An Online Learning Collaborative Method for Traffic Forecasting and Routing Optimization

Author(s):   Guo, Zhengang; Zhang, Yingfeng; Lv, Jingxiang; Liu, Yang; Liu, Ying

Title: An Online Learning Collaborative Method for Traffic Forecasting and Routing Optimization

Year: 2021

Version: Accepted version

Please cite the original version:

An Online Learning Collaborative Method for Traffic Forecasting and Routing Optimization

Zhengang Guo, Yingfeng Zhang*, *Member, IEEE,* Jingxiang Lv, Yang Liu*, Ying Liu

Abstract—**Recent advances in technologies such as the Internet of Things (IoT) and Cyber-Physical Systems (CPS) have provided promising opportunities to solve problems in urban traffic. With the help of IoT technologies, online data from road segments is captured by monitoring devices, while real-time data from vehicles is collected through preinstalled sensors. Based on these data, a CPS model is constructed to depict real-time status and dynamic behavior of road segments and vehicles. An online learning data-driven model is developed to extract prior knowledge and enhance collaboration between road segments and vehicles by combining short-term traffic forecasting and real-time routing optimization. A case study based on Xi'an city is proposed to demonstrate the feasibility and efficiency of the proposed method, showing a reduction in the travel time with reasonable computation time, without much compromising the fuel consumption and travel distance. This work potentially strengthens the transparency and intelligence of urban traffic.**

Index Terms—**Online learning, collaborative optimization, traffic forecasting, vehicle optimization, Cyber-Physical Systems.**

NOMENCLATURE

t	Time.
i, h	Subscript of road nodes.
j	Subscript of routes.
S	A finite set of road segments.
(i, h)	Road segment.
R	A finite set of feasible routes.
r_j	Feasible route.
FF	Free flow.
NT	Non-saturated traffic.
ST	Saturated traffic.
TC	Traffic congestion.
TV	Traffic volume.
AVS	Average vehicle speed.
RO	Road occupancy.
PT	Period of time.
DW	Day of the week.
P	Probability distribution function.
L	Likelihood function.
T	Time duration.
TD_{r_j}	Total travel distance of route r_j.
TT_{r_j}	Total Travel time of route r_j.
FC_{r_j}	Total Fuel consumption of route r_j.
$d_{(i,h)}$	Travel distance of road segment (i, h).
$\Delta t_{(i,h)}$	Expected time cost of road segment (i, h).
$\alpha_{(i,h)}$	Expected fuel consumption rate of road segment (i, h).
ω^{TD}	Travel distance weighting coefficient.
ω^{TT}	Travel time weighting coefficient.
ω^{FC}	Fuel consumption weighting coefficient.
$\beta_{(i,h)}$	Boolean variable, $\beta_{(i,h)}=0$, when road traffic events occur on road segment (i, h); otherwise, $\beta_{(i,h)}=1$.

I. INTRODUCTION

THE worldwide fast-growing car ownership has adverse consequences of traffic problems in urban areas, such as traffic congestion, vehicle emissions and exhaust, and waste of resources. For example, in 2014, traffic congestion caused an estimated 6.9 billion hours of travel delay and 3.1 billion gallons of extra fuel consumption with a total cost of $160 billion for urban areas of the United States [1]. These problems have contributed to migrating traditional transportation systems into next-generation variants. This area of research is referred to as intelligent transportation systems (ITS) [2]. In this regard, some of the advanced technologies such as the Internet of Things (IoT) [3], [4], cloud computing [5], [6], and Cyber-Physical Systems (CPS) [7], [8] can provide important concepts and technologies for this transition. By installing monitor devices and sensors on roads and vehicles, vast amounts of accurate and timely traffic information can be collected and processed, leading to the real-time status models for vehicles. This topic area is known as big data [9]. CPSs with integrated computational elements and physical processes can be used to manage such big data architecturally [10]. An integration architecture of vehicular CPS and mobile cloud computing has been designed to provide services for drivers to access mobile traffic information stored in the cloud [11]. Additionally, a large number of IoT-based models have been developed to simulate and optimize traffic systems. For example, a novel intelligent

Z. Guo, Y. Zhang, and J. Lv are with the Key Laboratory of Industrial Engineering and Intelligent Manufacturing, Ministry of Industry and Information Technology, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: guozg@mail.nwpu.edu.cn, zhangyf@nwpu.edu.cn, lvjx@nwpu.edu.cn).

Yang Liu is with the Department of Management and Engineering at Linköping University, SE-581 83 Linköping, Sweden and the Department of Production at University of Vaasa, 65200 Vaasa, Finland (e-mail: yang.liu@liu.se).

Ying Liu is with the School of Engineering at University of Glasgow, Glasgow G12 8QQ, UK (e-mail: Ying.Liu@glasgow.ac.uk).

traffic administration system based on IoT has been proposed to improve road traffic [12].

For traffic forecasting and routing problems, there are many existing models including linear models [13], [14] and nonlinear models [15], [16] for short-term traffic flow forecasting. Traffic forecasting supports proactive dynamic traffic control to ease traffic congestion and reduce travel time. Nevertheless, very little research in the literature combines both the forecasting and routing aspects for traffic problems, and at the same time showing that the solutions are available and effective in real-life environments [17], [18]. Despite the significant progress, existing approaches are insufficient to address the emerging challenges and problems that arise in future urban traffic systems. These challenges are summarized as follows.

(1) How to make effective used of IoT technologies to collect and aggregate online and real-time data from road segments and vehicles for enhanced transparency and intelligence of urban traffic systems?

(2) How to establish a CPS model with the integration of distributed physical entities and computational resources to depict real-time status and dynamic behavior of road segments and vehicles?

(3) How to construct an online learning data-driven model to extract prior knowledge from historical and online traffic data and strengthen collaborative relations between road segments and vehicles?

To solve these problems, an online learning collaborative method by combining short-term traffic forecasting and real-time routing optimization. Fig. 1 illustrates the design principle of the proposed method. By leveraging the IoT technologies, online traffic data from road segments and real-time data from vehicles are perceived by the monitoring devices and on-board units in an active manner. For example, online traffic data is captured from road segments by roadside monitoring devices such as automatic number plate recognition (ANPR) cameras, while real-time traffic data is collected from vehicles through preinstalled embedded devices which consists of sensors, processors, and communication modules. To store and analyze these data, a cloud computing platform is constructed, where semi-structured and unstructured data is processed by Hadoop clusters, while structured data is processed by data warehouses. Based on data analysis, an online learning data-driven model is developed to extract prior knowledge from the collected data, while the online learning mechanism includes model learning based on historical data and parameter learning based on online data. Besides, a self-adaptive collaborative mechanism is designed to combine short-term traffic forecasting and real-time routing optimization for traffic diversion and vehicle navigation. To evaluate the feasibility and efficiency of the proposed method, a case study based on Xi'an city is presented while three key performance indicators (KPI) are considered, namely the travel distance, travel time, and fuel consumption.

