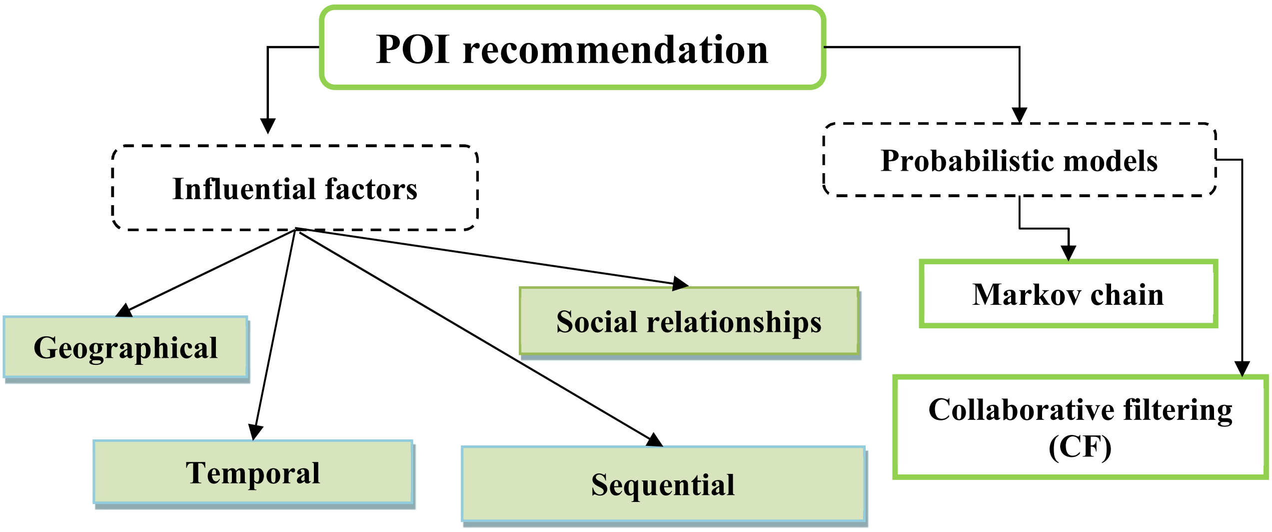

Figure 7 : Taxonomies for POI recommendation.

➢ Influence factors

We categorize POI suggestion research based on numerous characteristics that influence user check-in activity.

a. Geographical factor

Geographical influence is an essential aspect that distinguishes POI recommendations from standard item recommendations, as evidenced by check-in behavior data.in Ye et al. [22] proposed a power-law distribution model to capture the geographical influence, and proposed a collaborative POI recommendation algorithm based on geographical influence via naive Bayesian. Because it is difficult to find an anchor point to derive a reasonable distance for the new POI. As well as Zhang et al. [23] propose a kernel function for modeling geographical influence. The two-dimensional kernel function is more reasonable than the one-dimensional distance power-law distribution. The distribution of user check-ins is uneven. Some previous studies assume that user-checked locations follow a Gaussian distribution with numerous centers. Moreover, Liu et al.[24] create a broad geographical probabilistic factor model (Geo-PFM) to capture the geographical influence on user mobility behaviors, which they then integrate with Bayesian nonnegative matrix factorization (BNMF) to describe user preferences. Finally, POIs are suggested to users by aggregating the influence of several elements.

b. Temporal factor

The importance of temporal effects in POI suggestions cannot be overstated. Yuan et al .[25] found out that most users tend to visit different POIs at different times in a day, and the check-in behaviors between neighbor time slots are similar. In addition to that, He et al.[26] propose a spatial-temporal topic model (STM), which embedded the temporal and spatial patterns in users' check-in activities. Besides, Oppokhonov et al.[27] developed a recommendation system based on a directed graph. The algorithm of the system considers both the temporal factor and the distance for recommending a new POI for the next hours.

c. Social relationships factor

Users' choices of new POIs are also influenced by their social relationships. Based on the friends in their networks and the places they have visited Cheng et al.[28] used probabilistic matrix factorization with social regularization, and the social constraints ensured that latent traits of friends remained near to the latent subspace. Furthermore, Li et al.[29] established three types of friends in LBSN (i.e., social friends, location friends, and nearby friends) and proposed a two-step methodology to use friend information to increase POI recommendation accuracy.

d. Sequential factor

in the actual world, users' actions commonly follow one another, and the next action is frequently tied to the prior one. During the last few years, Several research has concentrated on several sequential recommendations tasks. Li et al.[30] introduce a novel neural network model named TMCA that employed the LSTM-based encoder-decoder framework for the next POI recommendation. Furthermore, Li et al.[30] provide TMCA, a novel neural network model that uses an LSTM-based encoder-decoder framework to recommend the next POI.

➢ Probabilistic models

In this section, we categorize the POI recommendation systems by the probabilistic models using the influential factors mentioned above.

a. Collaborative filtering

Collaborative filtering (CF) recommendation method is widely used in recommendation systems, in Wang et al.[26]. In the context of the social internet of things, they proposed developing a trust-enhanced collaborative filtering framework for POI recommendation in LBSNs. Also Yin et al. [28]. To provide time-aware POI recommendations, a collaborative filtering technique is proposed. tensor decomposition is used to tackle data-sparse issues, and in the tensor decomposition phase, location is substituted with category.

b. Markov Chain

Markov chain recommendation method is widely used in recommendation systems in LBSN. in Yin M et al.[28]. SONG is a unique sequential prediction approach based on the Markov chain model. They employ a variable order additive Markov chain to jointly represent user behaviour and geographical influence.

