20 Years of Mobility Modeling & Prediction: Trends, Shortcomings & Perspectives

Vaibhav Kulkarni
Distributed Object Programming Lab
University of Lausanne, Switzerland
Vaibhav.Kulkarni@unil.ch

Benoît Garbinato
Distributed Object Programming Lab
University of Lausanne, Switzerland
Benoit.Garbinato@unil.ch

ABSTRACT

In this paper, we present the insights drawn from a comprehensive survey of human-mobility modeling research based on 1680 articles that can serve as a roadmap for research and practice in this area. Mobility modeling research has accelerated the advancement of several fields of studies such as urban planning, epidemic modeling, traffic engineering and contributed to the development of location-based services. However, while the application of mobility models in different domains has increased, the credibility of the research results has decreased. We highlight two significant shortfalls observed in our reviewed studies: (1) data-agnostic model selection resulting in a poor tradeoff between modeling accuracy vs. computational complexity, and (2) failure in identifying the source of empirical gains, due to adoption of erroneous validation methodologies. We also observe troubling trends with respect to the application of Markov models for modeling mobility, despite the questionable association between Markov processes and mobility dynamics. We offer the literature meta-data and the associated tools to the community in order to improve the reliability and credibility of human mobility modeling research.

CCS CONCEPTS

• **Information systems → Spatial-temporal systems**.

KEYWORDS

mobility modeling, literature review, validation methodology

ACM Reference Format:
Vaibhav Kulkarni and Benoît Garbinato. 2019. 20 Years of Mobility Modeling & Prediction: Trends, Shortcomings & Perspectives. In *27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19), November 5–8, 2019, Chicago, IL, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3347146.3359110

1 INTRODUCTION

Over the last two decades, we have seen a large number of studies on human mobility modeling and prediction by the Geographic Information Systems community. This testifies to the importance of mobility prediction in context-aware systems, where a user's future location is used to seamlessly trigger service execution. These systems span services such as ride sharing, traffic prediction, point of interest recommendation, resource/urban planning, and network optimization among others. To understand the current methodologies behind the selection of prediction models and performance validation strategies, we performed a systematic literature review spanning last two decades amounting to a total of 1680 articles.

Based on the our literature survey, mobility modeling an be defined as *the process of estimating the probability distribution over an individual's future movement by minimizing the negative log-likelihood over the currently known user trajectory.* Research in this domain can be classified in three distinct categories: (1) theoretical modeling of mobility dynamics, (2) quantifying the uncertainty in next-place prediction, and (3) leveraging stochastic optimization algorithms to model human-mobility and benchmark the next-place forecasting capability. This paper focuses on the third category, where the models used for next-place forecasting fall into three categories: (1) Markov model variants (2) data mining techniques, and (3) neural network architectures, of which we exclusively focus on the Markov model variants. Despite the large number of studies, it is not trivial either to objectively compare cross-model performance, nor to identify the source of the empirical gains provided by the proposed models. This difficulty stems from the fact that each modeling approach is implemented with distinct search heuristics which introduces a range of different inductive biases. In order to validate the model's predictive performance, several types of datasets are used in these works that either contain GPS trajectories of pedestrians, recurrent WiFi connections, Bluetooth records or social network check-ins. This results in delivering different performance depending upon the dataset attributes. We find that for a majority of instances, empirical gains predominantly stem from erroneous validation methodology and selection of datasets with opportune characteristics rather than modeling/architectural amendments. We present the survey methodology in Section 2, followed by the insights drawn from the survey in Section 3. We discuss the shortcomings in Section 4 and conclude the paper in Section 5.

2 SURVEY METHODOLOGY

The increasing number of research articles published on human mobility-modeling and prediction testifies to the growing interest from researchers and practitioners. Figure 1 shows the most significant prediction models and the datasets used for mobility prediction within a given span of years. The goal of the systematic literature review is to classify the state-of-the-art in the domain of human mobility modeling and answer the following questions:

SIGSPATIAL '19, November 5–8, 2019, Chicago, IL, USA

ACM ISBN 978-1-4503-6909-1/19/11…$15.00
https://doi.org/10.1145/3347146.3359110

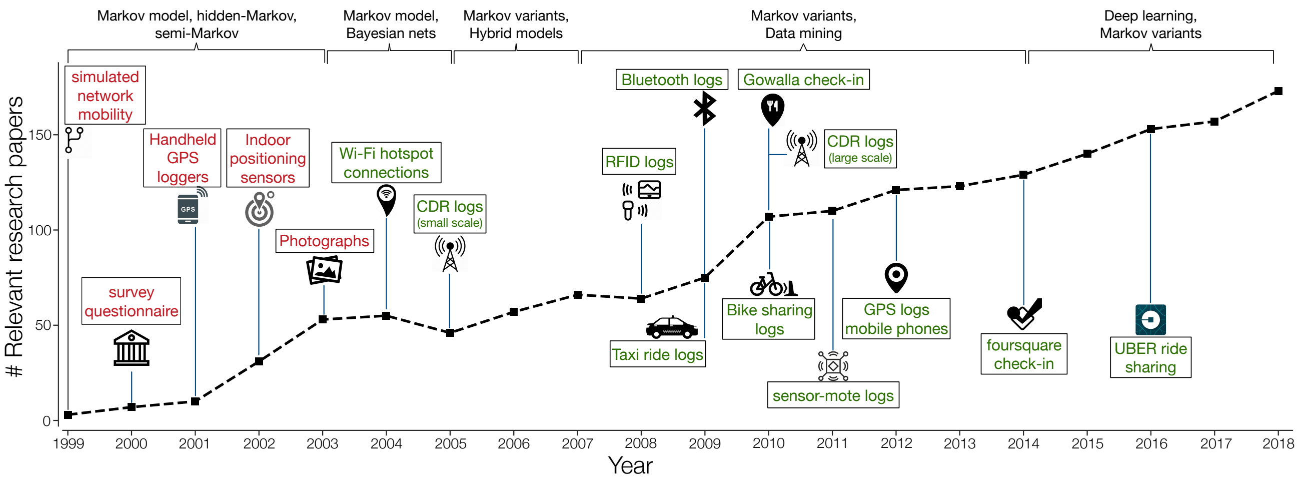

Figure 1: Mobility modeling and prediction: 20 years in review. The figure presents a summary of the review, highlighting the key techniques dominant (in terms of number of papers) in the respective era and the dataset (private datasets in red, public in green) driving this research.

(1) which techniques are used for constructing the next-place prediction models?, (2) which methodologies are adopted to validate the performance of the models?, and (3) which datasets are used to perform the performance quantification?.

2.1 Search Strategy & Article Selection

To find relevant studies, we selected three major academic search engines: Web of Science, Microsoft Academic Search and Semantic Scholar. Distinct platforms were chosen to ensure result completeness, as a single platform does not cover all major publisher venues. The search terms included, *human-mobility prediction, human-mobility modeling, next place forecasting, Predicting Significant locations* among others. The search terms were used for the title, abstract and keywords. These search domains consisted of *mobile and ubiquitous computing, geographic information systems and knowledge discovery and data mining*. Furthermore, only the studies that were peer reviewed, published in an international venue, written in English language were included. Additional papers were identified by using the citation and reference list of a given paper to minimize the risk of omitting relevant studies. We selected the top five cited articles in the domain of human mobility prediction [4, 7, 11, 12, 16] to bootstrap the snowballing procedure. At this stage, studies in the pool were ready for application of inclusion/exclusion criteria.

2.2 Preprocessing & Quality Assessment

To eliminate duplicates, the article titles were preprocessed by removing the stop words and normalization (stemming). String similarity was estimated by using Damerau-Levenshtein distance that measures the edit distance between two sequences. To calibrate the distance threshold, we manually assessed the top 20 top cited articles and labeled them according to their relevance. The prediction technique used in each paper was identified by first normalizing the abstract, followed by stop-words removal, tokenization, vectorization and using a count vectorizer to analyze relative frequency

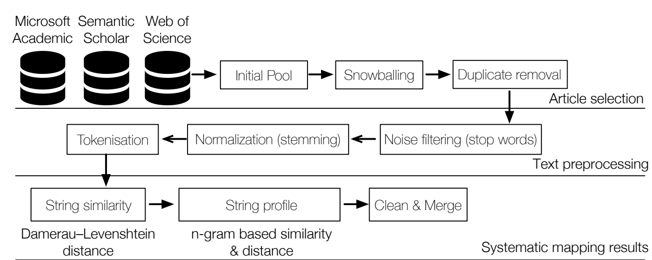

Figure 2: Overview of the search & data extraction strategy.

of the tri-grams. This is a widely used statistical approach that analyzes the position of word in the abstract, the term frequency and inverse document frequency. The preprocessing and quality assessment pipeline is summarized in Figure 2. This resulting list of articles was used to perform the analysis presented in the next section of the paper.