The remainder of this paper is organized as follows. Section II introduces the background and motivation. Section III outlines the overall system architecture of the proposed online learning collaborative method. Section IV proposes the online learning data-driven model and collaborative optimization mechanism. Section V demonstrates the proposed method with a case study based on Xi'an city. Section VI presents conclusion and future directions of this research.

II. BACKGROUND AND MOTIVATION

The state of the art of this topic is briefly reviewed, beginning with traffic forecasting and routing optimization approaches, and then continuing with vehicular networks such as vehicular ad-hoc networks (VANET), Internet of Vehicles (IoV), mobile cloud computing (MCC) and vehicular cloud computing (VCC). Finally, big data analytics tools such as Hadoop, Storm, Spark, and their applications in traffic are introduced and discussed. The main contributions of each research work are highlighted. Besides, research motivation of this research is also presented.

A. Traffic Forecasting and Routing Optimization Approaches

As an integral part of ITS applications, short-term traffic forecasting has received a wide range of research interests since the early 1980s [19]. Traffic forecasting models can be generally categorized into linear models and nonlinear models. For example, autoregressive integrated moving average (ARIMA) models are classic linear parametric models overwhelmingly used in time series analysis of traffic flow [20]. To improve the prediction accuracy, Kalman filters with discrete wavelet analysis have been investigated as a means of reducing local noise disruptions [21]. Besides, a local linear regression model was also used for short-term traffic prediction, showing better performance than the k-nearest neighbor and kernel smoothing models [22]. While nonlinear models include Bayesian network models [23], neural network models [24], deep learning models [25], support vector regression models [26], and particle swarm optimization models [27]. Apart from these models, research efforts have been conducted on hybrid models [28]. A hybrid modeling approach was proposed by combining artificial neural networks with a statistical approach to forecast urban traffic flow [29]. Nevertheless, existing traffic forecasting models are insufficient to deal with emerging challenges and problems in future urban traffic systems, due to limitations of integrating traffic data and vehicular data and collaboration between road segments and vehicles. While for vehicle routing approaches, geographic information systems (GIS) and optimization tools have been integrated with loading and distance requirements [30]. In urban areas, most of traffic navigation systems assist vehicles in route selection with real-time traffic information [31]. Besides, the research on time-dependent routing problems has been surveyed in Ref [32].
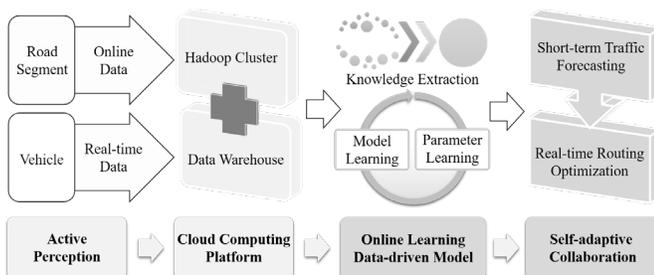

Fig. 1. Design principle of the online learning collaborative method.

However, these methods are passive or time-delayed strategies that cannot meet the need for real-time routing optimization in future urban traffic systems. For large-scale road networks with thousands or even millions of nodes, the classical Dijkstra's algorithm and its improved version A* algorithm are very slow. While the PHAST (hardware-accelerated shortest-path trees) algorithm is much faster than Dijkstra's algorithm. PHAST uses contraction hierarchies (CH) algorithm, which is a hierarchical speedup technique. PHAST can be further accelerated with multiple cores [33]. Besides, dynamic traffic assignment (DTA) models have been widely studied by researchers and practitioners [34], [35]. For example, a simulation-based DTA model was used to develop demand estimation methods in congested networks for lower-level problems [36]. An integrated framework for online and real-time advanced traveler information systems was also proposed by using DTA models [37]. However, most existing DTA models use the mean travel times as the route choice criteria but ignore their variations [38].

B. VANET, IoV, MCC, and VCC

Modern vehicles are equipped with a variety of sensors and communication devices [39]. These vehicles can collect traffic information from the environments as well as other vehicles and then feed it back to vehicles and traffic systems to assist in safe navigation, pollution control, and traffic management [40]. As wireless communication is provided among moving vehicles by using dedicated short-range communication, VANET allows vehicles to share different kinds of information such as safety or traffic jam information. However, service costs constrained communications for VANET is high due to the high mobility of the vehicles [41]. While IoV extends VANET's structure and applications and emphasizes information interaction among roadside infrastructures, vehicles, and humans [42]. In addition, MCC is another solution that integrates cloud computing into mobile environments to enable mobile users and mobile application providers to elastically use resources in a pay as you go manner [43]. Consequently, VCC is introduced to construct a cloud by using the collection of vehicles' computing resources and extending the capability of interactions amongst vehicles [44]. However, VCC is a local optimization method and causes a heavy load on the local Internet. Moreover, mobile devices are constrained by computing power, memory capacity and battery lifetime [45].

C. Hadoop, Storm, Spark, and Their Applications in Traffic

Current data processing tools on a single personal computer (PC) cannot meet the requirements of big data analytics. These tools are insufficient to capture, transfer, visualize, store, and analyze a huge number of real-time and multi-source data [46]. Several software frameworks have been used to develop big data applications, including Hadoop, Storm, and Spark. Hadoop is an open-source distributed computing platform for batch processing of large-scale data on commodity hardware. Hadoop and Hadoop Distributed File System (HDFS) by Apache have been extensively adopted in memorizing and controlling big data, while MapReduce is extensively applied for productivity exploration of big data [47]. In urban traffic systems, Hadoop

has been used to solve data-intensive problems, such as the processing of massive floating car data for traffic surveillance [48]. The storm is focused on real-time stream data processing for a broad variety of applications. For example, the Storm has achieved efficient data access and data cleaning of streaming mass spatial-temporal vehicle data [49]. Besides, a distributed stream processing model has been proposed with improved predictability by extending Storm [50]. To deal with computing workloads, many types of programming models have been developed to address the challenges. For instance, MapReduce can be used for batch processing, while Spark is designed for distributed data processing [51]. The programming model of Spark is similar to MapReduce but with an extension of resilient distributed datasets [52]. Hence, Spark can capture a wide range of processing workloads, including streaming, batch, and interactive workloads. For mobile big data analytics, Spark has been used to construct a scalable learning framework [53]. Also, data integration and sharing between heterogeneous traffic monitoring systems and vehicles are insufficient for current traffic applications, since many traffic monitoring systems are separate and independent for security purposes and the collected data is thus scattered among different systems. To develop applications that interact with open systems, real-time Java-centric architecture has been introduced as a homogeneous programming platform for industrial systems [54]. Therefore, a framework of traffic data integration and interaction is essential and needs to be further investigated, which helps in integrating traffic data and sharing data between traffic monitoring systems and vehicles.

