115

Physics Based Data Preprocessing to Increase the Accuracy of Deep LearningBased Traffic State Estimation

Physics informed deep learning (PIDL) based approach, a hybrid approach that strengthens the data-driven only deep learning approaches with the wealth of physics knowledge, have been widely applied by the urban planners and policymakers in traffic state estimation (TSE) to perceive the congestion levels, understand traffic demand, and recognize gridlocks and bottlenecks of a road network. However, the accuracy of these approaches have been severely hampered by physically inconsistent data, which were collected by malfunctioned devices or in noisy noncooperative environment such as windy, foggy, or heavy rain weather. Currently, the existing TSE schemes using PIDL-based approaches mainly focus on physics-based regulation in model training. None of these studies have specifically examined the effectiveness of data preprocessing to advance the accuracy of TSE. In this research, we investigated the benefit of physics-based data preprocessing to increase the accuracy of TSE via computer simulations. The synthetic traffic density data generated from a physics model was modified in random locations and times by adding Gaussian or binary noise to create physics inconsistent data. These data were first screened with the Lighthill-Whitham- Richards (LWR) model of traffic physics and the fundamental Greenshields' diagram. Then only physics consistent data were used in the PIDL-based approach to train the deep learning model, which was then used to estimate the complete dataset. To evaluate the effectiveness of the physics-based data preprocessing, the estimation accuracy was calculated in terms of the Frobenius norm based normalized error measurement and compared with that of the approach which uses all training data in model building. Simulation results flaunt the superiority of the physics-based data preprocessing to increase the accuracy of traffic state estimation. When 10% of physics inconsistent data were in the training data, physics-based data preprocessing increased the TSE accuracy by 9% for pure deep learning approach and by 4% for physics informed deep learning approach.