International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

STA-STF based Recurrent Neural Network for Next Location Prediction

Niharika Dinesh[1], Atul Pandey[2], Shrey[3], Shashikiran kummari[4], Priyanka Jadhav[5]

[1]Software Quality Engineer 1, Dell Technologies, Bangalore, India
[2]Btech Graduate, Electrical & Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
[3]Systems Engineer Data Analyst, Tata consultancy services, New Delhi, India
[4]Btech Graduate, Electrical & Electronics Engineering, G.pullaiah College of Engineering and Technology Kurnool, Andhra Pradesh
[5] Deputy Manager, Data Analytics and PMO, Cipla Ltd, Mumbai, India

Abstract

In the evolving landscape of technology and smart devices, the focus on modeling spatial correlations, temporal dynamics, and friendship influence in point-of-interest (POI) check-ins has intensified. Existing works center on capturing user check-in behavior, emphasizing spatial and temporal dependencies of POIs. Markov chain-based methods address instance-level interactions, while recurrent neural network (RNN) approaches excel in handling variable-length check-in sequences. However, the former struggles with high-order POI transition dependency, and the latter cannot discern individual POI contributions in a historical check-in sequence. Additionally, RNNs propagate local and global information through a single bottleneck—hidden states. To address these limitations, a novel model is presented, enforcing contextual constraints on sequential data. The design incorporates spatial and temporal attention mechanisms over an RNN, highlighting the significance of POIs visited by users within specific time intervals and geographical distances during successive check-ins. The attention mechanism aids in identifying crucial POIs based on time difference and spatial distance in user check-in history for predicting the next POI. Periodicity and friendship influence are also considered in the model design. Experimental results on the BrightKite location-based social network demonstrate the proposed method's outperformance over existing state-of-the-art deep neural network methods for the next POI prediction and understanding of user transition behavior. Sensitivity analyses of parameters, including context windows for capturing sequential effects and estimating temporal and spatial attention, further validate the model's effectiveness. In parallel, this paper introduces a Space-Time Features-based Recurrent Neural Network (STF-RNN) for predicting individuals' next movements using mobility patterns from GPS device logs. Unlike traditional approaches, the STF-RNN automatically extracts internal representations of space and time features, enhancing efficiency in uncovering valuable insights into human behavior. Leveraging the sequence-representing ability of RNN structures, the model keeps track of user movement history, enabling the discovery of more meaningful dependencies and, consequently, improving performance. Collectively, these contributions advance the understanding and prediction capabilities in the realms of spatial dynamics, temporal dependencies, and human mobility patterns.

Keywords: Space-Time Features, Spatio Temporal Modeling, Next Location Prediction, User Transition Behavior

1. INTRODUCTION

In the era of ubiquitous smart devices, individuals commonly share their point of interest (POI) on social networks through check-in activities. The temporal and spatial contexts inherent in users' check-in behaviors present a valuable resource for constructing personalized recommendation systems. These systems leverage sequential check-in histories to assist users in navigating to POIs at subsequent time points. The modeling of user transition preferences holds significance not only in the realm of personalized POI prediction but also extends its utility to resource allocation, budgeting, service enhancement, and transportation planning.

The recommendation of Point of Interest (POI) check-ins is influenced by various factors, encompassing spatial, temporal, social, and chronological dimensions [1]. Users typically exhibit a preference for check-ins at proximate POIs, exemplified by the inclination to choose a closer grocery store (d1) over a more distant one (d2) when faced with alternatives. This introduces a spatial constraint in the modeling of check-in sequences. Similarly, check-ins occurring in close temporal proximity exert a notable influence on subsequent POI choices, such as the tendency to visit a coffee shop immediately after lunch or a bar following dinner. Additionally, the social dimension plays a significant role, as the behavior of user u1 may be influenced by the places previously visited or recommended by a friend (u2). Figure 1(a) visually represents the inverse correlation between distance and check-in frequency, elucidating users' propensity to check in at nearby locations. This depiction is derived from plotting the density of distances between chronologically sorted consecutive check-ins for each user. To further elucidate the impact of social influence on a user's POI check-in behavior, we quantify, for each user, the fraction of check-ins at POIs shared with friends. The cumulative distribution function curve in Figure 1(b) exhibits a steep slope, providing empirical evidence of the substantial influence exerted by friends on user check-in behavior. This analysis contributes to a nuanced understanding of the multifaceted dynamics that shape user preferences in POI check-ins.