Based on the above discussion, the motivation for this research is to address the collaborative problems of urban traffic systems and vehicles by combing short-term traffic forecasting and real-time routing optimization. For future traffic systems, online traffic data from road segments and real-time data from vehicles can be collected by monitoring devices and sensors in an active manner in an IoT-based environment. While a CPS model is constructed to integrate distributed physical resources and computational capabilities. In cyber space, real-time status and dynamic behavior of road segments and vehicles are depicted using the collected data. An online learning data-driven model including model learning and parameter learning is established by extracting prior knowledge from historical and online traffic data. A collaborative optimization mechanism is investigated and designed to enhance the collaboration of road segments and vehicles by combining short-term traffic forecasting and real-time routing optimization. To demonstrate the feasibility and efficiency of the proposed method, a case study based on Xi'an city is presented with the consideration of the travel distance, travel time, and fuel consumption.

III. OVERALL ARCHITECTURE OF THE PROPOSED METHOD

To make effective use of advanced technologies such as IoT and CPS, a three-layer system architecture of online learning collaborative optimization is designed as the theoretical basis and practical guide for potential applications in the future urban traffic. Specifically, the proposed architecture consists of the physical world, cyber space, and system layer. The plan-do-
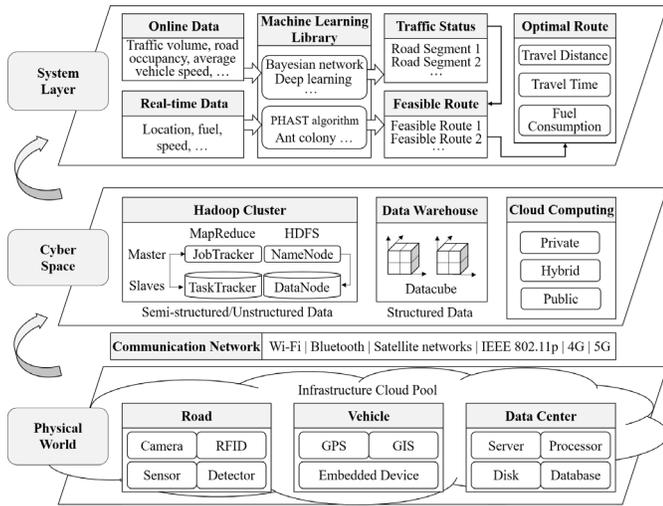
Fig. 2. Architecture of online learning collaborative optimization.

check-act (PDCA) cycle is used as the design cycle principle, which includes four steps, namely modeling, implementation, feedback, and improvement.

A. Physical world

The physical world consists of road segments, vehicles, and data centers. Physical infrastructure resources of urban traffic systems include road segments and roadside monitoring devices such as ANPR cameras, sensors, radio frequency identification (RFID) devices, and inductive-loop or microwave detectors. For vehicles, many assistant functional modules and on-board units are preinstalled such as the Global Positioning System (GPS) module, Geographic Information System (GIS) module, and other embedded devices. In addition, data centers including servers, processors, disks, and databases provide data services for data storage, processing, and analysis. However, most of the management and control systems for these physical resources are distributed and separate. To integrate these physical entities, a cloud computing platform is established by cloud computing, which can encapsulate capabilities of physical resources into cloud services and form an infrastructure cloud pool to provide these services in a plug-and-play manner. Fig. 3 illustrates the active perception of road segments and vehicles in an IoT-based environment. With the help of IoT technologies, online traffic data of road segments and real-time data of vehicles are actively collected by monitoring devices, preinstalled embedded devices, and sensors, which are transmitted through the communication networks such as Wi-Fi, Bluetooth, satellite networks, IEEE 802.11p, the fourth-generation (4G) and fifth-generation (5G) mobile communication networks. In Fig. 3, yellow dash lines denote the communication between vehicles.
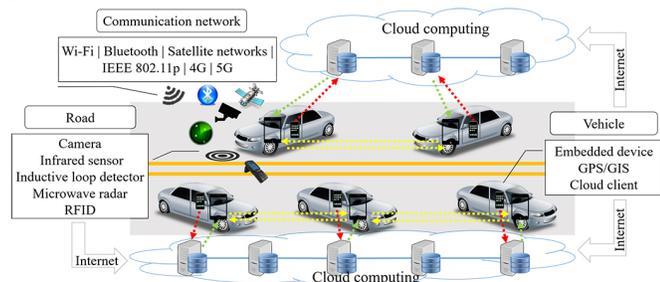

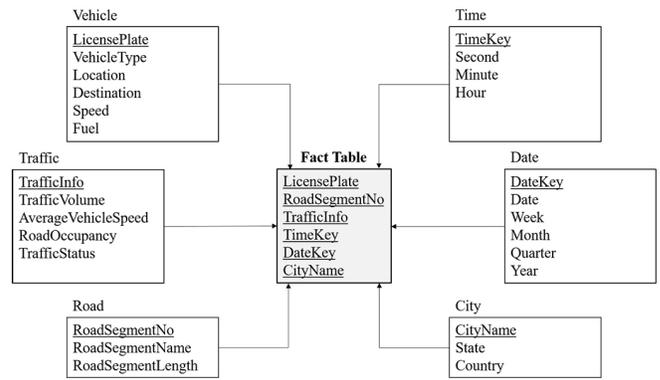Fig. 3. Active perception of road segments and vehicles in IoT-based setting.

Fig. 4. A star schema of multi-dimensional data model in data warehouses.

Red dash lines denote real-time status information transmission from vehicles to the cloud computing platform, while green dash lines denote decision-making information transmission from the cloud computing platform to vehicles.

B. Cyber space

The main function of the cyber space is to process and further analyze multi-source, multi-dimensional, and heterogeneous data. Semi-structured/unstructured data is stored and processed by Hadoop clusters, while structured data is managed by data warehouses. In data-intensive distributed applications, Hadoop is used to handle sensory and less structured data. Hadoop is an open-source distributed architecture system that can be used for structured and unstructured data search, data analysis, and data mining. It consists of two core parts, namely MapReduce and HDFS. MapReduce is a distributed and parallel framework proposed by Google and can be used to process large-scale data, which distributes data and parallelizes the computation across a cluster of computers. HDFS is used to store the input data for Map tasks and output data of the Reduce tasks. For structured data, it is stored as data cubes in data warehouses, where online analytical processing (OLAP) tools such as drill-up, drill-down, slice, dice, and pivot can be used to deal with multi-dimensional data. Fig. 4 illustrates a star schema of the multi-dimensional data model in data warehouses, which consists of a fact table and six-dimension tables. In the fact table, each tuple includes a pointer to each dimension, while the columns in each dimension table correspond to attributes of the dimension. While for cloud computing, a cloud computing platform is constructed to manage cloud resources and provide on-demand cloud services in a distributed environment. Specifically, three types of cloud services are provided with differentiated access permissions, namely public, private, and hybrid clouds.