7.1.Literature review for route prediction and POI recommendation

In this section, we present a table that shows a comparison between different The route, destination, and POI recommendation prediction algorithms discussed above. When these Models are compared according to certain criteria such as the technology used and type prediction and the type of data used in empirical evaluation, accuracy,

Study	Type prediction	Technique	Data	Accuracy	Type technique / Influential factors
Amirat et al. (2017) [18]	Route and location prediction	GMG and PMG (Graph-Dependency)	GPS	76% PMG 83.5 GMG	PM
Wang et al.(2020)[31]	Location	Collaborative Filtering (CF)	check-in Foursquare Gowalla	precision@k= 4.7% recall@k =3.3%	Geographic -al Temporal
Gan et al. (2018)[32]	Route and location prediction	parallel sequential pattern mining (PSMP)	trajectory data the sequence database	/	DM
Yin M et al. (2021) [33]	Location	Collaborative Filtering (CF)	Global positioning system (GPS) Trajectory check-in locations	Then TDCF algorithm nearly improves by 244%, 103%, and 45.2%	Temporal
Francisco et al (2016) [34]	Route and location prediction	Prediction by Partial Matching (PPM)	smartphone Global positioning system (GPS)	the precision of 76.9%	PM
Mourchid et al (2016)[35]	Location	Markov chain model	mobile location	Vmin and Vmax are set	Temporal social

			services Foursquare GPS check-in	to 40% and 70%,	
Kaur et al (2022)[36]	Route and location prediction	Hidden Markov Model	training dataset	54.23% of improvement in the lookup success ratio and 59% improvement in the maintenance	PM
Zhong et al .(2020)[37]	Location	Collaborative filtering (CF)	Dataset(Yelp, Gowalla, Foursquare)	$\lambda = 0.125$. Note that it is better	Social
Qiao et al.(2017)[38]	Route and location prediction	Hidden Markov Model	Actual plan traffic control collected from a GSM LTE, UMTS, GPS or Wireless	over 80%	PM
Li et al.(2021)[39]	Location	Markov chain model	Datasets(Beijing Jiepang Shanghai Foursquare). data to predict future check-in.	13%, 6%, 15%, and 10%	Special-sequential
Amirat et al. (2019)[40]	Route and location prediction	CPT model	Real GPS (driving) Data set synthetic	35% - 96%	DM
Yang et al.(2017)[41]	Location	Collaborative filtering (CF)	the data scarcity on labeled user preference, i.e., the check-in data	99.860% sparsity	Social

DM data mining, PM probabilistic model, Location = POI recommendation..

Table 2 :Literature review for route prediction and POI recommendation.

7.2.Discussion

The table above provides a comparison of the path prediction and positioning methods discussed in this section, as well as a comparison of our methodology. In this table, models are compared using three basic criteria: (1) prediction type (2) technique, and (3) technique type/influencing factors infleuntial observations can be made from this table. First, we look at the prediction. It is clear to us that the highest accuracy and correctness of the results is when using the GMG, PMG, and PPM technology of more than 76%, similar to the Graph-Dependency technology, which has a very higher health rate because it has the ability to handle GPS data unlike most sensitive proposals preceding the noise. As for the recommendation, the results differ according to the different influencing factors, as we used two techniques, namely, cooperative filtering and Markov design, where we note that what affects cooperative filtration the most is the social and thermal factors. As for the Markov design, the chances are greater in the Markov chain in terms of accuracy. Special forecasting work is used based on the movement of vehicles and people's visits to places and is inspired by the DG predictor.

Chapiter 2:

The MyLoc model for route and Location prediction .

2.1.Introduction

Mobility prediction of human and objects have a variety of application benefits such as tracking of a mobile device, vehicle route optimization, human activity time optimization, etc. To perform mobility prediction, some research focused on predicting the movements of moving bodies, which is important and has many applications in several fields, to obtain information. Most of these predictions depend on the short-term prediction of the next few potential locations. In this chapter, a new location predictor and recommender is designed to perform short and long term predictions. Our proposal adopts a geograhical representation of the movement by using the DG predictor.
This chapter is organized as follows . First, the set of definitions is defined followed by formulation the problem we are dealing with. Next, we provided a detailed explanation of the system architecture for *My Loc* and finally, we end this chapter with a conclusion.

2.2.Background

2.1.Problem Definition

Definition 1 (Road segment)

A road segment is an abstraction of a vehicle or a driver location. A road segment r_i is a directed edge between two junctions For example, Figure9 depicts road segments r_x and r_y connecting two junctions [18].

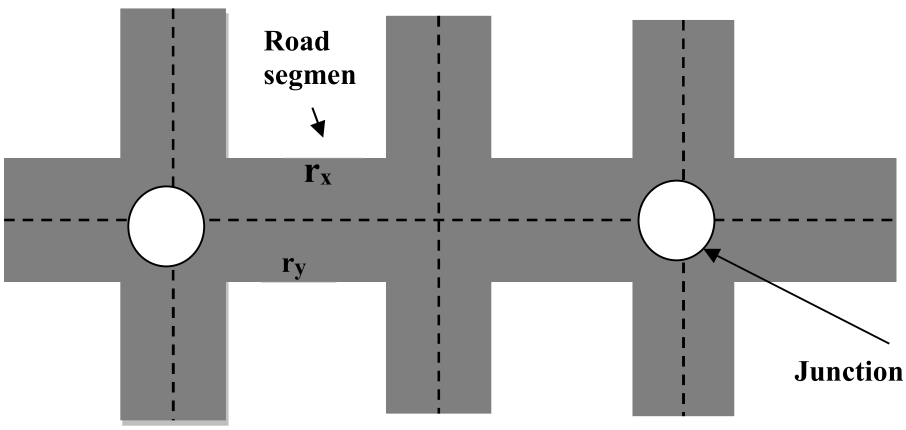

Figure 8: Illustration of a road segment.