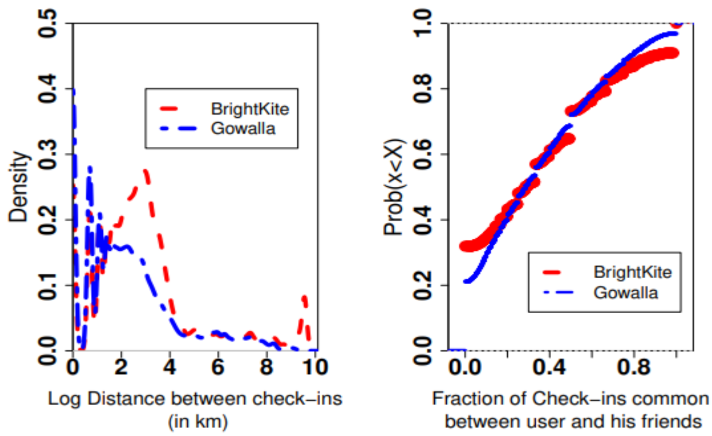


Fig. 1. Induce of distance and friendship on user check-in behavior

In contemporary research, the prediction of individuals' forthcoming locations has garnered considerable attention. Notably, the Markov Chain (MC) model, elucidated in [16] and [17], has been instrumental in inferring users' subsequent destinations by computing transition probabilities derived from mobility data logs. Moreover, the hidden Markov model has found application in predicting trip destinations, as exemplified in [18] and [19], where location characteristics or user activity transitions are considered as latent parameters. Another approach, articulated in [20]– [22], adopts a rule-based methodology to discover associations from movement transaction databases. In the domain of location prediction within cellular communication networks, Neural Networks (NN) have been extensively employed, aiming to mitigate traffic loads through the automatic updating of mobile user location information [23]– [27]. Notably, Recurrent Neural Networks (RNN) have emerged as a prominent method, showcasing successful applications in various domains, including sequential click prediction [28], word embedding [29], and time series prediction [30], [31]. Recent studies, exemplified by [32] and [33], leverage RNNs to model individuals' mobility patterns. Noteworthy findings from these investigations underscore the superior performance of RNNs compared to traditional NN models in the context of predicting people's future locations.

In this research endeavor, we advocate for the comprehensive consideration of both geographical and temporal dependencies within the context of predicting individuals' next locations. Our proposed approach entails the development of a novel ranking model, grounded in attention mechanisms applied over a recurrent neural network (RNN). To achieve this, a spatial attention layer is instantiated to systematically capture the geographical correlation between past and future Points of Interest (POIs), drawing insights from their spatial proximities. Simultaneously, temporal dynamics inherent in the existing POI sequences are addressed through the implementation of a temporal attention layer. This layer focuses on discerning the correlation and significance of POIs checked in short durations, complemented by the contextual dependency captured by the recurrent neural network. Furthermore, we incorporate the modeling of user friendships and POI co-occurrence by employing word2vec pre-trained embeddings. Specifically, these embeddings are applied to users and POIs on friendship edges and POI sequences, respectively. To account for periodicity, the time of visit for each POI is systematically integrated into our model. Finally, we advocate for the joint optimization of the next POI prediction task by using back-propagation through time (BPTT). The Bayesian pairwise ranking (BPR) algorithm is concurrently employed for effective ranking optimization, enhancing the overall predictive capabilities of the proposed model.