C. System layer

The system layer includes a CPS model of road segments and vehicles as well as an online learning data-driven collaborative model of short-term traffic forecasting and real-time routing optimization. By integrating the physical infrastructures and computational resources, a CPS model is established to depict real-time status and dynamic behavior of road segments and vehicles. The main objective is to enable road segments and vehicles to be 'smart' for active perception, active response, and autonomous decision-making based on their real-time status and actively adjusting their behavior. With the help of the CPS model, an online learning data-driven model including model learning and parameter learning is developed to extract prior

knowledge from historical and online data for short-term traffic forecasting and real-time routing optimization. At first, the online learning data-driven model of each road segment for traffic status forecasting is developed by extracting knowledge from the collected online traffic data and historical traffic data, such as traffic volume, road occupancy, and average vehicle speed. The models and parameters are updated based on the extracted knowledge. A machine learning library including a broad variety of machine learning algorithms is constructed to support data preprocessing, data analytics, short-term traffic forecasting, and real-time routing optimization. Then, the collaboration relationship between road segments and vehicles is realized by combining short-term traffic forecasting and real-time routing optimization. According to the real-time status information of vehicles including current location, destination, and speed, a finite set of feasible routes is initialized. Finally, the optimal route is generated considering three KPIs, namely travel distance, travel time, and fuel consumption, while the result is transmitted to the cloud client on the vehicle side such as thin clients, mobile apps, and web browsers.

IV. ONLINE LEARNING DATA-DRIVEN MODEL AND COLLABORATIVE OPTIMIZATION MECHANISM

To integrate multi-source traffic data and vehicular data from separate and independent management subsystems, an online learning data-driven model is thus constructed to extract prior knowledge from online traffic data and real-time vehicular data. Specifically, two parts are considered, namely model learning and parameter learning. Besides, to enhance the collaboration relations between road segments and vehicles, a self-adaptive collaboration mechanism is proposed by combing short-term traffic forecasting and real-time routing optimization.

A. Model learning

The main objective of model learning to select proper traffic characteristics from a range of features and develop a short-term traffic forecasting model for each road segment from a library of traffic forecasting models. Generally, four types of traffic status are considered, namely free flow, non-saturated traffic, saturated traffic, and traffic congestion, while the characteristics of traffic data include the period of time, day of the week, percent time-spent-following, average vehicle speed, average headway, time occupancy, road occupancy, and traffic volume.

To sort out traffic characteristics with strong correlations to traffic status, backward elimination and schemata search are implemented to assess these potential variables and generate a finite subset of characteristics. The backward elimination begins with the complete set and successively eliminates one characteristic at a time [55]. At each stage, the possibility of deleting a characteristic is considered. The significance level is set at 0.15, which determines the inclusion and exclusion of characteristics from the subset. Schemata search is a method for quickly finding the subset of relevant characteristics, which aims to reduce the computation time of backward elimination especially if many characteristics are irrelevant. Consequently, traffic volume, average vehicle speed, road occupancy, the period of time, and day of the week are selected as the top five
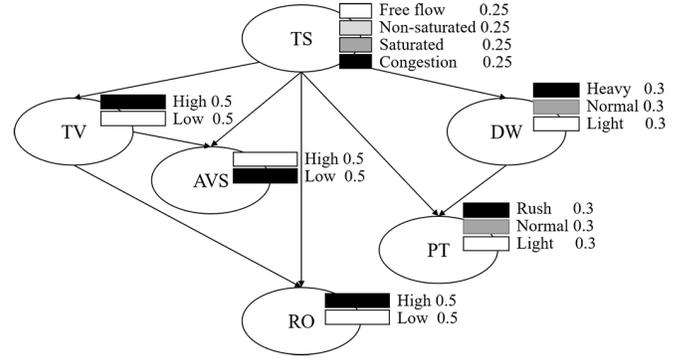

Fig. 5. Network structure of Bayesian networks for model learning.

characteristics.

Mathematically, the road network is defined as a graph $G = (N, S)$ with a finite nonempty set N of road nodes and a set S of edges representing road segments between nodes. Consider a road network with a finite non empty set of road nodes $N = \{1, 2, \dots, i, \dots, h, \dots n\}$ and a finite set of road segments $S = \{(1, a), \dots, (i, h), \dots, (n, c)\}$. Note that (i, h) denotes the road segment from road node i to node h. For a finite set of feasible routes $R = \{r_1, r_2, \dots, r_m\}$, each route r_j of R consists of a set of road segments from S, namely S_{r_j}. Thus, the matrix TI encapsulates multi-source and multi-dimensional traffic data, which can be used to formulate the online traffic status of road segments by an information model as follows.

$$TI = \begin{bmatrix} TV_{(1,a)} & AVS_{(1,a)} & RO_{(1,a)} & PT_{(1,a)} & DW_{(1,a)} \\ TV_{(2,b)} & AVS_{(2,b)} & RO_{(2,b)} & PT_{(2,b)} & DW_{(2,b)} \\ \dots & \dots & \dots & \dots & \dots \\ TV_{(n,c)} & AVS_{(n,c)} & RO_{(n,c)} & PT_{(n,c)} & DW_{(n,c)} \end{bmatrix}$$
(1)

where TI denotes traffic information of road networks. $TV_{(i,h)}$ denotes traffic volume of road segment (i, h). $AVS_{(i,h)}$ denotes average vehicle speed of road segment (i, h). $RO_{(i,h)}$ denotes road occupancy of road segment (i, h). $PT_{(i,h)}$ denotes the period of time of road segment (i, h). $DW_{(i,h)}$ denotes day of the week of road segment (i, h).

While Bayesian networks have proven to be a very powerful method for dealing with uncertainty and complexity, which can be formulated as a directed acyclic graph model that represents conditional independencies between a large set of variables [56]. Besides, the scale of the traffic forecasting model is adjusted by increasing or cutting down the number of vertices in Bayesian networks. Fig. 5 illustrates the network structure of Bayesian networks for model learning of each road segment, which consists of six vertices, namely traffic status, traffic volume, average vehicle speed, road occupancy, the period of time, and day of the week. A probability table of each vertex defines the traffic characteristics probability distribution. For example, the traffic status (TS) node consists of four status types, namely free flow, non-saturated traffic, saturated traffic, and traffic congestion. The criteria are calculated using cluster analysis and k-means algorithms from historical traffic data. The sum of probability values for each vertex is 1.

For numerical characteristics, such as traffic volume, average vehicle speed, and road occupancy, these characteristics satisfy the Gaussian distribution, the probability density function is

formulated as follows.

$$P_{(i,h)}(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{2}$$

where $P_{(i,h)}$ denotes the traffic status probability function of road segment (i,h), $x \in \{TV, AVS, RO\}$. μ denotes the mean value and σ denotes the standard deviation.

