Spatiotemporal Representation Learning for Driving Behavior Analysis: A Joint Perspective of Peer and Temporal Dependencies

*Submitted in fulfilment of the requirements for the degree of Bachelor of Technology*

*by*

Atishay Jain (167108)

Ankit Kumar (167107)

Prudhvi Malhotra (167129)

*Under the esteemed guidance of*

**Dr. R.B.V. SUBRAMAANYAM**



Department of Computer Science and Engineering

National Institute of Technology Warangal

2019-2020

# APPROVAL SHEET

The project work entitled **Spatiotemporal Representation Learning for Driving Behavior Analysis: A Joint Perspective of Peer and Temporal Dependencies** by Atishay Jain, Ankit Kumar, Prudhvi Malhotra is approved for the degree of Bachelor of Technology in Computer Science and Engineering at National Institute of Technology Warangal during the year 2019-20.

**Examiners**

——————————

——————————

——————————

——————————

——————————

**Supervisor**

——————————

Dr. R.B.V. SUBRAMAANYAM

Professor, CSE Dept.

Date: Place:

# DECLARATION

We declare that this written submission represents our ideas in our own words and where others’ ideas or words have been included. We have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misinterpreted or fabricated or falsified any ideas/ data / fact / source in our submission. We understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

—————————————————

(Signature)

Atishay Jain 167108

Date:

—————————————————

(Signature) Ankit Kumar 167107

Date:

—————————————————

(Signature)

Prudhvi Malhotra 167129

Date:

### ACKNOWLEDGEMENT

We would like to express heartfelt gratitude and regards to our project guide **Dr. R.B.V. SUBRAMAANYAM,** Department of Computer Science and Engineering, NIT Warangal. We convey a humble thanks to him for his valuable cooperation, support and suggestion throughout the project work which made this project successful. I shall remain indebted throughout my life for his noble help and guidance.

We are thankful to all the faculties of the Department of Computer Science and Engineering, NIT Warangal, for their encouraging words and valuable suggestions towards the project work. Last but not the least we want to acknowledge the contribution of our parents, family members, and friends for their constant and never-ending motivation.

We would like to take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of the project.

Atishay Jain (167108)

Ankit Kumar (167107)

Prudhvi Malhotra (167129)

#### ABSTRACT

As the world is growing at a great pace, so are humans' dependencies on electronic gadgets and automobiles. Transportation has been one such field which has revolutionized and changed the life of millions. The major changes from walking in past to multi-featured vehicles in present sets our path to autonomous driving in future. As every major change comes with its pros and cons, so to ensure driving remains a safe activity, assessing a driver's behavior in different circumstances and providing improvement measures becomes a necessity.

Our project aims to achieve significant knowledge from various drives conducted in a variety of situations to analyze the driver's behavior. We are using a Peer and Temporal Aware Representation Learning based framework (PATRL) for driving behavior analysis using GPS trajectory data. We ﬁrst detect the driving operations and states of each driver from their GPS traces. Then, we derive a sequence of multi-view driving state transition graphs from the driving state sequences, in order to characterize a driver’s driving behaviors that vary over time. In addition, we develop a peer and temporal-aware representation learning method to learn a sequence of time-varying yet relational vectorized representations from the driving state transition graphs..

### CONTENTS

1. Introduction 7
2. Related Work 8
3. Problem Statement 9
4. Attributes Definition 10
5. Model Framework 14
6. Experimental Results 17
7. Conclusion 23
8. References 24

**CHAPTER 1**

#### INTRODUCTION

**Driving Behavior Analysis**

As driving comprises of various speed and direction related operations, such as acceleration, deceleration, keeping constant speed, turning left, turning right, moving straight. By analyzing driving behaviors, we can try to provide driving assistance, enhance driving comforts, develop intelligent and resilient transportation systems, assessing driver's performances and enhance traffic safety. Our project aims to obtain quantitative representation of driver's behavioral trend from his raw GPS trajectories.

The studies conducted till date have scope of improvements as some of them are based on biased and expensive data sources, others considering data which might have privacy issues e.g. usage of CAN bus data, hence GPS data remains the best alternative. However CAN bus data is more accurate to quantify driving operations, but it receives data from equipment installed in vehicles and even consist of information which driver might not be willing to reveal. Due to the pervasiveness of GPS sensors, GPS data can be collected from location-based apps (e.g. Google Maps), that are granted permissions by user themselves. Thus, we are convinced to use GPS data for our task. It is highly promising to use high-resolution widely-available GPS trajectories with representation learning for the task of driving behavior analysis.