This study proposes the utilization of Recurrent Neural Networks (RNN) to model individuals' movement behavior, facilitating the prediction of their subsequent locations. Spatial and temporal dimensions are integrated into the network as features, with their internal representations autonomously learned by the network, eliminating the reliance on manually crafted representations. The spatial feature signifies specific locations visited by the user, while the temporal feature denotes the corresponding visiting times. Before constructing the prediction model, the previously gathered GPS points transform a sequence of interest points, encapsulating the series of locations visited by the user. During the training phase, trainable input features are forward-propagated into the hidden layer, concomitant with the incorporation of the previously accumulated hidden state. Empirical evaluations on a substantial real-life mobility dataset from the Geolife project demonstrate that the employed RNN structure enhances model effectiveness compared to contemporary methodologies such as Neural Networks (NN) and Markov-based approaches.

2. RELATED WORK

In contemporary research, recurrent neural networks (RNNs) have emerged as a focal point for modeling the sequence history and transitional patterns of user movements, achieving notable acclaim as a state-of-the-art methodology. In the context of predicting a user's next location at a specific time (t) based on their check-in history, the ST-RNN framework, as described in [2], integrates the consideration of time and spatial differences between consecutive check-ins. This is accomplished by partitioning both space and time into bins and subsequently learning a transition matrix for each temporal and spatial bin. However, a limitation of this approach lies in its consideration of only a fixed number of previous inputs, thereby neglecting the incorporation of long-range dependencies. Furthermore, the necessity to manually define the bounds of transition matrices renders the model impractical for deployment across diverse datasets. An alternative model, SERM [3], focuses on learning the dynamics of Points of Interest (POIs) through the embedding of different elements such as locations, users, time-bins, and textual information. Despite these advancements, SERM does not explore the spatial or temporal influences on user check-ins and inadvertently overlooks long-range dependencies by segmenting a user's check-in sequence into multiple segments, treating them as independent entities. This limitation hampers the comprehensive understanding of the intricate spatial and temporal dynamics influencing user movement behaviors.

Numerous models for Point of Interest (POI) recommendation have been constructed based on matrix factorization, concentrating on distinct facets such as geographical influence [4], temporal influence [5], and semantic influence [6]. However, a common limitation in these approaches is the oversight of sequential dependencies. Tensor Factorization (TF) has demonstrated success in time-aware recommendation and modeling spatial and temporal information [7]. TF incorporates both time bins and locations as additional dimensions in the factorized tensor. Nevertheless, this inclusion poses a challenge known as the cold start problem, particularly in predicting user behavior with new time bins. Within the realm of recommendation systems, ranking techniques play a pivotal role in predicting user preferences for various entities, including books, items, and locations. Rende et al. [8] introduced a Bayesian pairwise ranking (BPR) approach, wherein the relative preference of implicit items over non-observed items for each user is leveraged to enhance predictive accuracy. This approach contributes to the nuanced understanding of user preferences within the recommendation domain.

A recent contribution in the domain, as detailed in [34], introduces the utilization of Recurrent Neural Networks (RNN) in the form of a global prediction model termed Spatial Temporal Recurrent Neural Network (ST-RNN). This model is designed to predict the subsequent locations users are likely to visit. The efficacy of the ST-RNN model is evaluated using two distinct datasets, namely the Global Terrorism Database (GTD) and the Gowalla dataset. The recurrent structure embedded in ST-RNN is adept at capturing both local temporal contexts and periodic temporal patterns. To facilitate the modeling of temporal and spatial dependencies, the values are discretized into bins, allowing for the generation of time-specific and distance-specific transition matrices. Each specific temporal value within a time bin and each specific spatial value contribute to the calculation of the corresponding transition matrix. It is noteworthy that, in contrast to ST-RNN, our proposed model adopts a distinct approach wherein space and time features are directly fed into the network. The network, in turn, autonomously learns its internal representations, distinguishing itself from the discretization-based approach employed by ST-RNN.