3 SURVEY FINDINGS

The first application of human mobility prediction was in the context of ad-hoc wireless networks by Gerla [8]. The knowledge about the user's next-location was used to anticipate topological changes and minimize the connectivity disruption caused by mobility. Application of human mobility prediction in wireless networks became prominent after Su et al. in [17], proposed a location-aware routing scheme and demonstrated its effectiveness using simulations. The seminal work however, in the context of prediction location on road networks using GPS information, was made by Ashbrook et al. [3]. They first identified significant places (points of interest) by clustering the raw GPS trajectories and then built a Markov-based predictor to forecast the next significant place. In their next article,

Approach	Dataset (#participants, duration, type, location)	Validation methodology
Markov model [3]	1 participant, 4 months, GPS traces	Not specified
semi-Markov model [6]	10 participants, unknown duration , GPS, GSM, WiFi	Not specified
hidden Markov model [11]	GeoLife (182 participants, 5 years, GPS traces)	Leave one out validation
mixed Makov model [2]	Simulated data: 691 participants, 1hr;31 minutes, GPS	k-fold cross validation (k=10)
mobility Markov chain [7]	GeoLife, synthetic dataset, private dataset (6 researchers)	Holdout validation, 50% split train-test
extended mobility Markov chain [1]	1 participant, 54 weeks, CDR	Not specified
variable-order Markov model [19]	GeoLife (182 participants, 5 years, GPS traces)	Holdout, Random selection, 90%-10% train-test
hidden semi-Markov model [21]	Simulated data	No empirical validation
hierarchical semi-Markov model [5]	GeoLife (182 participants, 5 years, GPS traces)	Holdout validation (no split % specified)
spherical hidden-Markov model [23]	Simulated data, Twitter dataset (geotagged tweets)	Random selection 70%-30% train-test split
Adaboost-Markov model [18]	GeoLife (182 participants, 5 years, GPS traces)	Random selection 90%-10% train-test split

Table 1: Markov-based mobility models, utilized datasets and cross-validation strategies.

Ashbrook et al. [4] extended this approach to perform mobility prediction across a dataset having multiple users. Several approaches along these lines were presented in the subsequent years using geolocation datasets consisting of data-points from different mobile phone sensors [13–15].

The reviewed studies can be categorized based on the techniques applied for mobility prediction as follows: (1) Markov model variants, (2) neural network techniques, and (3) data-mining based approaches. In this paper, we focus explicitly on the models derived using Markov process variants. Table 1 presents a meta-summary of the Markov model variants, the dataset used to quantify the model performance and the validation methodology used. These studies have used several datasets differing with respect to the data type, number of users, collection duration and geographic regions. We also highlight that in several cases the datasets are obtained privately from the telecommunication operators or generated synthetically using unspecified mobility simulators.

Accessibility of larger datasets prompted development and application of several variants of predictors based on Markov model such as hidden Markov model [11], mixed Markov model [2], semi-Markov model [6], hidden semi-Markov model [21], mobility Markov chain [7], extended mobility Markov chain [1], variable order Markov model [20], hierarchical hidden Markov model[5] and spherical hidden Markov model [5]. Each of the variant claims to address and account for different aspects of mobility trajectories, such as missing data from some time intervals [21], correlation with the stay duration [6], memory requirement [20], user behavioral characteristics [2], spatiotemporal associations with the path connecting the stay-points [5], semantic trace data [23] or location specific characteristics [11].

In addition to the usage of different datasets for performance validation, we observe that the applied validation methodologies also differ widely from one another or bootstrapped using differing parameters. We categorize the currently used cross-validation techniques for assessing model predictability as: (1) random shuffling before holdout validation, (2) train-test split ratio ranging from 90%-10% to 50%, (3) 3-fold to 30-fold cross-validation, and (4) weekday/month/year-based splitting. In the last case, model training is typically performed on the first four days of the week and tested on the remaining three. We also found several occurrences where the validation methodology was not specified. Selecting an appropriate validation approach to correctly assess the model performance on a given dataset is imperative to draw legitimate conclusions. To the best of our knowledge, we did not find any argumentation to select a particular dataset or a particular validation strategy in the reviewed literature. In the next Section, we delve deeper in these shortcomings as we highlight that selecting arbitrary validation technique is detrimental to the incremental process of model improvement and provides misleading measures.