Definition 2 Location

A location is a geographic area, which has a unique identifier L_i and has been visited by a user u_i where L_i in L (the set of locations in a geographical area) and u_i in u (the set of users in an LBSN).

Definition 3(mobility or movement sequence)

A mobility sequence *ms* is a sequence of road segments traversed by a user during a trip and the location he has visited during his mobility. For instance, Table 3 depicts five sample of mobility sequences ms:{ms_1,ms_2,ms_3,ms_4,ms_5} performed by three users u:{u_1,u_2,u_3}, representing their paths. For instance, the user u_2 has a mobility sequence ms_3 which indicates that u_2 has visited L_1 and route segment r_1 followed respectively by r_2 and r_3 then visiting L_2.

User ID	Mobility sequence
u_1	ms_1:(L_1) , r_1 , r_2 , r_3 , (L_2) ms_2:(L_1) , r_1 , r_2 , r_3 , r_4 (L_3)
u_2	ms_3:(L_1) ,r_1 , r_2 , r_3, r_4 , (L_2)
u_3	ms_4:(L_3) , r_5, r_6 , (L_2) ms_5:(L_3) , r_4 , r_5 , r_6, r_3 (L_2)

Table 3:A sample of mobility sequences.

Definition 4 Location History

The location history (LH) of a user u_i is the set of all location sequences of U_i.

2.3.System architecture

In this section, we introduce our proposed framework *MyLoc*, which forecasts the driver route and location relying on a graph representation. *MyLoc* is originally designed to regard the sequential nature of human mobility and could be easily extended to incorporate all the influence factors on user mobility, i.e. temporal and sequential. In the following, of the framework describe its design in detail.

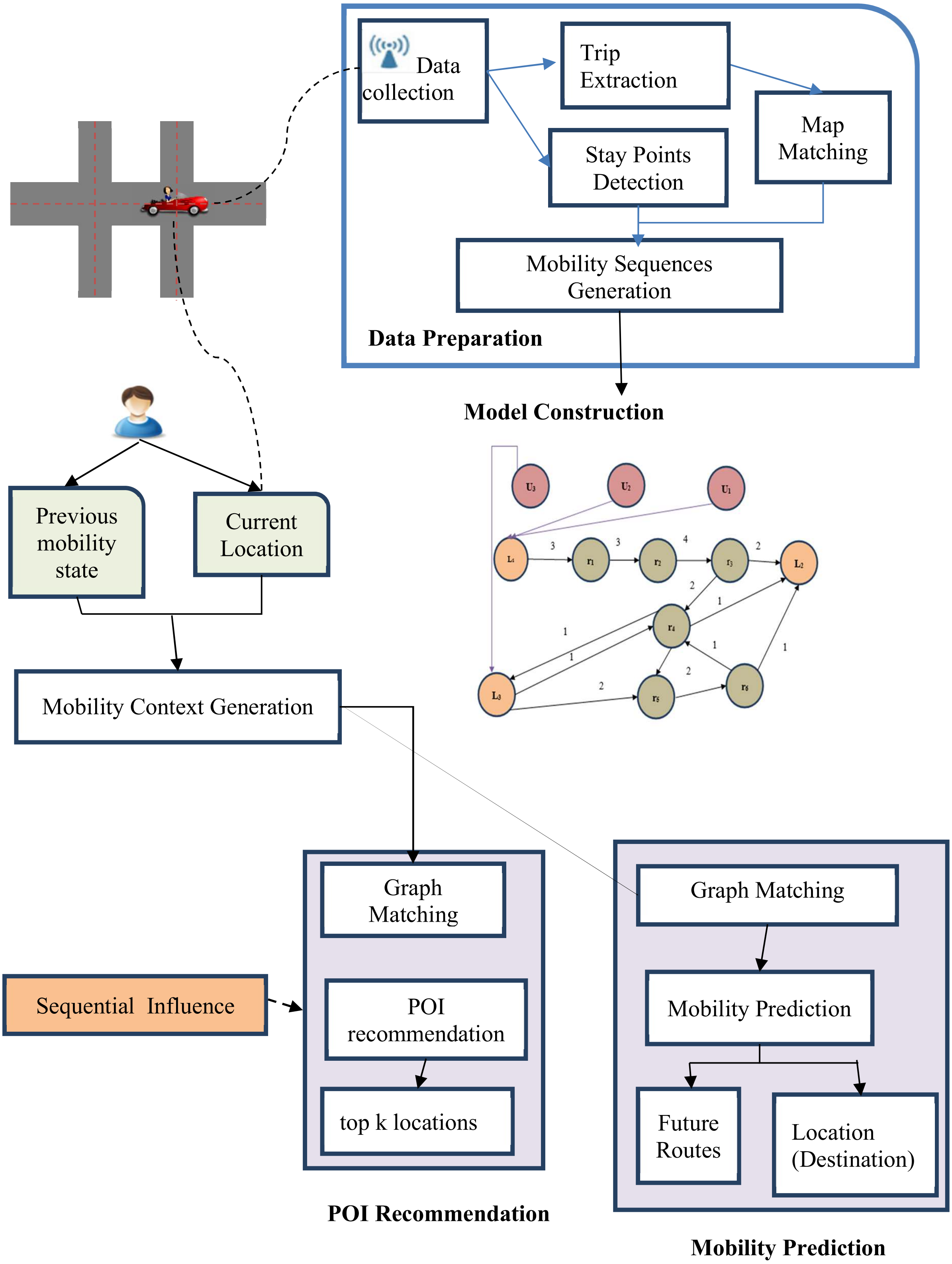

Figure 9: Architecture of *My Loc*.

3.1.Data preparation

It comprises two-sub steps: data collection and mobility sequence generation.