The research presented in [27] proposes the development of local and global predictors utilizing Neural Networks (NN) to forecast an individual's subsequent movements. The evaluation of the neural predictor is conducted using the movement histories of four individuals within the research group at the University of Augsburg. The model employs a straightforward multi-layer perceptron with a single hidden layer, trained using the Backpropagation algorithm. A bit encoding mechanism is employed to represent both rooms and individuals. In the local predictor, each NN is individually trained with the movement patterns of a single person. Consequently, the input to the network comprises solely the codes corresponding to the last visited rooms. Conversely, in the global predictor, a single NN is trained using the collective movement data of all individuals. In this case, both the codes representing individuals and the codes for the last visited rooms serve as inputs to the network. Following a series of experiments, the optimal configuration for the NN is determined, featuring two and three neurons in the input and hidden layers, respectively. The evaluations conducted reveal that the local predictor outperforms the global predictor, achieving accuracy rates of 92.32% and 87.3%, respectively. This comparative analysis underscores the effectiveness of the local predictor in accurately forecasting an individual's next movements.

3. RESEARCH METHODOLOGY

In STF-RNN, the trajectory is represented as a sequence of tuple (xt, ht) where x is the centroid ID of the interest point visited at time t and h is the time unit part in hours of the leaving time from the interest point, $t = 1, 2, \ldots, n$, and n is the length of the trajectory. The task is to predict the future location of the mobile user at a specific time t based on historical mobility records. The architecture of the STF-RNN model is shown in Figure 1. It consists of four layers: input layer, lookup table layer, hidden layer (with recurrent connection), and output layer. The input layer consists of two vectors. The first one is $xt \in R^N$ which represents the centroid ID of the interest point at timestamp t. This vector is encoded using 1- of-N (or one-hot encoding) where N is the number of interest points. The second vector represents the time unit part in hours of leaving time from the interest point at timestamp t. We denote this vector by $ht \in R^M$ and it is encoded also using a 1-of-M encoding technique where M is the number of different time intervals. The time intervals represent the number of hours per day in which there are 24-time intervals (hour). In the one-hot vector representation, the interest points (or leaving times) are equidistant from each other without preserving any relationship among them. The lookup table layer maps the vectors of the centroid IDs and leaving times into real value vectors. The lookup table layer aims to learn a meaningful representation of the interest points and the leaving time's input features. This representation enables the model to capture the embedded semantic information about user behavior and consequently improve the prediction performance. Therefore, the trainable features will be used as input to further layers in the network rather than using one-hot vectors. More formally, let $X \in R^{N*d}$be the embedded matrix that represents a set of interest points, where d is the dimensionality of the embedded vector of the interest point.

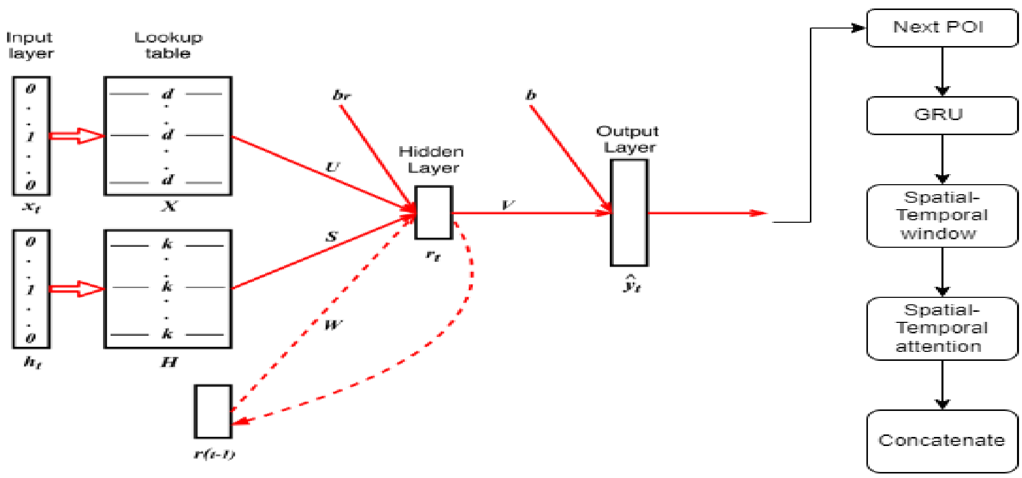


Fig. 2. STA-STF-RNN architecture.