Otherwise, the characteristics such as the period of time and day of the week are transformed into nominal characteristics using entropy-based discretization. For nominal characteristics, traffic status probability of different nominal characteristics for road segment (i,h) is formulated using Bayes' rule as follows.

$$P_{(i,h)}(A|B) = \frac{P_{(i,h)}(A) \cdot P_{(i,h)}(B|A)}{P_{(i,h)}(B)} \tag{3}$$

$$P_{(i,h)}(B) = \sum_A P_{(i,h)}(B|A) \cdot P_{(i,h)}(A) \tag{4}$$

$$P_{(i,h)}(B|A) = \frac{T_{(i,h)}(B|A)}{T_{(i,h)}(A)} \tag{5}$$

where $P_{(i,h)}$ denotes the traffic status probability function of road segment (i,h), $A \in \{PT, DW\}$, and $B \in \{FF, NT, ST, TC\}$. $T_{(i,h)}$ denotes the time duration of traffic status on the road segment (i,h).

The likelihood function of traffic status probability for road segment (i,h) and its standardization are formulated as follows.

$$L_{(i,h)}(y) = \prod_{k=1}^{5} P_{(i,h)}(x_k|y) \tag{6}$$

$$P_{(i,h)}(y) = \frac{L_{(i,h)}(y)}{\sum_{k=1}^{4} L_{(i,h)}(y_k)} \tag{7}$$

where $L_{(i,h)}$ denotes the likelihood estimation function and $P_{(i,h)}$ denotes the normalized probability function of traffic status for road segment (i,h). x_k denotes the kth element of x, $x \in \{TV, AVS, RO, PT, DW\}$. y_k denotes the kth element of y, $y \in \{FF, NT, ST, TC\}$.

B. Parameter learning

Parameter learning is another core part of the online learning data-driven model, which can quantify and adjust the mapping parameters between traffic characteristics input vector of each road segment and traffic status output vector using online traffic data. Based on the proposed system architecture, online traffic data is transformed into labeled data, which can be used to construct a multi-layer perceptron (MLP) network. MLP is one of the deep neural networks, while deep learning is an emerging data-driven paradigm and has been used for solving complex and nonlinear traffic problems [57]. Compared with deep belief network (DBN) and stacked autoencoder (SAE), MLP shows promising results when dealing with labeled data.

Based on the traffic characteristics generated from model learning, Fig. 6 illustrates the network structure of an MLP with two hidden layers for parameter learning of each road segment. The input vector $\boldsymbol{x} = \{x_1, x_2, x_3, x_4, x_5\}$ passes through the hidden layers and generate output vector $\boldsymbol{y} = \{y_1, y_2, y_3, y_4\}$. $\boldsymbol{W_1}$, $\boldsymbol{W_2}$, and $\boldsymbol{W_3}$ denote the weighting coefficients of the MLP. These weighting coefficients are trained using backpropagation in a supervised manner, which are thus quantified and adjusted to form the mapping between input vector and output vector. Thus, the proposed MLP can be formulated as follows [58].
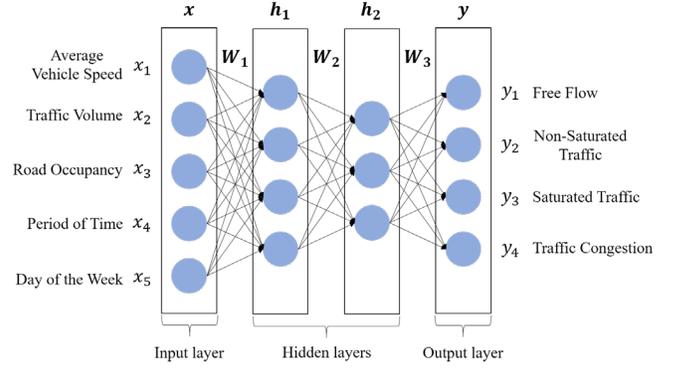

Fig. 6. Network structure of a multi-layer perception for parameter learning.

$$\boldsymbol{h_1} = g_1(\boldsymbol{W_1} \cdot \boldsymbol{x} + \boldsymbol{a}) \tag{8}$$

$$\boldsymbol{h_2} = g_2(\boldsymbol{W_2} \cdot \boldsymbol{h_1} + \boldsymbol{b}) \tag{9}$$

$$\boldsymbol{y} = g_3(\boldsymbol{W_3} \cdot \boldsymbol{h_2} + \boldsymbol{c}) \tag{10}$$

where $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ denote weights. g_1, g_2, and g_3 denote non-linear activation functions such as *sigmoid*, hyperbolic tangent *tanh*, and rectified linear function *ReLU*.

For each road segment (i,h), the errors between the observed value $f_{(i,h)}$ and the forecasted value $\hat{f}_{(i,h)}$ are quantified in terms of the mean absolute error (MAE), the mean relative error (MRE), and the root means square error (RMSE). The formulation of these indices is shown as follows [25].

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} |f_{(i,h)} - \hat{f}_{(i,h)}| \tag{11}$$

$$MRE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{f_{(i,h)} - \hat{f}_{(i,h)}}{f_{(i,h)}} \right| \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (f_{(i,h)} - \hat{f}_{(i,h)})^2} \tag{13}$$

C. Collaborative optimization mechanism

To enhance collaborative relations between road segments and vehicles, a collaborative optimization mechanism is thus proposed by combining short-term traffic forecasting and real-time routing optimization, with the consideration of three KPIs, namely travel distance, travel time, and fuel consumption. Real-time vehicular data is used to determine and update the optimal route for vehicles until the vehicle arrives at the destination.

Based on the online learning data-driven model, the traffic status probability of each route from R is formulated as follows.

$$P_{r_j}(y) = \frac{\prod_{(i,h) \in S_{r_j}} \beta_{(i,h)} \cdot \left[P_{(i,h)}(y) + \varepsilon_{r_j} \right]}{\sum_{k=1}^{4} \prod_{(i,h) \in S_{r_j}} [P_{(i,h)}(y_k) + \varepsilon_{r_j}]} \tag{14}$$

where P_{r_j} denotes the traffic status probability function of route r_j. If $P_{r_j}(y = FF \cup NT) > \varepsilon_{r_j}$, route r_j will be chosen as an alternative route and added to a new set R^*. $\beta_{(i,h)}$ is an indicator of road events, which is set as 0 if these is an event such as a traffic accident on road segment (i,h), otherwise it is set as 1. ε_{r_j} is a user-specified positive value.