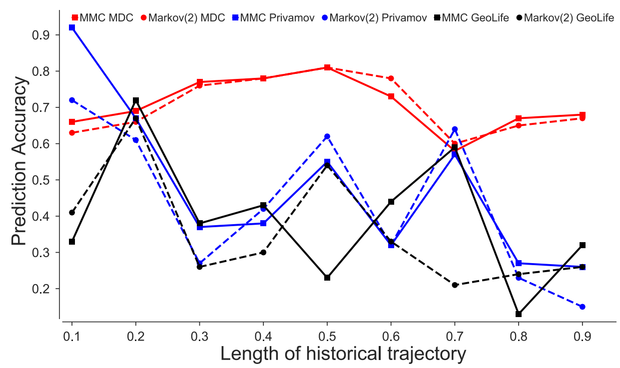

Figure 3: Prediction accuracy of Markov models variants on three datasets. The horizontal axis signifies proportion of trajectory length considered for the train-test split and vertical axis signifies the precision of the prediction model.

4 SHORTCOMINGS

In this section, we experimentally investigate the ill-effects resulting due to the shortcomings described in Section 3. We categorize these shortcomings in two domains: (1) inefficient accuracy vs. complexity trade-off arising from data-agnostic prediction model selection, (2) inconclusive model performance quantification due to adoption of inaccurate validation methodologies. We also expose the systematic bias involved in the model assessment due to the selection of the dataset and validation methodology devoid of any heuristics. Our experiments are based on PrivaMov dataset [9], Nokia mobile dataset (MDC) [10] and the GeoLife dataset [22].

Dataset	Holdout cross-validation			K-fold cross-validation		
	80-20	70-30	60-40	3-fold	5-fold	10-fold
MDC	0.78	0.81	0.66	0.63	0.72	0.65
PrivaMov	0.63	0.65	0.45	0.68	0.57	0.52
GeoLife	0.83	0.65	0.63	0.75	0.70	0.61

Table 2: Prediction accuracies derived by using different splits and values for validation.

4.1 Data-agnostic Model Selection

In order to analyze the generalizability of the model performance, we apply the same prediction model on mobility datasets differing widely with respect to some key characteristics. We also compare the accuracy of different prediction models on the same dataset using a fixed cross-validation methodology. Instead of presenting the average accuracy over the test-set, we adopt the canonical approach to perform sequential data cross-validation by rolling through the dataset.

In Figure 3, we compare the prediction accuracy of two approaches specified in Table 1: (1) mobility Markov chain (MMC), and (2) second-order Markov model (Markov(2)). We observe that the average accuracy and the variation trend of the Markov models differ by a large extent over the three datasets. Interestingly, the accuracy variation across the trajectory length is substantial for GeoLife and PrivaMov datasets as opposed to the MDC dataset. These accuracy variations stem from the fluctuation in dependencies between the POIs in a given dataset. From these experiments, we emphasize that the performance of the same prediction approach can differ widely across datasets. Moreover, a prediction model that performs poorly on one dataset can provide sufficiently favorable results on another. As a result, it is evident that conclusions regarding algorithmic performance cannot be justified without defining the dataset characteristics.

4.2 Flawed Validation Methodology

In order to highlight the misleading nature of the above validation approaches in the context of application to mobility modeling, we apply the holdout and k-fold cross-validation to access the performance (see Table 2). We observe that different train-test split ratios result in different prediction accuracies in case of holdout validation for all the three datasets under consideration. A similar behavior is observed in case of k-fold cross-validation for distinct values of k. Thus, it is clear that the accuracy results computed by these validation measures are neither a conclusive evidence of model performance, nor do they provide a comparative measure to analyze performance with respect to another model. Based on the above, we argue that application of flawed validation methodology is detrimental to the advancement of human-mobility modeling.

5 CONCLUSION

In this paper, we have highlighted the inconsistencies and pitfalls in human mobility modeling and prediction research through a large scale systematic review. Through this review, we have attempted to systematize knowledge and provide guidelines towards performing credible mobility modeling research. We have exposed the consequences of relying on data-agnostic model selection and adopting inaccurate validation methodologies through experiments on three real-world mobility datasets, while focusing on mobility models derived using Markov process variants. We offer these results, tools and the literature meta-data to the community with the hope of improving the credibility of mobility modeling research. [1]

Acknowledgement. This research work was partially supported by the Swiss National Science Foundation grant 157160.