Embedding:
This layer serves the purpose of mapping object identifiers to vectors of real numbers, establishing semantic similarity among entities, namely users, time bins, and Points of Interest (POIs), within a vector space. To initialize the user and POI embedding layer, we employ pre-trained embeddings obtained through word2vec [9], as introduced by Mikolov. Specifically, these embeddings are derived from the user's social links graph and the user's POI sequences, respectively. Within our model architecture, four distinct embedding layers are employed: one for user identification, one for the target POI identification, one for the contextual embedding of the sequence of POIs, and one for the time bins associated with the sequence of POIs. Subsequently, we concatenate the representations of time bins and POIs within the check-in sequence. This concatenation is performed to facilitate the learning of a compact representation encapsulating the user's check-in behavior.

Recurrent Neural Networks:
Gated Recurrent Units (GRUs) are used in this experiment which is a more robust variant of recurrent neural networks (RNN) and works better in capturing long-term dependencies. A GRU has two gates, a reset gate (r) to determine how to combine the current input xt and previous memory h_{t-1}, while an update gate (z) to define how much of the previous information to keep.

The Attention Mechanism RNN operates by compressing all visited Points of Interest (POIs) within a check-in sequence into a fixed-length vector, represented by the last hidden state. This vector serves as the determinant for predicting the next POI at each time step. In contrast, the attention mechanism is designed to comprehend the sequence of POI check-ins, selectively focusing on crucial and relevant POIs for the subsequent POI prediction at each time step. In this mechanism, all preceding hidden states are utilized as input to compute the probability distribution of previously checked-in POIs. Subsequently, a context vector is generated as a weighted sum of visited POIs, offering a nuanced representation for the prediction of the next POI. This approach enables the model to capture global information from the entire sequence, as opposed to relying solely on a single hidden state, thereby enhancing its capacity to infer and understand intricate patterns within the data.

In the proposed architecture as shown in Fig 2, two modes of attention mechanism are used namely *Temporal Attention* and *Spatial Attention*, respectively. In both modes, the attention component takes $t - 1$ hidden states of GRU as arguments $\{h1, h2, \dots, ht - 1\}$ and a context vector ht and context window τ in our case. It returns a vector z, which is the summary of all the previous hidden states focusing on the information linked to context ht and context window τ. More formally, it returns a weighted arithmetic mean of the hidden states of GRU $\{hi\}\ t - 1\ i = 1$ and the weights are chosen according to the relevance of each hi given the context ht.

Given a collection of user-POI check-ins Q, the objective is to represent users, time-bins and POIs in a new unified embedding space, where the new space captures the latent mobility patterns in the raw data. We jointly learn the embeddings for users and POIs using a Bayesian pairwise ranking (BPR) loss and backpropagation through time (BPTT) and formulate the training objective function as

$$J_t = \sum_{Q} \left(1 - \sigma(o_{u,t,q} - o_{u,t,q'})\right) + \lambda \|(\theta)\|^2$$

where q' is a negative location sample, λ is the regularization parameter and θ indicates all the parameters. As per the STF-RNN, the values of the hidden layer and the output layer are computed as below:

$$r_t = f(x_{et}U + h_{et}S + r_{t-1}W + b_r)$$

4. RESULTS & DISCUSSION

Comparison Methods:

To benchmark the efficacy of our proposed model, we conduct comparative evaluations with the following methodologies:

- Factorizing Personalized Markov Chains (FPMC) [10]. introduces a personalized transition matrix and factorizes the transition cube using tensor decomposition.
- Bayesian Pairwise Ranking (BPR) [11]. an implicit feedback method by using pairwise item preferences
- Gated Recurrent Unit (GRU) [12]. a robust variant of RNN and has an in-built memory along with gates to determine what information should be passed to the output.
- Spatio Temporal Recurrent Network (ST-RNN) [13]. models the user sequence of check-ins by considering the spatial and temporal contexts using transition matrices of spatial and time differences between successive check-ins.
- Semantics Enriched Recurrent Model (SERM) [14]. a variant of SERM, which only models' location, time, and user factors without using textual information.
- Attentional Recurrent Networks (DeepMove) [15]. learns the periodic contribution of user check-in history by attention mechanism, on sequence-level.