Real-time routing optimization is then implemented with the consideration of three KPIs, namely travel distance, travel time, and fuel consumption. The objective of real-time routing optimization is to determine an optimal route with the minimum sum of weighted total travel distance, total travel time, and total fuel consumption, which is formulated as follows.

$$\min_{r_j \in R^*} \omega^{TD} \cdot \frac{TD_{r_j}}{TD_{max}} + \omega^{TT} \cdot \frac{TT_{r_j}}{TT_{max}} + \omega^{FC} \cdot \frac{FC_{r_j}}{FC_{max}} \quad (15)$$

$$\text{subject to } TD_{r_j} = \sum_{(i,h) \in S_{r_j}} d_{(i,h)} \quad (16)$$

$$TT_{r_j} = \sum_{(i,h) \in S_{r_j}} \frac{d_{(i,h)}}{AVS_{(i,h) + \sum_{k=1}^{i} \Delta t_{(i,h),k}}} \quad (17)$$

$$FC_{r_j} = \sum_{(i,h) \in S_{r_j}} [\alpha_{(i,h) + \sum_{k=1}^{i} \Delta t_{(i,h),k}} \cdot d_{(i,h)}] \quad (18)$$

where TD_{r_j}, TT_{r_j}, and FC_{r_j} denote the total travel distance, travel time, and fuel consumption of route r_j. TD_{max}, TT_{max}, and FC_{max} denote the maximum travel distance, travel time, and fuel consumption of routes from R^*. $AVS_{(i,h) + \sum_{k=1}^{i} \Delta t_{(i,h),k}}$ denotes the predicted average vehicle speed of road segment (i,h) at the predicted arrival time $t + \sum_{k=1}^{i} \Delta t_{(i,h),k}$, while the time interval of short-term forecasting is determined by the arrival time of the vehicle on the next road segment.

To generate a finite set R of feasible routes and the optimal route for vehicles, the ant colony optimization (ACO) algorithm is used, which is a meta-heuristic method for solving vehicular routing problems in an efficient manner [59]. Each artificial ant finds the shortest path from the start point to the destination in road networks. At each step, the next node is chosen based on the pheromone trails and heuristic information. The possibility $P_{(i,h)}^k$ of the kth ant moving from road node i to the next node h at time t is formulated as follows.

$$P_{(i,h)}^k(t) = \begin{cases} \frac{[\tau_{(i,h)}(t)]^{\alpha} \cdot [\eta_{(i,h)}(t)]^{\beta}}{\sum_{h \in N_k} [\tau_{(i,h)}(t)]^{\alpha} \cdot [\eta_{(i,h)}(t)]^{\beta}}, & h \in N_k \\ 0, & otherwise \end{cases} \quad (19)$$

where $\tau_{(i,h)}$ denotes the pheromone value on road segment (i,h), $\eta_{(i,h)}$ denotes the heuristic information.

During the iterative process, the optimal route is stored, and pheromone trails are updated considering evaporation ρ, until the termination rule is reached. The updating process can be formulated as follows.

$$\tau_{(i,h)}(t+1) = (1-\rho)\tau_{(i,h)}(t) + \sum_{k=1}^{n} \Delta \tau_{(i,h)}^k \quad (20)$$

where $\Delta \tau_{(i,h)}^k$ denotes pheromone value on road segment (i,h), released by the kth ant. ρ denotes evaporation rate, $0 < \rho < 1$.

V. CASE STUDY

To validate the feasibility and efficiency of the online learning collaborative method, a case study based on Xi'an city is presented. Roadside monitoring devices such as ANPR cameras, inductive-loop and microwave detectors were used for online traffic data collection. The traffic statistical data includes 2018 annual data and traffic data in 5-minute intervals, which was obtained from the Traffic Bureau of Shaanxi Province. Based on the Xi'an city road networks and the analysis of these data, simulation experiments were designed and conducted in the laboratory. Specifically, the uncertain impacts of dynamics such as traffic congestion and some road events such as traffic accidents were considered in this case study.

Fig. 7 illustrates a part of the Xi'an city road networks, which is composed of 358 road nodes and 593 road segments. The online information of all road segments for feasible routes is given in Table I. t denotes the arrival time on the road segment (i,h). Totally 49 road segments were under the saturated traffic
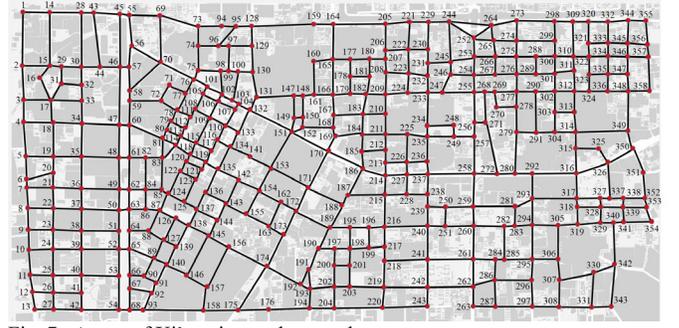

Fig. 7. A part of Xi'an city road networks.

or traffic congestion conditions, while 2 road events were specially set on the road segments (5, 19) and (119, 135).

Computational experiments were conducted with Hadoop 3.1.3 cluster on two servers each includes two Xeon E5-2690V4 CPU and 64 GB RAM. To begin with, initialization was executed, and the start time was set as 0s. The initial set of feasible routes was $R = \{r_1, r_2, r_3, r_4, r_5\}$ and ε_{r_j} was set at 0.1. The weighting coefficients $(\omega^{TD}, \omega^{TT}, \omega^{FC})$ were set at (0.3, 0.5, 0.2). In the simulation experiments, five test vehicles of the same type were driven from the start point to the destination, while the online learning data-driven model was running on the servers at the same time. All test vehicles were preinstalled with embedded devices including the sensors, processors, and communication modules. Real-time vehicular data was collected by on-board units and transmitted to the cloud computing platform through communication networks. Meanwhile, real-time optimal route generated by the online learning data-driven model was transmitted from the cloud computing platform to the vehicles.

The optimization results of the online learning data-driven model are shown in Table II. The free flow and non-saturated traffic status probabilities of routes r_1 and r_3 are 0 due to the road events on the road segments (119, 135). Consequently, the set of alternative routes is $R^* = \{r_2, r_4, r_5\}$.

The set of road node sequence for route r_2 is {2, 15, 29, 30, 44, 46, 47, 48, 49, 50, 63, 86, 87, 126, 138, 145, 156, 174, 192, 193, 194, 204, 220, 243, 242, 262, 263}. The total travel distance is 10.30km. The total travel time is 3444s, while the total fuel consumption is 0.63L.

The set of road node sequence for route r_4 is {2, 3, 4, 5, 6, 7, 8, 22, 37, 50, 63, 86, 87, 126, 138, 145, 156, 174, 192, 193, 194, 204, 220, 243, 242, 262, 263}. The total travel distance is 10.20km. The total travel time is 3211s, which is around 6.7% shorter than that of the route r_2. The total fuel consumption is 0.63L.

The set of road node sequence for route r_5 is {2, 15, 29, 30, 28, 43, 45, 55, 56, 70, 76, 101, 102, 103, 104, 132, 151, 170, 186, 214, 227, 237, 258, 259, 260, 261, 262, 263}. The total travel distance is 10.60km, which is roughly 3% longer than that of the route r_2 and roughly 4% longer than that of the route r_4. The total travel time is 2924s, which is roughly 15% shorter than that of the route r_2 and roughly 9% shorter than that of the route r_4. The total fuel consumption is 0.66L.