A. Data collection :

The first module periodically collects driver location data (GPS records) and sends it to a server site. During collection, location data is split into trips by defining stay points. A stay point is a geographic area expressed as a set of consecutive GPS records where the distance from the first and last GPS records exceeds a distance threshold D_{thre} and the driver spent more time than a threshold T_{thr}. The resulting trips are then converted into mobility sequences by map-matching GPS trajectories using a cloud map-matching based API [42].

Dissemination mode is to whether the collected GPS logs are broadcast to surrounding RSUs or only sent to the nearest RSU in a unicast manner. It determines whether the navigation is

The data is collected and then automatically sent to RSU by car or upon request from RSU.

B. Mobility sequence generation

Definitions Mobility sequence generation aims at transforming GPS trajectories collected from the previous step into a set of mobility sequences that comprises both locations visited by user and routes taken to reach those locations. This step consists of three sub-stages namely stay point detection, map-matching and sequence generation. The stay point detection stage aims at identifying geographic areas where the person remains stationary for a definite period of time. Stay locations are can be determined as a set of consecutive GPS points, where the distance from the first and last GPS point exceeds a distance threshold D_{thre} and the driver spends more time than a threshold T_{thr} in that area. Once the set of stay points are determined, they will be extracted from trajectories to obtain the traveled trips on the road network. These resulting trips are then transformed in the map matching phase stage, into sequences of roads using a cloud map-matching based API [42]. The Map matching process consists of finding the corresponding road segment,in a road network, for GPS records. At the end of this step, mobility sequences could be generated. A mobility sequence ms=(Li) , r_1 ,r_2 ,..,r_n ,(L_j) represents the whole travelled trip. It comprises the stay location from where the driver starts its movement (L_i) followed by an ordered list of traversed road segments (r_1, r_2 ,.., r_n) and then his intended destination (L_j). However, time is continuous. Generating mobility sequences it would considerably decrease the quality of prediction, as only few users would visit exactly the same place at exactly the same time. To address this issue, the time is discretized using bins.

3.2.Model construction

The problem of route prediction and POI recommendation(location), we suggest a sample navigation graph inspired by the Dependency graph (DG) predictor. DG is a graph-based sequence prediction model that was initially proposed for prefetching from the web, which represents system dependencies such as graph nodes, as the corresponding Figure 11 shows mobility graph (MG) constructed from the mobility history depicted in Table 3.

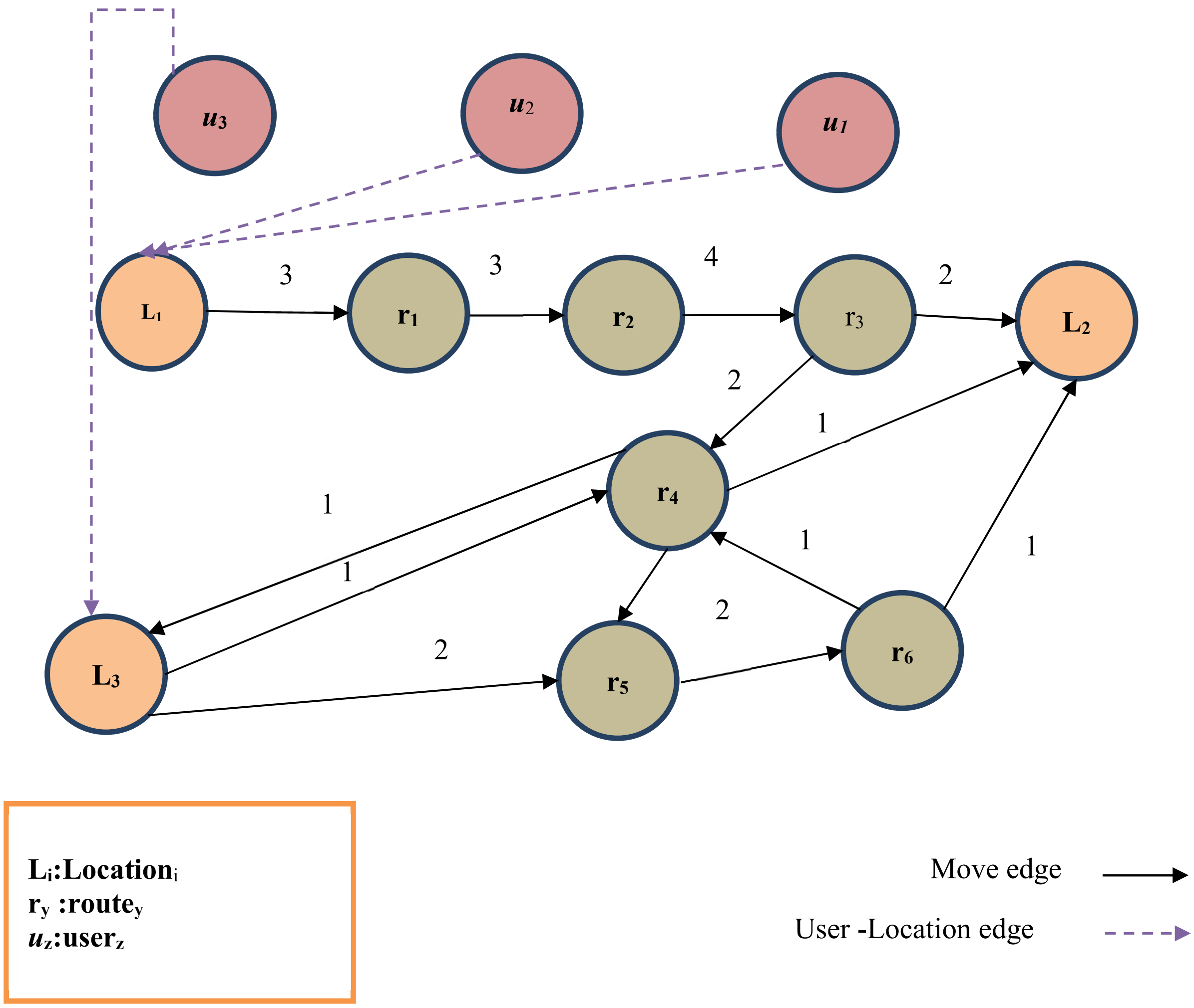

Figure 10: Mobility Graph model.