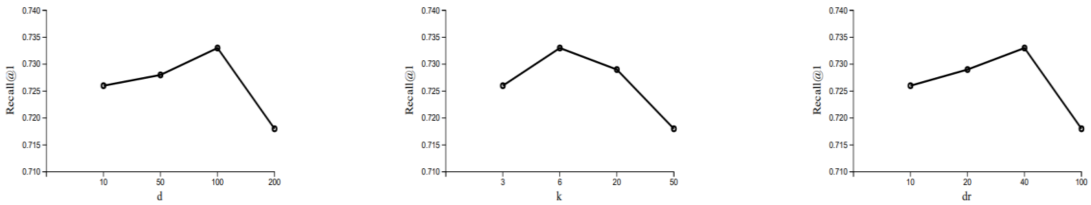
In Table 1, we show the results of models based on GRU networks with different attention mechanisms. GRU+SA considers only spatial attention while GRU+TA considers only temporal attention. GRU+STA combines both spatial and temporal attention. From the results, we can observe that GRU+STA outperforms the other three methods, which proves the effectiveness of the spatio-temporal attention mechanism. Besides, temporal attention shows better performance than spatial-level attention and basic GRU.

Table 1: Comparison of Different Attention Mechanisms using BK datasets

METHOD	AUC	MRR	recall
GRU	0.92	0.42	0.81
GRU+SA	0.95	0.45	0.86
GRU+TA	0.97	0.47	0.87
GRU+STA	0.98	0.49	0.89
STA+STF RNN	0.99	0.50	0.90

The table provides a comprehensive comparison of different attention mechanisms applied to the BK dataset, utilizing three key evaluation metrics: Area Under the Curve (AUC), Mean Reciprocal Rank (MRR), and recall. The baseline model, GRU (Gated Recurrent Unit), serves as a reference point, while subsequent enhancements include GRU with Spatial Attention (GRU+SA), GRU with Temporal Attention (GRU+TA), GRU with Spatio-Temporal Attention (GRU+STA), and the more sophisticated STA+STF RNN (Spatio-Temporal Attention + Space Time Features-based Recurrent Neural Network). Across the board, the metrics demonstrate a consistent performance improvement with the incorporation of attention mechanisms. Specifically, the STA+STF RNN model emerges as the top performer, achieving the highest AUC, MRR, and recall values. This underscores the effectiveness of integrating spatio-temporal attention and the STF-RNN architecture in capturing intricate spatial and temporal patterns, leading to superior predictive capabilities in the context of point-of-interest sequence modeling.

Within the STF-RNN framework, the parameters $d, k,$ and dr assume critical roles in determining the dimensionality of the embedded vectors corresponding to location, time, and hidden layers, respectively. These parameters significantly influence the model's efficiency. To comprehensively investigate the impact of these parameters and identify optimal settings, we conduct a series of experiments, as illustrated in Figure 3. The experimentation involves systematically varying the value of one parameter while keeping the others fixed, allowing for an examination of how alterations in each parameter influence the model's performance. This procedure is iteratively applied to all three parameters, providing a nuanced understanding of their individual contributions to the overall efficiency of the STF-RNN model.



(a) Impact of parameter d. (b) Impact of parameter k. (c) Impact of parameter dr.

Fig. 3. Parameters impact.

As depicted in the figures, the optimal performance for STF-RNN is achieved when the parameter k is set to 6, and the peak performance is observed under the parameter dr with a value of 40. This observation suggests that the model attains its highest accuracy with a relatively small k value, indicating that a limited amount of time information suffices in capturing the model dependencies. In contrast, the requirement for a larger dimensionality (dr) underlines the significance of more detailed location features. These results affirm the crucial role played by the parameters k and dr in the construction of an accurate location prediction model based on the Recurrent Neural Network (RNN). The findings provide insights into the optimal configurations of these parameters, shedding light on the nuanced balance needed to enhance the performance of the STF-RNN model.