The efficiency of the proposed method was also evaluated, and the results show that the computation time of each process is less than 1s, which can meet the requirements of applications with the proposed method in real-time environments.

TABLE I
REAL-TIME INFORMATION OF ROAD SEGMENTS FOR FEASIBLE ROUTES

R	(i,h)	$d_{(i,h)}$(km)	t(s)	$AVS_{(i,h)}$(km/h)	$\alpha_{(i,h)}$(L/100km)	$\beta_{(i,h)}$	$P_{(i,h)}(FF)$	$P_{(i,h)}(NT)$	$P_{(i,h)}(ST)$	$P_{(i,h)}(TC)$
r_1	(2,15)	0.40	0	20	6.5	1	0.8	0.1	0.1	0
r_1	(15,29)	0.19	72	5	6.0	1	0.1	0.1	0.6	0.2
r_1	(29,30)	0.29	209	10	6.0	1	0.1	0.4	0.2	0.3
r_1	(30,44)	0.29	313	15	6.0	1	0.5	0.1	0.3	0.1
r_1	(44,46)	0.29	383	25	6.5	1	0.4	0.3	0.2	0.1
r_1	(46,47)	0.90	425	10	6.0	1	0.3	0.1	0.5	0.1
r_1	(47,60)	0.17	749	15	6.0	1	0.2	0.3	0.4	0.1
r_1	(60,81)	0.55	789	10	6.0	1	0.2	0.2	0.4	0.2
r_1	(81,114)	0.24	987	5	6.0	1	0.1	0.1	0.6	0.2
r_1	(114,118)	0.26	1160	10	6.0	1	0.1	0.2	0.5	0.2
r_1	(118,119)	0.27	1254	5	6.0	1	0.1	0.2	0.4	0.3
r_1	(119,135)	0.23	1448	5	6.0	0	0.1	0.1	0.5	0.3
r_1	(135,142)	0.35	1614	15	6.0	1	0.4	0.3	0.2	0.1
r_1	(142,154)	0.35	1698	15	6.0	1	0.3	0.4	0.1	0.2
r_1	(154,162)	0.40	1782	10	6.0	1	0.4	0.3	0.1	0.2
r_1	(162,172)	0.27	1926	20	6.5	1	0.4	0.2	0.3	0.1
r_1	(172,189)	0.60	1974	5	6.0	1	0.1	0.1	0.3	0.5
r_1	(189,195)	0.30	2406	10	6.0	1	0.2	0.3	0.4	0.1
...	...	...	...	...	...	...	....	...	...	...
r_1	(261,262)	0.40	3504	20	6.5	1	0.3	0.4	0.2	0.1
r_1	(262,263)	0.30	3576	15	6.0	1	0.3	0.3	0.2	0.2
r_2	(2,15)	0.40	0	20	6.5	1	0.8	0.1	0.1	0
r_2	(15,29)	0.19	72	5	6.0	1	0.1	0.1	0.6	0.2
r_2	(29,30)	0.29	209	10	6.0	1	0.1	0.4	0.2	0.3
...	...	...	...	...	...	...	....	...	...	...
r_2	(262,263)	0.30	3390	20	6.5	1	0.4	0.3	0.2	0.1
r_3	(2,3)	0.50	0	20	6.5	1	0.2	0.5	0.1	0.2
r_3	(3,4)	0.35	90	25	6.5	1	0.3	0.4	0.2	0.1
r_3	(4,18)	0.45	140	5	6.0	1	0.1	0.2	0.6	0.1
...	...	...	...	...	...	...	....	...	...	...
r_3	(119,135)	0.23	1610	5	6.0	0	0	0.1	0.8	0.1
...	...	...	...	...	...	...	....	...	...	...
r_3	(262,263)	0.30	3757	15	6.0	1	0.3	0.3	0.2	0.2
r_4	(2,3)	0.50	0	20	6.5	1	0.2	0.5	0.1	0.2
r_4	(3,4)	0.35	90	25	6.5	1	0.3	0.4	0.2	0.1
r_4	(4,5)	0.50	140	5	6.0	1	0.1	0	0.3	0.6
...	...	...	...	...	...	...	....	...	...	...
r_4	(262,263)	0.30	3139	15	6.0	1	0.3	0.3	0.2	0.2
r_5	(2,15)	0.40	0	20	6.5	1	0.8	0.1	0.1	0
r_5	(15,29)	0.19	72	5	6.0	1	0.1	0.1	0.6	0.2
r_5	(29,30)	0.29	209	10	6.0	1	0.1	0.4	0.2	0.3
...	...	...	...	...	...	...	....	...	...	...
r_5	(262,263)	0.30	2852	15	6.0	1	0.3	0.3	0.2	0.2

TABLE II
OPTIMIZATION RESULTS OF THE ONLINE LEARNING DATA-DRIVEN MODEL

Route	r_1	r_2	r_3	r_4	r_5
Road node sequence	{2,15,29,30,44,46,47, 60,81,114,118,119, 135,142,154,162,172, 189,195,196,216,240, 251,260,261,262,263}	{2,15,29,30,44,46,47, 48,49,50,63,86,87,126, 138,145,156,174,192, 193,194,204,220,243, 242,262,263}	{2,3,4,18,34,47,60,81, 114,118,119,135,142, 154,162,172,189,195, 196,216,240,251, 260,261,262,263}	{2,3,4,5,6,7,8,22,37, 50,63,86,87,126,138 ,145,156,174,192, 193,194,204,220, 243,242,262,263}	{2,15,29,30,28,43,45, 55,56,70,76,101,102, 103,104,132,151,170, 186,214,227,237,258, 259,260,261,262,263}
$P_{r_j,t}(FF \cup NT)$	0	0.64	0	0.82	0.89
Total travel distance (km)	9.40	10.30	9.39	10.20	10.60
Total travel time (s)	3648	3444	3829	3211	2924
Total fuel consumption (L)	0.57	0.63	0.58	0.63	0.66
Computation time (s)	0.40	0.35	0.46	0.42	0.43

To compare with the existing methods, including the GIS-based routing methods [30] and real-time routing methods [31], experimental performance simulation was also conducted based on the same benchmark. GIS-based routing method uses spatial data and chose route r_3 as the optimal route, while the real-time routing method uses real-time traffic data and chose r_2 as the optimal route. The comparison results are shown in Fig. 8. The proposed method outperformed the GIS-based routing method and real-time routing method by reducing the total travel time of 23.6% and 15% respectively. While to compare with other

prediction-based routing methods using limited traffic data from taxi GPS traces [60] or detectors [61], the proposed online learning collaborative method mainly focuses on unsolved problems of data integration from separate and independent traffic subsystems as well as the collaboration relations between road segments and vehicles.