**CHAPTER 2**

#### RELATED WORK

Prior studies in driving behaviors analysis can be categorized into:

1. **Causal Analysis**: Researchers identify the causal factors of driving behaviors and explain how these factors inﬂuence road safety [20].
2. **Predictive Analysis**: Researchers mine the patterns from driving data and apply machine learning models (e.g., SVM, naive Bayesian, etc.) to predict risky scores [36].
3. **Descriptive Analysis**: Transportation experts deﬁne measurements (e.g., harsh or frequent acceleration/braking, sharp turn, acceleration before turn) based on transportation theory to describe driving behaviors [4].

## CHAPTER 3

#### PROBLEM STATEMENT

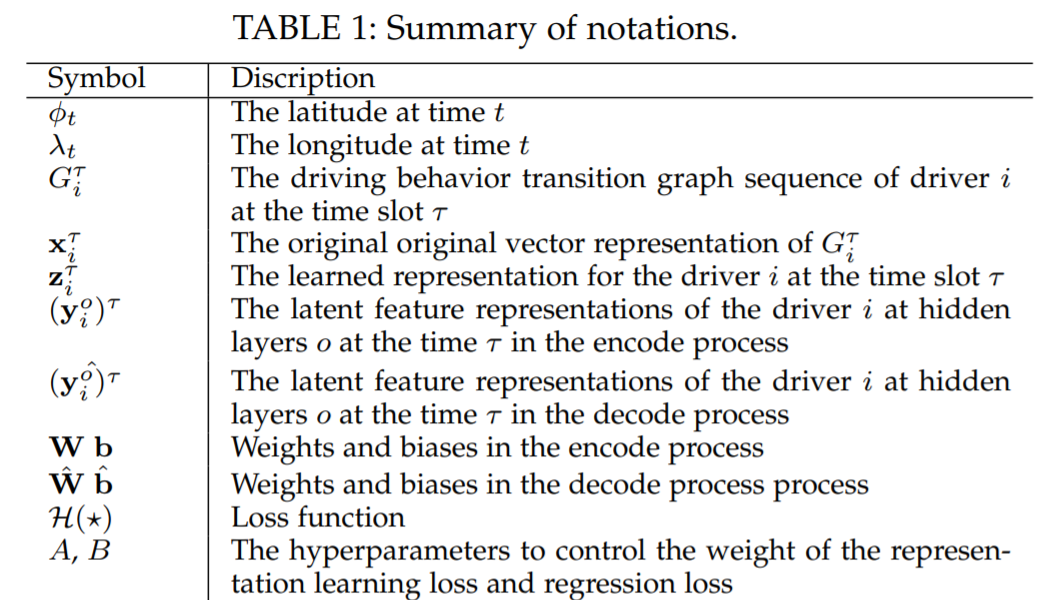
##### In our project, we explore the problem of automated driving behavior profiling with GPS traces. Formally, given a driver (a vehicle) and corresponding GPS trajectories, we aim to find a mapping function that takes the GPS trajectories as inputs, and outputs a sequence of time-varying yet relational vectorized representations in order to quantify the dynamics of the driver’s driving behavior. We formulate this problem as a task of spatio-temporal representation learning. Essentially, we first construct a sequence of driving state transition graphs from GPS trajectories, and then learn the latent representations of driving behavior from the graphs.

**CHAPTER 4**

**ATTRIBUTES DEFINITION**

1. *Driving State*: It is a tuple with two values, the first being speed status and second as direction status. E.g. <acceleration, turning-left>.

2.*State Transition Graphs*: Driver's movement from one driving state to another is depicted as state transition graph which inherently is a directed graph.



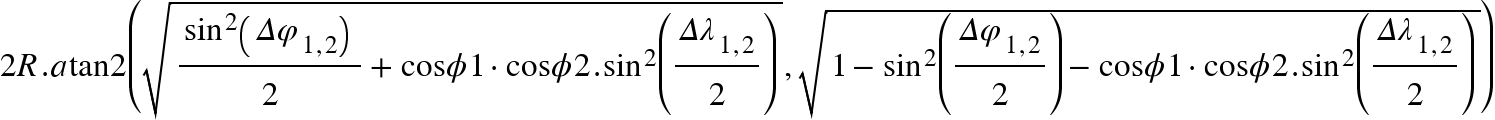
## CHAPTER 5

#### MODEL FRAMEWORK

To address the problem of consideration of both peer and temporal dependencies simultaneously, a framework called Peer and Temporal Aware Representation Learning Framework (PTARL) has been proposed and is being used in this project. Raw GPS trajectories are converted into state transition graphs and are used with aforementioned framework which can learn a sequence of time-varying yet relational vectorized representations.