The MG mobility graph is a directed graph of G = (N, E) where NR(Road nodes) and NL(Location nodes) is the set of graph nodes and E= N × N means the set of edges denotes the set of edges linking graph. The set includes N = {u ∪ r ∪ L} for nodes of the graph. The set E of the edges of the graph is defined as E = { EM ∪ EuL}.

EM: the set of moving-edges used to represent/model the movements in mobility sequence (e.g.ms=(Li),r_i,, r_x, r_y, r_m,(L_j)) that could be 1) from a road r_x to another r_y ,2) from a starting location Li to the following/first ri road immediately traversed after Li or 3) from a road r_m to the final destination (L_j).

EUL: Defines a set of user location links/edges between the user node user interface and the site node location. User- location uL_i is created between the user node u_i and my location file if there is a journey through the u_i where L_i is the start. MG allows to represent system dependencies between disabled and location methods. In the driver mobility sequence where the given arc source must appear before its destination. The mobility prediction graph was generated by stepwise insertion of navigation sequences into the graph, where each movement of a sequence is represented as a path comprising the roads it contains. If you share the sequences then a significant reduction in space can be achieved by using MG representation. For example, a graph created using the mobility sequence in the first table with the lookahead window defined by w = 2.

3.3.Mobility Prediction

The module of mobility prediction is used to predict the future movement of the user's destination, as it studies the current state Mobility Pattern (MP).of mobility with the user with any other precedent, as it includes the user's mobility context so that the user's global location determines the way and the location of stay point. The mobility state can extend or include others for contextual information such as user knowledge if the personal prediction.

3.3.1 Route and location prediction

Having the mobility context (MP), the prediction of the future route segment or location can be performed as follows.

1) Graph matching:

This step consists of discovering the path P is similar to MP in that it corresponds to MP. P is said to be similar to MP however, the use of graph matching is limited MG can only find a path that is comparable to its own. context of navigation if a similar path cannot be discovered, the prediction was not made or failed to be made. Eases we can overcome this issue by making our structure more flexible. It is suggested that a recursive partial matching process be implemented. This final excludes the last location of each step to partial context or the context becomes empty trying to discover a path that is comparable to the partial path, created context if the path is the same, the matching stops.

2) Prediction:

The prediction is made using the similar path set collected in the previous step (route and/or location prediction). If the goal is to predict the route or the location, or both the future route and the final planned destination of the route and to predict the location, it aims to retrieving future routes and future locations. In the input deputy, we explain how route and location prediction is accomplished regardless of the context of mobility. The last step in the prediction, given the associated route, is to retrieve sorted animated edges, indicated by in the scenario where the acquired context is a route, a traversal graph the user's intended position must be determined by following the edges with higher weights. If the context is a location, however, *MyLoc* begins at that position and keeps the edge of the moving edge with the highest weight between continuous movement edges to determine the best future route. Several selection criteria, such as the retention route with the highest frequency in the mobility data, can be utilized if the weights of the candidate edges are equal. equals indicate that only series integration was discussed. These parameters have a significant impact on elements such as the route forecast and location.

4.1.POI recommendation

MyLoc uses the users travel graph as input, and the output is a list of top k locations, but the social impact is a sub-graph of the user and their general friend network rather than the entire graph. Regarding the navigation prediction method, our software first selects a set of candidate locations, but depends on the geographical proximity of the candidate locations closest to the current location, and then displays the closest candidate places as viable destinations for the user.

2.4.Conclusion

In this chapter, we have introduced our *MyLoc* system that allows the prediction of the user's future movement in the short and long term. Also provides suggestions on locations that may interest the users in visiting while the user is moving. It also takes into account the sequential context as we have explained.

Chapter 3:

Experimental evaluation

3.1 Introduction

This chapter presents an empirical study o evaluate the performance of the *MyLoc* prediction model and to compare it with to another prediction model, which is the Markov model (Makov-Pre model.) as a basic representative predictor.In this chapter, we first give the description of datasets used followed by the experimental settings in section 2. In the third section, the formulation of the evaluation metrics is given. In the fourth section, a description of the experiments conducted is presented. We conclude this chapter with a conclusion.

3.2 Experimental seetings

The experimental evaluation conducted in this section is performed using a realistic large scale dataset called Gowalla [43]. The latter is well used for the task of location prediction such as in dataset a subset of check-in locations from 1000 users was utilized. For both datasets, a 10-fold cross validation was used.. Our experimental environment is composed of an Intel(R) Core(TM) i5 CPU 2.67 GHz with 4 GB of RAM with Windows10 (64bit). The proposed model has been implemented in Java by adapting the DG implementation available in the SPMF library [44].

3.3 Evaluation Metrics

To measure the performance of *MyLoc*, the two following metrics were computed .

- ***Overall Accuracy***

It is defined as the number of successfully predicted routes and location, divided by the total number of test mobility sequences.

$$\text{Overall Accuracy} = \frac{\text{number of successful predictions}}{\text{number of testing mobility sequences}}$$

- ***Coverage***

It is calculated as the number of testing sequences where a matching path was found for each current trajectory of a driver divided by the total number of test sequences.

$$\text{Coverage} = \frac{\text{number of matching path}}{\text{number of testing mobility sequences}}$$

3.1 Experiments

In this section, we present the experiments conducted to measure the performance of our proposal for the task of route and location prediction. *MyLoc* is compared with the Makov-Pre model.