TABLE 2: Performance of STF-RNN evaluated by Recall@N with varying window size.

Parameter	Recall@1	Recall@2	Recall@3
d	0.729	0.884	0.939
k	0.734	0.878	0.932
dr	0.687	0.854	0.897

Each row in the table represents the performance of the STF-RNN model for different recall values (Recall@1, Recall@2, and Recall@3) corresponding to specific window sizes. The values in each cell indicate the recall rate, which is the proportion of relevant items (correctly predicted next POIs) retrieved by the model among the top N predictions. The results provide insights into the model's effectiveness under different parameter configurations and help identify the optimal settings for achieving higher recall rates.

5. CONCLUSION

This study introduces the Spatio-Temporal Attention over Gated Recurrent Units (STA-GRU) for modeling Point of Interest (POI) sequences, presenting a novel neural network architecture leveraging self-attention mechanisms. This architecture is designed to effectively capture user transition behavior and improve the next POI prediction. Experimental results demonstrate the superior performance of STA-GRU compared to existing state-of-the-art models, particularly on the BrightKite benchmarks. Recognizing the challenges associated with long-range dependencies in recurrent neural networks, even with gating mechanisms, we further propose the integration of spatial and temporal attention over recurrent neural networks. Additionally, a Space-Time Features-based Recurrent Neural Network (STF-RNN) is proposed in this paper for predicting the future states of individuals' movements. A lookup table layer is employed to efficiently discover internal representations of space and time input features, enhancing the model's ability to capture embedded semantic information about user behavior. This STA-STF RNN proposed approach works better as compared to existing research. The recurrent structure is integrated with space and time interval sequences to uncover long-term dependencies, thereby increasing the efficiency of the proposed model. Performance evaluations conducted on a large real-life mobility dataset from the project substantiate that the STF-RNN model significantly improves prediction effectiveness compared to state-of-the-art models.

REFERENCES

1. Ramesh Baral, Dingding Wang, Tao Li, and Shu-Ching Chen. Geotecs: exploiting geographical, temporal, categorical, and social aspects for personalized poi recommendation. In Information Reuse and Integration (IRI), 2016 IEEE 17th International Conference on, pages 94–101. IEEE, 2016.
2. Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. Predicting the next location: A recurrent model with spatial and temporal contexts. In AAAI’16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pages 194–200, 2016.
3. Di Yao, Chao Zhang, Jianhui Huang, and Jingping Bi. Serm: A recurrent model for next location prediction in semantic trajectories. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pages 2411–2414. ACM, 2017.
4. Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun 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, pages 325–334. ACM, 2011.
5. Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. Exploring temporal effects for location recommendation on location-based social networks. In Proceedings of the 7th ACM conference on Recommender systems, pages 93–100. ACM, 2013.
6. Bo Hu and Martin Ester. Spatial topic modeling in online social media for location recommendation. In Proceedings of the 7th ACM conference on Recommender systems, pages 25–32. ACM, 2013.
7. Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R Lyu, and Irwin King. Stellar: Spatial-temporal latent ranking for successive point-of-interest recommendation. In AAAI’16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pages 315–322, 2016.

8. Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence, pages 452–461. AUAI Press, 2009.
9. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems, pages 3111–3119, 2013.
10. Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In Proceedings of the 19th international conference on World wide web, pages 811–820. ACM, 2010.
11. Jure Leskovec and Andrej Krevl. Snap datasets: Brightkite checkin data set. http://snap.stanford.edu/data, June 2011.
12. Kyunghyun Cho, Bart Van Merrienboer, Caglar Gulcehre, Dzmitry Bah- ¨ danau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078, 2014.
13. Jure Leskovec and Andrej Krevl. Snap datasets: Gowalla checkin data set. http://snap.stanford.edu/data, June 2011.
14. Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. Where you like to go next: Successive point-of-interest recommendation. In IJCAI '13 Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, volume 13, pages 2605–2611, 2013.
15. Jie Feng, Yong Li, Chao Zhang, Funing Sun, Fanchao Meng, Ang Guo, and Depeng Jin. Deepmove: Predicting human mobility with attentional recurrent networks. In Proceedings of the 2018 World Wide Web Conference on World Wide Web, pages 1459–1468. International World Wide Web Conferences Steering Committee, 2018.
16. Sebastien Gambs, Marc-Olivier Killijian, and Miguel N ´ u´nez Del Prado ˜ Cortez. Next place prediction using mobility Markov chains. Proceedings of the First Workshop on Measurement Privacy and Mobility MPM 2012, pages 1–6, 2012.
17. Akinori Asahara, Kishiko Maruyama, Akiko Sato, and Kouichi Seto. Pedestrian-movement prediction based on mixed Markov-chain model. GIS '11 Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 25–33, 2011.
18. Wenhao Huang, Man Lit, Weisong Hut, Guojie Song, and Kunqing Xie. Hierarchical Destination Prediction Based on GPS History. pages 972–977, 2013.
19. Wesley Mathew, Ruben Raposo, and Bruno Martins. Predicting future locations with hidden markov models. Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pages 911–918, 2012.
20. Conor Ryan and Kenneth N Brown. Occupant Location Prediction Using Association Rule Mining. pages 23–28, 2012.
21. L. Chamek M. Lalam S. Hamrioui M. Daoui, M. Belkadi and A. Berqia. Mobility Prediction and Location Management Based on Data mining. (December):1–4, 2013.
22. S. P. Kedia. An object tracking scheme for wireless sensor networks using data mining mechanism. 2012 IEEE Network Operations and Management Symposium, pages 526–529, 2012.
23. S. Parijaa, S. K. Nanda, P. K. Sahu, and S. S. Singh. Location Prediction of Mobility Management Using Soft Computing Techniques in Cellular Network. International Journal Computer Network and Information Security (IJCNIS), (May):27–33, 2013.

24. S. Parijaa, R. K. Ranjan, and P. K. Sahu. Location prediction of mobility management using neural network techniques in cellular network. 2013 International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT), pages 1–4, 2013.
25. S. Parijaa, P. K. Sahu, S. K. Nanda, and S. S. Singh. A functional Link Artificial Neural Network for Location Management in Cellular Network. International Conference on Information Communication and Embedded Systems (ICICES), page 11601164, 2013.
26. C. Leca, I. Nicolaescu, and C. Rˆıncu. Significant Location Detection & Prediction in Cellular Networks Using Artificial Neural Networks. Computer Science and Information Technology, 3(3):81–89, 2015.
27. J. Petzold L. Vintan, A. Gellert and T. Ungerer. Person movement prediction using neural networks. In First Workshop on Modeling and Retrieval of Context, Ulm, Germany, 114(4):1–12, 2004.
28. Y. Zhang, H. Dai, C. Xu, J. Feng, T. Wang, J. Bian, B. Wang, and T. Liu. Sequential Click Prediction for Sponsored Search with Recurrent Neural Networks. pages 1369–1375, 2014.
29. T. Mikolov, S. Kombrink, A. Deoras, L. Burget, and J. Cernocky. RNNLM - Recurrent Neural Network Language Modeling Toolkit. pages 196–201, 2011.
30. S. Yumlu, F. Gurgen, and N. Okay. A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange (ISE) prediction. Proc Pattern Recognition Letters, 26:2093–2103, 2005.
31. T. G. Barbounis and J. B. Theocharis. Locally recurrent neural networks for long-term wind speed and power prediction. Neurocomputing, 69:466–496, 2005.
32. F. Kamoun H. Kaaniche. Mobility Prediction in Wireless Ad Hoc Networks using Neural Networks. Journal of Telecommunications, 2(1):95–101, 2010.
33. Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts. pages 194–200, 2016.
34. Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A., et al. Theano: A python framework for fast computation of mathematical expressions. preprint arXiv:1605.02688, 2016.