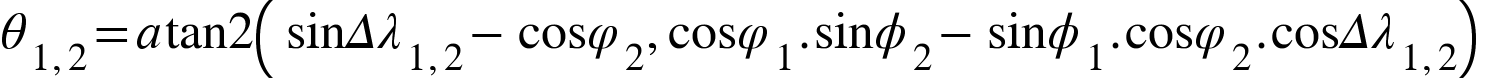
PTARL framework consists of following five essential steps:

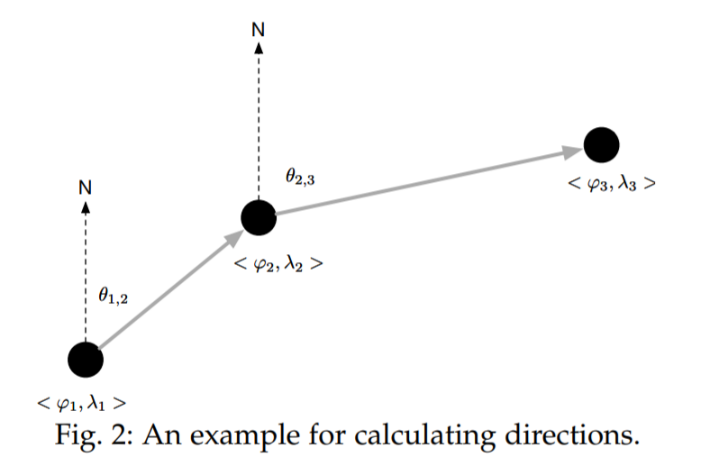
1. Construction of Multi-View Driving State Transition Graphs: To obtain driving state sequence and then driving state transition graph from GPS trajectories, we first detect different speed and direction related operations for each driver and then identify driving states and obtain driving state sequence. Later, the driving state sequence is segmented into small subsequences with a fixed time window, and then it is converted into a driving state transition graph. For each driving state subsequence, we get a driving state transition graph and hence it can be used to characterize the time-varying driving behavior of a driver.
   1. *Detection of speed-related operations*: Let ∆ϕ1,2 be the difference of ϕ1 and ϕ2, ∆ϕ2,3 be the difference of ϕ2 and ϕ3, ∆λ1,2 be the difference of λ1 and λ2, ∆λ2,3 be the difference of λ2 and λ3, and R be the radius of the earth. Then, the distance d1,2 between the two GPS points <ϕ1, λ1> and <ϕ2, λ2> is given by following equation:



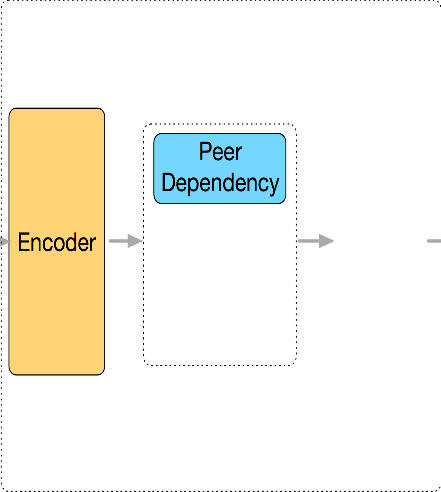
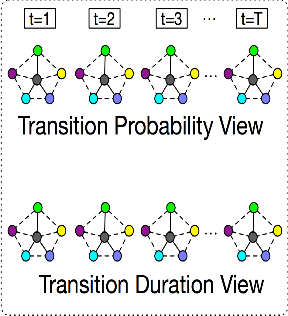
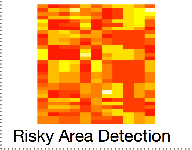
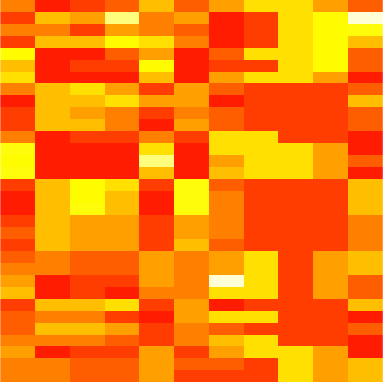
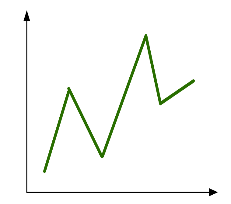
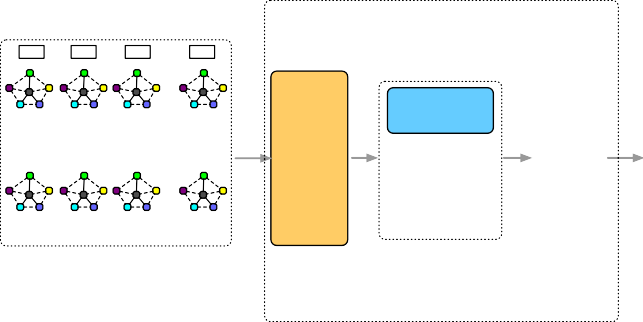
Similarly, the distance d2,3 between the two GPS points <ϕ2, λ2> and <ϕ3, λ3> can also be calculated. Then, given the time stamps of the three GPS points, denoted by t1, t2 and t3, the speed s2 at t2 is given by s2 = d1,2/(t2 − t1), the speed s3 at t3 is given by s3 = d2,3/(t3 − t2). For t3, if s3 > s2, the operation is detected as acceleration; if s3 < s2, the operation is deceleration; otherwise, the operation is “constant speed”. In practice, due to the noise caused by GPS devices, we introduce a loosing boundary €s into the calculation. For t3, if s3 > s2 and | s3 - s2| > €s, the operation is detected as acceleration; if s3 < s2 and |s3 −s2| > €s, the operation is deceleration; otherwise, the operation is “constant speed”.

* 1. *Detection of direction-related operations:* To detect the direction-related operations, we calculate the bearing θ1,2 between the two GPS points <ϕ1, λ1> and <ϕ2, λ2> by the following equation:





Similarly, we can obtain the bearing θ2,3 between the two GPS points <ϕ2, λ2> and <ϕ3, λ3>.Therefore, as shown in Figure2, at t3, if θ2,3 > θ1,2, then the operation is “turning right”; if θ2,3 < θ1,2, then the operation is “turning left”; otherwise, the operation is “moving straight”. In practice, we also introduce a loosing boundary €d to estimate directions. At t3, if θ2,3 > θ1,2 and |θ3 −θ2| > €d, then the operation is “turning right”; if θ2,3 < θ1,2 and |θ3 − θ2| > €d, then the operation is “turning left”; otherwise, the operation is “moving straight”.



Jointly Optimization

Decoder

……

Prediction and Historical Assessment of Driving Scores

Temporal Dependency

GPS Trajectory

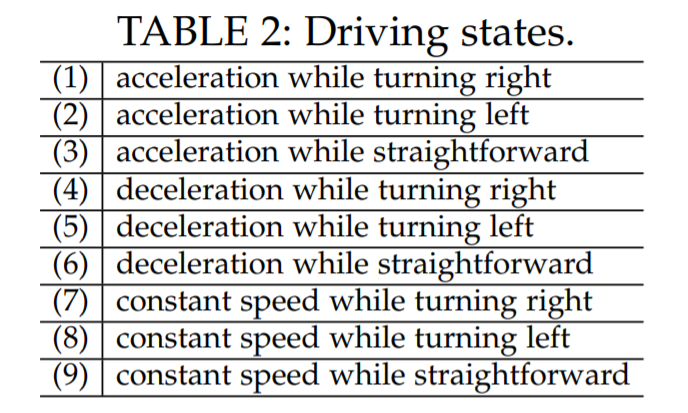
Representation Result (Vectors)

Driving State Transition Graph Sequence

Step-by-Step Optimization

An overview of the proposed analytic framework

1. Extraction of Driving State Sequences: Based on the two speed-related operations and the three direction-related operations, we can define the following driving states as per the table:



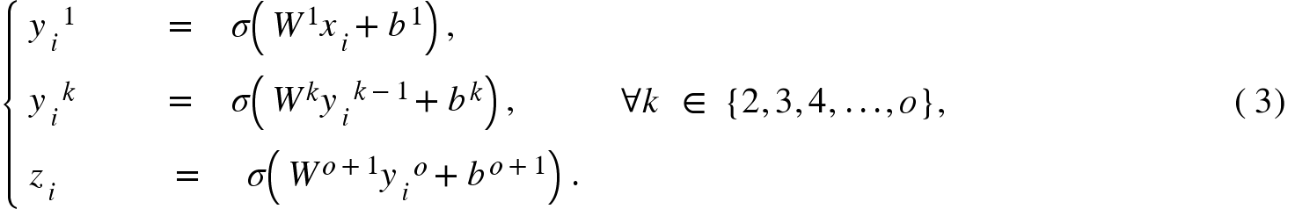
With the above definitions, we can identify the driving state of a driver at each time stamp. In other words, each trajectory is associated with a driving state sequence, which is denoted by {(ID, tn, Sn)} Nn=1, where ID is the identity of the driver, N is the size of the driving state sequence, tn is the nth time stamp, and Sn is the driving state at tn.

1. Peer and Temporal-Aware Representation Learning: There are peer and temporal dependencies among driving behavior. Therefore, in this project, we model the representation of driving behavior based on the following intuitions.

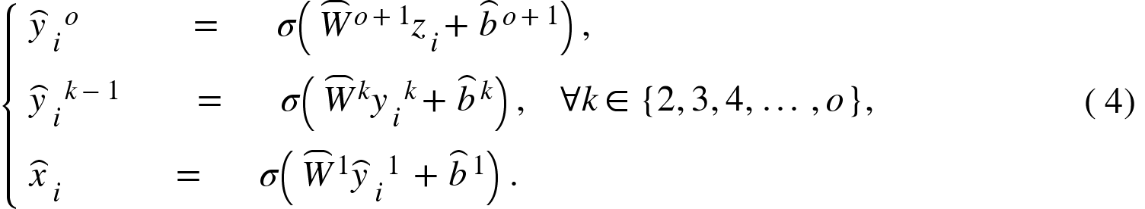
* Intuition 1: *Structural Reservation*: After reducing driving behavior into graphs, we need a representation learning based method to transform graphs into vectors in a latent feature space for automated quantification and profiling. Consequently, the method should be able to project graphs into lower-dimensional vectors while reserving corresponding characteristics and structures.
* Intuition 2: *Peer Dependency*: If two drivers exhibit similar driving habits, and the vehicle operation patterns of two corresponding trajectories are similar, then the driving state transition graphs of these two trajectories share a lot in terms of structures and characteristics. As a result, the learned representations of driving behavior should be close to each other. Consequently, the method should be able to model the graph-graph peer dependency in representation learning.
* Intuition 3: *Temporal Dependency*: The driving operations of the current time slot have autocorrelation with previous driving states. For example, if a driver decelerates while straightforward at t, and if ∆(t, t+1) is small enough, then he is likely to accelerate at t + 1. Consequently, the method should be able to model the current-past temporal dependency in representation learning.

1. Base Model: We utilize the deep Auto-Encoder model [3] as our base model. The motivation of using Auto-Encoder is that we aim to model the structural information of the driving state transition graph. Auto-Encoder shows the good performance of modeling structural information [1]. In addition, previous studies [18], [27] show that autoencoder is effective in modeling human mobility data, which fits the scenario in our work.

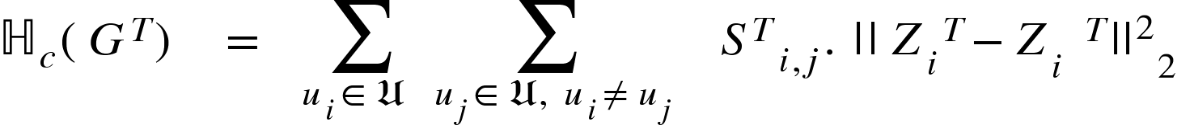
Auto-Encoder is an unsupervised neural network model, which projects the instances in original feature representations into a lower-dimensional feature space via a series of non-linear mappings. The Auto-Encoder model involves two steps: encode and decode. The encode part projects the original feature vector to the objective feature space, while the decode step recovers the latent feature representation to a reconstruction space. In the auto-encoder model, we need to ensure that the original feature representation of instances should be as similar to the reconstructed feature representation as possible. Formally, let xi be the original feature representation of the ith driver, and y 1 , y 2 , · · · , y o be the latent feature representations of the diver at hidden layers 1, 2, · · · , o in the encode step respectively, the encoding result in the objective lower-dimension feature space can be represented as zi ∈ R d with dimension d. Formally, the relationship between these vector variables is denoted by:



Meanwhile, in the decode step, the input will be the latent feature vector zi (i.e., the output of the encode step), and the final output will be the reconstructed vector xˆi . The latent feature vectors at each hidden layers can be represented as yˆ o i , yˆ o−1 i , · · · , yˆ 1 i . The relationship between these vector variables is denoted by:



where Ws and bs are the weight matrices and bias terms to be learned in the model. The objective of the auto-encoder model is to minimize the loss between the original feature vector x and the reconstructed feature vector xˆ. Formally, the loss function is



where ui denotes the i th driver and U denotes the driver set.

## CHAPTER 6

#### EXPERIMENTAL RESULTS

## CHAPTER 7

#### CONCLUSION

## CHAPTER 8

#### REFERENCES

[1] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. IEEEtransactionson pattern analysis and machine intelligence, 35(8):1798–1828, 2013.

[2] Yoshua Bengio et al. Learning deep architectures for ai. Foundations and trends Rin Machine Learning, 2(1):1–127, 2009.

[3] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy layer-wise training of deep networks. In Advances in neural information processing systems, pages 153–160, 2007.

[4] Adrian B Ellison, Michiel CJ Bliemer, and Stephen P Greaves. Evaluating changes in driver behaviour: a risk proﬁling approach. Accident Analysis & Prevention, 75:298–309, 2015.

[5] Yanjie Fu, Yong Ge, Yu Zheng, Zijun Yao, Yanchi Liu, Hui Xiong, and Jing Yuan. Sparse real estate ranking with online user reviews and ofﬂine moving behaviors. In Data Mining (ICDM), 2014 IEEE International Conference on, pages 120–129. IEEE, 2014.

[6] KlausGreff,RupeshKSrivastava,JanKoutn´ık,BasRSteunebrink, and J¨urgen Schmidhuber. Lstm: A search space odyssey. IEEE transactions on neural networks and learning systems, 2017.

[7] IsmailBulentGundogdu. Applyinglinearanalysismethodstogissupported procedures for preventing trafﬁc accidents: Case study of konya. Safety Science, 48(6):763–769, 2010.

[8] JungwookJun,JenniferOgle,andRandallGuensler. Relationships betweencrashinvolvementandtemporal-spatialdrivingbehavior activity patterns: use of data for vehicles with global positioning systems. Transportation Research Record: Journal of the Transportation Research Board, (2019):246–255, 2007.

[9] Tapas Kanungo, David M Mount, Nathan S Netanyahu, Christine D Piatko, Ruth Silverman, and Angela Y Wu. An efﬁcient kmeans clustering algorithm: Analysis and implementation. IEEE transactions on pattern analysis and machine intelligence, 24(7):881– 892, 2002.

[10] JiaweiZhangYanjieFuKunpengLiu,PengyangWangandSajalK. Das. Modeling the interaction coupling of multi-view spatiotemporal contexts for destination prediction. In Proceedings of the 2014 SIAM International Conference on Data Mining. SIAM, 2018.

[11] SteveLawrence,CLeeGiles,AhChungTsoi,andAndrewDBack. Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks, 8(1):98–113, 1997.

[12] Miao Lin, Hong Cao, Vincent W Zheng, Kevin Chen-Chuan Chang, and Shonali Krishnaswamy. Mobility proﬁling for user veriﬁcation with anonymized location data. In IJCAI, pages 960– 966, 2015.

[13] HaiLong Liu, Tadahiro Taniguchi, Yusuke Tanaka, Kazuhito Takenaka, and Takashi Bando. Visualization of driving behavior based on hidden feature extraction by using deep learning. IEEE Transactions on Intelligent Transportation Systems, 18(9):2477–2489, 2017.

[14] 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.

[15] Hideaki Misawa, Kazuhito Takenaka, Tomoya Sugihara, Hailong Liu, Tadahiro Taniguchi, and Takashi Bando. Prediction of driving behavior based on sequence to sequence model with parametric bias. In 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pages 1–6. IEEE, 2017.

[16] Mingdong Ou, Peng Cui, Jian Pei, Ziwei Zhang, and Wenwu Zhu. Asymmetric transitivity preserving graph embedding. In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1105–1114. ACM, 2016.