Markov-Pre (First order markov): This model is based on Markovian property that the predictions are based on the last elements of a sequence. It can be represented as a graph where the prefix sub sequences are linked to the suffix sub sequences by outgoing arcs having transition probabilities.

Markov-Pre has been also modified in our work so it considers both types of mobility locations (routes and POI).

Experiment 1: Impact of varying the lookahead size

In the first experiment (Figure 12), we assesess to measure the perofmance of *MyLoc* by varying the lookahead value from 1 to 8. Overall, the peromance of *MyLoc* increase by increasing the value of lookahead. This may be due to the fact that by increasing the lookahead size the more additional links are created which give more flexibility for prediction and boost the coverage and accuracy by allowing selecting more choices in recommandation.

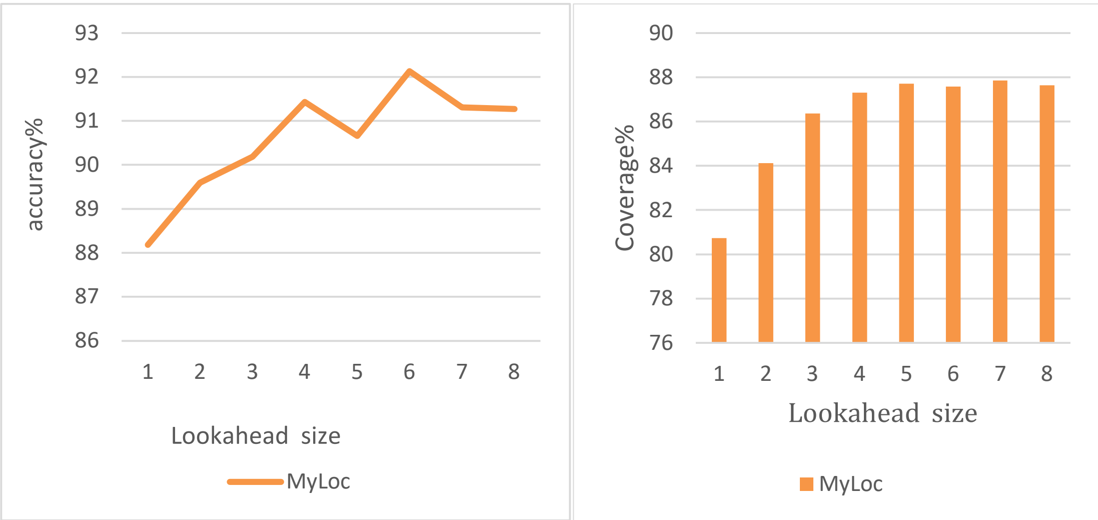

Figure 11 : Impact of varying the lookahead parameter.

Experiment 2:Impact of varying the number of mobility sequences (Scalability)

This experiment was also carried out to study the scalability of the proposed model by increasing the number of location sequences considered. For *MyLoc* a value of lookahead of 2 is fixed. From the obtained results, it is found that prediction performance is improved when the number of location sequences for training is increased. This is reasonable because as more location data is used, more can be learned from this data. Overall, the results also show that *MyLoc* outperforms

Markov-Pre in terms of accuracy and coverage. This may be due to the implemntation of the lookahead that allows creating links in the mobility graph which allows *MyLoc* to achieve the best coverage value of 82%.

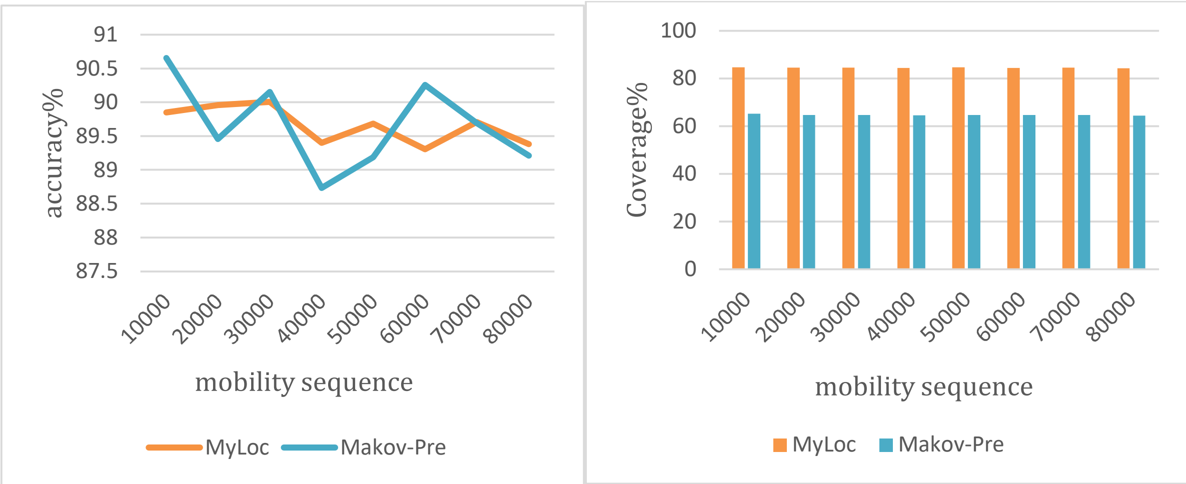

Figure 12: Impact of varying the number of mobility sequences.

3.2 Conclusion

In this chapter, we have conducted of experiments to compare the performance of our *MyLoc* model with Markov-Pre model prediction model using the Gowalla real dataset. The experimental results showed that the performance of *Myloc* model is generally better in terms of accuracy and coverage compared to the Markov-Pre model.

