Deep understanding of big geospatial data for self-driving cars

Shuo Shang (UESTC), Jianbing Shen (IIAI), Ji Rong Wen (Renmin Univ.),

Panos Kalnis (KAUST)

Self-driving cars are capable of sensing environment and moving with little or no human input. Effective control of self-driving cars based on big geospatial data is one of the promising future directions of intelligent transportation. Specifically, big geospatial data understanding is helpful in acquiring travel behavior, vehicle mobility, traffic flow, nearby environment, and traffic-aware navigation. This special issue contains 10 research articles that present solid and novel research studies in the area of geospatial data analytics for self-driving applications, and 1 survey article that investigates existing studies related to self-driving cars. All of the 11 papers went through at least two rounds of rigorous reviews by the guest editors and invited reviewers.

Zhang et al. [1] target the problem of virtual machine (VM) live migration in self-driving systems. They aim to reallocate resources among VMs running different self-driving services for load balance. The VM live migration problem is modeled as a bin pack problem. A cluster-based genetic algorithm that outputs an approximation result of the bin pack problem is proposed.

Li et al. [2] present a framework named GeoTrafficPredict to support the accurate spatio-temporal prediction of web traffic flow. They develop a machine learning platform to learn the spatio-temporal pattern of traffic flow and use the pattern to predict the trend in both spatial and temporal dimension.

Wang et al. [3] propose an urban hot region detection method by using massive geo-tagged image data. They define the hot-region detection problem as finding the regions with the highest hot degrees. To solve the problem, they devise a spatial-density correlation measure to evaluate the similarity between two regions, and develop a parallel search framework to process the hot region detection problem efficiently.

Chen et al. [4] focus on the trajectory generation problem, which generates synthetic trajectory dataset that is indistinguishable from real trajectories. The trajectory generation problem fulfills the needs of large-scale trajectory data by self-driving simulation and traffic analysis tasks in data-sparse cities. They utilize Long Short-Term Memory (LSTM) to model the characteristics of trajectories, and apply Generative Adversarial Network (GAN) and Variational Autoencoder (VAE) frameworks to generate trajectories.

Cao et al. [5] study the problem of processing semantic trajectory space keyword query. To answer the query efficiently, they propose a hybrid index structure that is capable of effectively organizing social, spatial, and semantic information of trajectory data.

Han et al. [6] aim to address the classification and detection of malware in self-driving systems. In particular, they extract multi-features from malicious programs by combining pixel and n-gram features. In the process of feature selection, they apply L1-L2 penalty into the Logistic Regression and utilize LDA to reduce dimensions of malware features.

Wang et al. [7] propose a multi-view visual-semantic representation method for few-labeled visual recognition. Their method helps learn reliable classifiers with limited labeled image data, and thus improving the accuracy for self-driving cars in recognizing objects on the road.

Xu et al. [8] focus on recommending the users with their desired POIs in self-driving services. They adopt a conceptual inference-based method to deduce implicit intentions of user. Specifically, a locality-aware inference model is developed to generate concepts by considering typicality, granularity, spatial distribution, and the hypernym-hyponym relationships in knowledge graphs. They propose an interactive framework to return meaningful results by learning the importance of different ranking features through interactions.

Hu et al. [9] target the POI recommendation problem, which is a fundamental problem in self-driving services. They propose a POI-based item recommendation model via graph embedding. The model accurately captures the joint effect of geographical and temporal influences on both POI-level and item-level recommendation in a shared space.

Wang et al. [10] study the problem of potential workers estimation for a spatial crowdsourcing task. They define a novel Worker Collaborative Group Estimation (WCGE) problem and propose efficient algorithms to find the Nash Equilibrium of the game. Their solution is capable of estimating the potential workers for a new task.

Finally, Zhang et al. [11] present a comprehensive survey regarding managing, mining, and analyzing geo-social data in the context self-driving applications. Specifically, they investigate existing studies that target geo-social data index, location-based search, route planning, and recommendation mechanisms.

Those 11 articles represent diverse directions in the fast-growing area of geospatial data analytics in the community of self-driving cars. We hope that these papers will foster the development of self-driving related techniques and inspire more research in this promising area.

References

[1] Binbin Zhang, Xiao Wang, Hao Wang, Virtual Machine Placement Strategy for Self-driving using Clusterbased Genetic Algorithm, Neurocomputing.

[2] Jingjing Li, Jun Li, Nan Jia, Xunchun Li, Wenzhen Ma, Shanshan Shi, GeoTraPredict: A Machine Learning System of Web Spatio-Temporal Traffic Flow, Neurocomputing.

[3] Dahan Wang, Jianmin Li, Shunzhi Zhu, Detecting Urban Hot Regions by Using Massive Geo-Tagged Image Data, Neurocomputing.

[4] Xinyu Chen; Jiajie Xu; Rui Zhou; Wei Chen; Junhua Fang; Chengfei Liu, TrajVAE: A Variation AutoEncoder Model for Trajectory Generation, Neurocomputing.

[5] Keyan Cao, Qimeng Sun, Haoli Liu, Yefan Liu, Gongjie Meng, Jingjing Guo, Social Space Keyword Query Based on Semantic Trajectory, Neurocomputing.

[6] Xiangyu Han, Fusheng Jin, Runan Wang, Shuliang Wang, Ye Yuan, Classification of malware for self-driving systems, Neurocomputing.

[7] Da-Han Wang, Jianmin Li, Shunzhi Zhu, Few-Labeled Visual Recognition for Self-Driving Using Multi-View Visual-Semantic Representation, Neurocomputing.

[8] Jiajie Xu, Jiabao Sun, Rui Zhou. Chengfei Liu, Lihua Yin, CISK: An Interactive Framework for Conceptual Inference based Spatial Keyword Search, Neurocomputing.

[9] Xiaojiao Hu, Jiajie Xu, Weiqing Wang, Zhixu Li, An Liu, A Graph Embedding Based Model for Fine-Grained POI Recommendation, Neurocomputing.

[10] Zhi Wang, Yubing Li, Kun Zhao, Wei Shi, Liangliang Lin, Jizhong Zhao, Worker Collaborative Group Estimation in Spatial Crowdsourcing, Neurocomputing.

[11] Binbin Zhang, Peng Han, Hao Wang, Big Geo-social Data Driven Routing for Self-driving: A Survey, Neurocomputing.