[17] Johannes Paefgen, Florian Michahelles, and Thorsten Staake. Gps trajectoryfeatureextractionfordriverriskproﬁling. InProceedings of the 2011 international workshop on Trajectory data mining and analysis, pages 53–56. ACM, 2011.

[18] Guannan Liu Yanjie Fu Charu Aggarwal Pengyang Wang, Jiawei Zhang. Ensemble-spotting: Ranking urban vibrancy via poi embedding with multi-view spatial graphs. In Proceedings of 2018 SIAM International Conference on Data Mining (SDM’18). SIAM, 2018.

[19] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 701–710. ACM, 2014.

[20] Sarah M Simmons, Anne Hicks, and Jeff K Caird. Safety-critical event risk associated with cell phone tasks as measured in naturalistic driving studies: A systematic review and meta-analysis. Accident Analysis & Prevention, 87:161–169, 2016.

[21] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. In WWW. ACM, 2015.

[22] Daixin Wang, Peng Cui, and Wenwu Zhu. Structural deep network embedding. In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1225–1234. ACM, 2016.

[23] Hongjian Wang and Zhenhui Li. Region representation learning via mobility ﬂow. In Proceedings of the 2017 ACM on Conference on InformationandKnowledgeManagement,pages237–246.ACM,2017.

[24] Pengfei Wang, Yanjie Fu, Guannan Liu, Wenqing Hu, and Charu Aggarwal. Human mobility synchronization and trip purpose detection with mixture of hawkes processes. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 495–503. ACM, 2017.

[25] PengfeiWang,JiafengGuo,YanyanLan,JunXu,andXueqiCheng. Multi-task representation learning for demographic prediction. In European Conference on Information Retrieval, pages 88–99. Springer, 2016.

[26] Pengfei Wang, Guannan Liu, Yanjie Fu, Yuanchun Zhou, and Jianhui Li. Spotting trip purposes from taxi trajectories: A general probabilistic model. ACM Transactions on Intelligent Systems and Technology (TIST), 9(3):29, 2018.

[27] Pengyang Wang, Yanjie Fu, Jiawei Zhang, Xiaolin Li, and Dan Lin. Learning urban community structures: A collective embedding perspective with periodic spatial-temporal mobility graphs. ACM Transactions on Intelligent Systems and Technology (TIST), 9(6):63, 2018.

[28] Yue Wang, Dawei Yin, Luo Jie, Pengyuan Wang, Makoto Yamada, Yi Chang, and Qiaozhu Mei. Beyond ranking: Optimizing wholepage presentation. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, pages 103–112. ACM, 2016.

[29] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. Driving with knowledge from the physical world. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 316–324. ACM, 2011.

[30] Jing Yuan, Yu Zheng, Chengyang Zhang, Wenlei Xie, Xing Xie, Guangzhong Sun, and Yan Huang. T-drive: driving directions based on taxi trajectories. In Proceedings of the 18th SIGSPATIAL Internationalconferenceonadvancesingeographicinformationsystems, pages 99–108. ACM, 2010.

[31] Nicholas Jing Yuan, Yu Zheng, Xing Xie, Yingzi Wang, Kai Zheng, and Hui Xiong. Discovering urban functional zones using latent activity trajectories. IEEE Transactions on Knowledge and Data Engineering, 27(3):712–725, 2015.

[32] Quan Yuan, Wei Zhang, Chao Zhang, Xinhe Geng, Gao Cong, and Jiawei Han. Pred: Periodic region detection for mobility modeling of social media users. In Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, pages 263–272. ACM, 2017.

[33] Chao Zhang, Liyuan Liu, Dongming Lei, Quan Yuan, Honglei Zhuang,TimHanratty,andJiaweiHan. Triovecevent:Embeddingbased online local event detection in geo-tagged tweet streams. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 595–604. ACM, 2017.

[34] Ruqing Zhang, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. Aggregating neural word embeddings for document representation. In European Conference on Information Retrieval, pages 303–315. Springer, 2018.

[35] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. Urban computing: concepts, methodologies, and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 5(3):38, 2014.

[36] Xiaoyu Zhu, Yifei Yuan, Xianbiao Hu, Yi-Chang Chiu, and YuLuen Ma. A bayesian network model for contextual versus noncontextual driving behavior assessment. Transportation Research Part C: Emerging Technologies, 81:172–187, 2017.