Conclusion general

This dissertation, we propose a *MyLo***c** routes prediction and location system that is designed to consider both spatial factors of the sequence of human location movements. The model is mainly based on the DG model and therefore it provides all its advantages including 1) our framework allows to predict the future movement of the user in the short and long term , 2) the property of no data loss that allows to preserve all the data for making prediction, 3) it provides suggestions about location that may interest users to visit while they are on the move 4) Its ability to predict situations with high accuracy and wide coverage. Our proposal uses a graphical representation of sequences where nodes are location and roads and arc represents a visitation between two location and two roads. In this dissertation, an empirical study was conducted using a set of real-world data. The results showed that the proposed model has promising results compared to the latest models.

Despite the promising results, many improvements, additions and predictions are possible. Therefore, as future work, we plan to extend the *MyLo***c** model so that they consider more temporal, social, geographical information at a POI of recommendation .

Bibliography

[1] Barbosa-Filho, H., Barthelemy, M., Ghoshal, G., James, C. R., Len Ormand, M., Louail, T., Menezes, R., Ramasco, J. J., Simini, F., & Tomasini, M. (2017). Human Mobility: Models and Applications.

[2] W. Huang, M. Li, W. Hu, G. Song, and K. Xie, "Hierarchical destination prediction based on GPS history," Proceedings - 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2013, pp.972–977, 2013.

[3]M. Vukovic and D. Jevtic, "Agent-based movement analysis and location prediction in cellular networks," Procedia Computer Science, vol. 60(1), pp.517–526, 2015.

[4] K. Farrahi and D. G. Perez, "Discovering human routines from cell phone data with topic models," Proceedings - International Symposium on Wearable Computers, ISWC, pp.29–32, 2008.

[5]J.Kolodziej, L. Wang, N. Ghani, S. U. Khan,N. Min-Allah, S. A. Madani, and H. Li, "An application of Markov jump process model for activitybased indoor Mobility Prediction in wireless networks," Proceedings - 9th International Conference on Frontiers of Information Technology, FIT, pp.51–56, 2011.

[6]L. D. M. Lam, A. Tang, and J. Grundy, "Predicting indoor spatial movement using data mining and movement patterns," IEEE International Conference on Big Data and Smart Computing, pp. 223–230, 2017.

[7] W. C. Peng, Y.Z. Ko, and W.C. Lee, "On mining moving patterns for object tracking sensor networks," Proceedings - IEEE International Conference on Mobile Data Management, pp. 0–3, 2006.

[8]W. H. Liao, K. C. Chang, S. P. Kedia, "An object tracking scheme for wireless sensor networks using data mining mechanism," IEEE Network Operations and Management Symposium, pp.526–529, 2012.

[9] L. O. Alvares, V. Bogorny, J. A. F. de Macedo, B. Moelans and S. Spaccapietra, "Dynamic modeling of trajectory patterns using data mining and reverse engineering," Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual Modeling, vol. 83, pp.149–154, 2007.

[10] G. Lee and U. Yun, "A new efficient approach for mining uncertain frequent patterns using minimum data structure without false positives," Future Generation Computer Systems, vol. 68, pp.89–110, 2016.

[11] M. Gregor, T. Miklušćák, and A. Janota, "Using neural networks for route and destination prediction in intelligent transport systems," Commun. Comput. Inf. Sci., vol. 329 CCIS, pp. 380–387, 2012, DOI: 10.1007/978-3-642-34050-5_43.

[12] A. De Brebisson, E. Simon, A. Auvolat, P. Vincent, and Y. Bengio, ´ "Artificial neural networks applied to taxi destination prediction," in Proceedings of the 2015th International Conference on ECML PKDD Discovery Challenge - Volume 1526, ser. ECMLPKDDDC'15. Aachen, Germany, Germany: CEUR-WS.org, 2015, pp. 40–51. [Online]. Available: http://dl.acm.org/citation.cfm?id=3056172.3056178.

[13] F. Terroso-Saenz, M. Valdes-Vela, and A. F. Skarmeta-Gomez, "Online route prediction based on clustering of meaningful velocity-change areas," Data Min. Knowl. Discov., vol. 30, no. 6, pp. 1480–1519, 2016, DOI: 10.1007/s10618-016-0452-3.

[14] Shubhendu Trivedi, Z. A. P. N. T. H. (n.d.). The Utility of Clustering in Prediction Tasks.

[15] Froehlich J, Krumm J (2008) Route prediction from trip observations. In: Proceedings of SAE world congress, vol 2193, p 53

[16]A. F. Merah, S. Samarah, A. Boukerche, and A. Mammeri, "A sequential patterns data mining approach towards vehicular route prediction in vanet," Mobile Networks and Applications, vol. 18, no. 6, pp. 788– 802, 2013.

[17]Qiao, S., Han, N., Wang, J., Li, R. H., Gutierrez, L. A., & Wu, X. (2018). Predicting Long-Term Trajectories of Connected Vehicles via the Prefix-Projection Technique. *IEEE Transactions on Intelligent Transportation Systems*, *19*(7), 2305–2315. https://doi.org/10.1109/TITS.2017.2750075.

[18]Amirat H, Lagraa N, Fournier-Viger P, Ouinten Y (2017) MyRoute: a graph-dependency based model for real-time route pre-diction. J Commun 12:668–676.
https://doi.org/10.12720/ jcm.12.12.668-676.

[19]Tiwari, V. S., Arya, A., & Chaturvedi, S. (2018). Scalable prediction by partial match (PPM) and its application to route prediction. *Applied Informatics*, *5*(1). https://doi.org/10.1186/s40535-018-0051-z

[20] S. Qiao, D. Shen, X. Wang, N. Han, and W. Zhu, "A self-adaptive parameter selection trajectory prediction approach via hidden Markov models," IEEE Trans. Intell. Transp. Syst., vol. 16, no. 1, pp. 284–296, Feb. 2015

[21]Zhou, Y., Wang, L., Zhong, R., & Tan, Y. (2018). A Markov Chain Based Demand Prediction Model for Stations in Bike Sharing Systems. *Mathematical Problems in Engineering*, *2018*. https://doi.org/10.1155/2018/8028714.