REFERENCES

[1] Nurul Ain Amirrudin, Sharifah HS Ariffin, NNN Abd Malik, and N Effiyana Ghazali. 2013. User's mobility history-based mobility prediction in LTE femtocells network. In *2013 IEEE International RF and Microwave Conference (RFM)*. IEEE, 105–110.
[2] Akinori Asahara, Kishiko Maruyama, Akiko Sato, and Kouichi Seto. 2011. Pedestrian-movement prediction based on mixed Markov-chain model. In *GIS*.
[3] Daniel Ashbrook and Thad Starner. 2001. Learning Significant Locations and Predicting User Movement with GPS. In *SEMWEB*.
[4] Daniel Ashbrook and Thad Starner. 2003. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing* 7 (2003), 275–286.
[5] Mitra Baratchi, Nirvana Meratnia, Paul J. M. Havinga, Andrew K. Skidmore, and Bert A. G. Toxopeus. 2014. A hierarchical hidden semi-Markov model for modeling mobility data. In *UbiComp*.
[6] Yohan Chon, Hyojeong Shin, Elmurod Talipov, and Hojung Cha. 2012. Evaluating mobility models for temporal prediction with high-granularity mobility data. In *2012 IEEE International Conference on Pervasive Computing and Communications*. IEEE, 206–212.
[7] Sébastien Gambs, Marc-Olivier Killijian, and Miguel Núñez del Prado Cortez. 2012. Next place prediction using mobility markov chains. In *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*. ACM, 3.
[8] Mario Gerla. 1999. IPv6 flow handoff in ad hoc wireless networks using mobility prediction.
[9] Inria. 2012. PrivaMOv Dataset. https://projet.liris.cnrs.fr/privamov/project/. [Online; accessed 26-July-2018].
[10] Juha K Laurila, Daniel Gatica-Perez, Imad Aad, Olivier Bornet, Trinh-Minh-Tri Do, Olivier Dousse, Julien Eberle, Markus Miettinen, et al. 2012. The mobile data challenge: Big data for mobile computing research. In *Pervasive Computing*.
[11] Wesley Mathew, Ruben Raposo, and Bruno Martins. 2012. Predicting future locations with hidden Markov models. In *UbiComp*.
[12] Anna Monreale, Fabio Pinelli, Roberto Trasarti, and Fosca Giannotti. 2009. WhereNext: a location predictor on trajectory pattern mining. In *KDD*.
[13] Pubudu N. Pathirana, Andrey V. Savkin, and Sanjay Jha. 2003. Mobility modelling and trajectory prediction for cellular networks with mobile base stations. In *MobiHoc*.
[14] Donald J. Patterson, Lin Liao, Dieter Fox, and Henry A. Kautz. 2003. Inferring High-Level Behavior from Low-Level Sensors. In *UbiComp*.
[15] Wee-Seng Soh and Hyong S. Kim. 2003. QoS provisioning in cellular networks based on mobility prediction techniques. *IEEE Communications Magazine* 41 (2003), 86–92.
[16] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. 2010. Limits of predictability in human mobility. *Science* 327 5968 (2010), 1018–21.
[17] William Su, Sung-Ju Lee, and Mario Gerla. 2001. Mobility prediction and routing in ad hoc wireless networks. *Int. Journal of Network Management* 11 (2001), 3–30.
[18] Hongjun Wang, Zhen Yang, and Yingchun Shi. 2019. Next Location Prediction Based on an Adaboost-Markov Model of Mobile Users. *Sensors* 19, 6 (2019), 1475.
[19] Xiao-Yong Yan, Xiao-Pu Han, Bing-Hong Wang, and Tao Zhou. 2013. Diversity of individual mobility patterns and emergence of aggregated scaling laws. *Scientific reports* 3 (2013), 2678.
[20] Jie Yang, Jian Xu, Ming Xu, Ning Zheng, and Yu Chen. 2014. Predicting next location using a variable order Markov model. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on GeoStreaming*. ACM, 37–42.
[21] Shunzheng Yu and Hisashi Kobayashi. 2003. A hidden semi-Markov model with missing data and multiple observation sequences for mobility tracking. *Signal Processing* 83 (2003), 235–250.
[22] Yu Zheng, Xing Xie, and Wei-Ying Ma. 2010. Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data Eng. Bull.* 33, 2 (2010), 32–39.
[23] Wanzheng Zhu, Chao Zhang, Shuochao Yao, Xiaobin Gao, and Jiawei Han. 2018. A Spherical Hidden Markov Model for Semantics-Rich Human Mobility Modeling. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

[1] Repository Link: https://github.com/vaibhav90/Mobility-Prediction-Literature