[22]M. Ye, P. Yin, W. C. Lee, and D. L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval ACM, pp. 325 –334, 2011.

[23]J. D. Zhang, C. Y. Chow, and Y. Li, "Lore: exploiting sequential influence for location recommendations," in Proceedings of the 22nd ACMSIGSPATIAL International Conference on Advances in Geographic Information Systems ACM, pp. 103–112, Dallas Texas, USA, 2014.

[24] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 5, pp. 1167–1179, 2015.

[25] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in Proceedings of the 36th international ACMSIGIR conference on Research and development in information retrieval ACM, pp. 363–372, Dublin, Ireland, 2013

[26] T. He, H. Yin, Z. Chen, X. Zhou, S. Sadiq, and B. Luo, "A spatial-temporal topic model for the semantic annotation of POIs in LBSNs," ACM transactions on Intelligent Systems and Technology, vol. 8, no. 1, pp. 1–24, 2016.

[27] S. Oppokhonov, S. Park, and I. K. Ampomah, "Current location-based next POI recommendation," in Proceedings of the International Conference on Web Intelligence ACM, pp. 831–836, Leipzig, Germany, 2017.

[28] C. Cheng, H. Yang, I. King, M.R. Lyu, Fused matrix factorization with geographical and social influence in location-based social networks, in Proceedings of the 26th AAAI Conference on Artificial Intelligence, 2012, pp. 17–23.

[29] H. Li, Y. Ge, R. Hong, H. Zhu, Point-of-Interest recommendations: learning potential check-ins from friends, in Proceedings of the ACM SIGKDD International Conference, 2016, pp. 975–9

[30]R. Li, Y. Shen, and Y. Zhu, "Next point-of-interest recommendation with temporal and multi-level context attention," in 2018 IEEE International Conference on Data Mining (ICDM), pp. 1110 –1115, Singapore, 2018.

[31]Wang, W., Chen, J., Wang, J., Chen, J., Liu, J., & Gong, Z. (2020). Trust-Enhanced Collaborative Filtering for Personalized Point of Interests Recommendation. *IEEE Transactions on Industrial Informatics*, *16*(9), 6124–6132. https://doi.org/10.1109/TII.2019.2958696.

[32]Gan, W., Lin, J. C.-W., Fournier-Viger, P., Chao, H.-C., & Yu, P. S. (2018). *A Survey of Parallel Sequential Pattern Mining*. https://doi.org/10.1145/3314107.

[33]Yin, M., Liu, Y., Zhou, X., & Sun, G. (2021). A tensor decomposition-based collaborative filtering algorithm for time-aware POI recommendation in LBSN. *Multimedia Tools and Applications*, *80*(30), 36215–36235. https://doi.org/10.1007/s11042-021-11407-9.

[34]Francisco Dantas Nobre Neto, C. de S. B., and C. E. C. C., IEEE ITSS, & Institute of Electrical and Electronics Engineers. (n.d.). *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) : 1-4 Nov. 2016.*

[35]Mourchid, F., ben Othman, J., Kobbane, A., Sabir, E., & el Koutbi, M. (2016). A Markov chain model for integrating context in recommender systems. *2016 IEEE Global Communications Conference, GLOBECOM 2016 - Proceedings*. https://doi.org/10.1109/GLOCOM.2016.7841514.

[36]Kaur, R. (2022). Hidden Markov Model for short-term churn forecast in the structured overlay networks. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-021-11831-x.

[37]Zhong, T., Zhang, S., Zhou, F., Zhang, K., Trajcevski, G., & Wu, J. (2020). Hybrid graph convolutional networks with multi-head attention for location recommendation. *World Wide Web*, *23*(6), 3125–3151. https://doi.org/10.1007/s11280-020-00824-9.

[38] Q. Lv and Y. Qiao, "Big Data-Driven Hidden Markov Model-Based Individual Mobility Prediction at Points of Interest," vol. 9545, no. c, pp. 1–12, 2016, DOI: 10.1109/TVT.2016.2611654.

[39]Li, R. (2021). Joint Modeling of User Behaviors Based on Variable-Order Additive Markov Chain for POI Recommendation. *Wireless Communications and Mobile Computing, 2021*. https://doi.org/10.1155/2021/4359369.

[40]Amirat, H., Lagraa, N., Fournier-Viger, P., & Ouinten, Y. (2020). NextRoute: a lossless model for accurate mobility prediction. *Journal of Ambient Intelligence and Humanized Computing*, *11*(7), 2661–2681. https://doi.org/10.1007/s12652-019-01327-w.

[41]Yang, C., Bai, L., Zhang, C., Yuan, Q., & Han, J. (2017). Bridging collaborative filtering and semi-supervised learning: A neural approach for POI recommendation. *Proceedings of the ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining, *Part F129685*, 1245–1254. https://doi.org/10.1145/3097983.3098094.

[42] F. Marchal, J. Hackney, and K. Axhausen, "Efficient map matching of large global positioning system data sets: Tests on speed-monitoring experiment in zürich," Transp. Res. Rec. J. Transp. Res. Board, vol. 1935, pp. 93–100, Jan

[43] SNAP: Network datasets: Gowalla (stanford.edu). visited in date(27/04/2022)

[44] Fournier-Viger P, Gomariz A, Gueniche T et al (2014) SPMF : a Java open-source pattern mining library. J Mach Learn Res 15:3569–3573