Fig. 9 illustrates a proof of concept prototype system of the proposed online learning data-driven model. In future urban traffic systems, online traffic data from road segments and real-time vehicular data from individual vehicles are easily collected
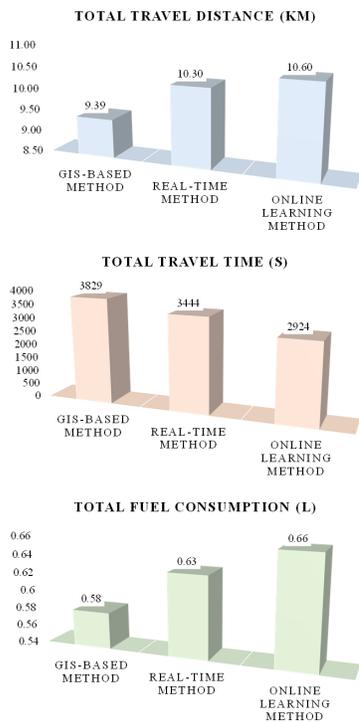
Fig. 8. Comparison between the existing methods and the proposed method.

Fig. 9. Proof of concept prototype system.

and aggregated by roadside monitoring devices and preinstalled on-board units (OBU). Thus, with the help of the CPS model and machine learning algorithms, road segments and vehicles are made 'smart', with the capability of active perception, active response, and autonomous decision-making. In Fig. 9, an OBU with a GPS tracker was installed on the vehicle to capture real-time vehicular data such as current location and speed, while roadside ANPR cameras and detectors were used to collect online traffic data. The collected data was uploaded to the cloud computing platform through wireless communication networks. The online learning data-driven model with model learning and parameter learning was executed on the cloud computing platform, extracting prior knowledge from the collected historical and online traffic data. The collaborative optimization mechanism was realized by combing short term traffic forecasting and real-time routing optimization. Consequently, the optimal route result was generated and transmitted to vehicular cloud clients such as thin clients, mobile apps, and web browsers.

For the scalability and latency problems, the optimal route results of computational experiments show that the proposed method has sub-second performance, thus it is reasonable to use in real-life applications. While data transmission is conducted using wireless communication networks, including Wi-Fi, satellite networks, IEEE 802.11p, and 4G/5G. IEEE 802.11p provides data rates ranging from 6 to 27 Mbps, while 4G provides high-speed Internet of 100 Mbps and the 5G edge rate ranges from 100 Mbps to 1 Gbps. Worldwide interoperable for Microwave Access (WiMAX) can be used to provide mobile broadband services at a vehicular speed of up to 350 km/h.

VI. CONCLUSION AND FUTURE WORK

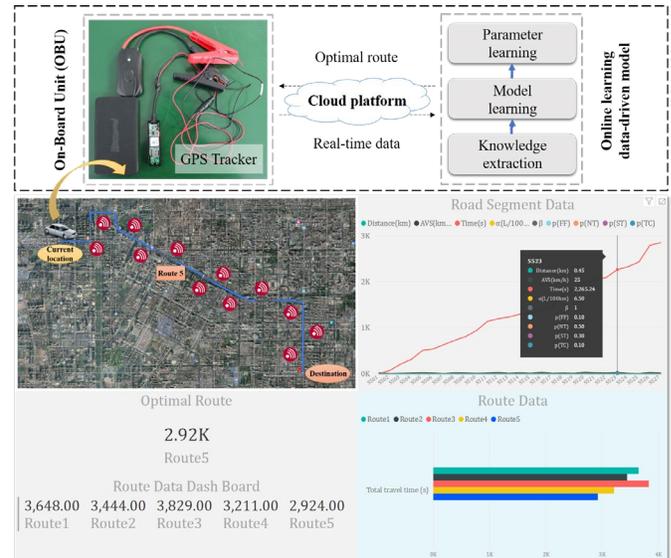To integrate distributed and separate traffic infrastructures and enhance collaboration between urban traffic systems and individual vehicles, an online learning collaborative method is proposed by combining short term traffic forecasting and real-time routing optimization. The main contributions of this work include the design of a three-layer system architecture of online learning collaborative optimization, traffic and vehicular data fusion and technologies integration, and a case study based on the Xi'an city with a proof of concept prototype system.

As the theoretical foundation and practical guide of the proposed online learning collaborative method, a three-layer system architecture is thus designed to integrate separate and independent traffic subsystems with computational resources. To establish the mapping between the physical entities and virtual models, a CPS model is developed to depict the real-time status and dynamical behavior of road segments and vehicles. Based on the online traffic data collected from road segments and real-time vehicular data captured from vehicles, an online learning data-driven model including model learning and parameter learning is constructed to extract prior knowledge from the collected data. Also, a collaborative optimization mechanism is proposed to strengthen the self-adaptive collaboration between road segments and vehicles by combining short-term traffic forecasting and real-time routing optimization.

To demonstrate the feasibility and efficiency of the proposed method, a case study based on Xi'an city is introduced with the consideration of three KPIs, namely travel distance, travel time, and fuel consumption. The experimental results indicate that the proposed method is effective to reduce the travel time and the computation time used is reasonable in real-life applications.

For future research, other types of models and methods for short-term traffic forecasting and real-time routing optimization will be further explored. Besides, the proposed method will also be adopted in real-life environments.

REFERENCES

[1] D. Schrank, B. Eisele, T. Lomax, and J. Bak, "2015 Urban Mobility Scorecard," Texas A&M Transp. Inst., College Station, TX, USA, 2015.

[2] G. Dimitrakopoulos and P. Demestichas, "Intelligent Transportation Systems," *IEEE Veh. Technol. Mag.*, vol. 5, no. 1, pp. 77–84, Mar. 2010.

[3] A. Whitmore, A. Agarwal, and L. Da Xu, "The Internet of Things—A survey of topics and trends," *Inf. Syst. Front.*, vol. 17, no. 2, pp. 261–274, Apr. 2015.

[4] Y. Zhang, Z. Guo, J. Lv, and Y. Liu, "A Framework for Smart Production-Logistics Systems based on CPS and Industrial IoT," *IEEE Trans. Ind. Informatics*, vol. 14, no. 9, pp. 4019 – 4032, 2018.

[5] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," *J. Netw. Comput. Appl.*, vol. 40, pp. 325–344, 2014.

[6] Y. Zhang, Z. Guo, C. Qian, and R. Li, "Investigation on Process-aware Based Intelligent Modeling of Bottom Layer Manufacturing Resources and Self-adaptive Collaborative Optimization Methodology," *J. Mech. Eng.*, vol. 54, no. 16, pp. 1–10, 2018.

[7] J. Wan, D. Zhang, S. Zhao, L. Yang, and J. Lloret, "Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 106–113, Aug. 2014.

[8] Z. Guo, Y. Zhang, X. Zhao, and X. Song, "A Timed Colored Petri Net Simulation-Based Self-Adaptive Collaboration Method for Production-Logistics Systems," *Appl. Sci.*, vol. 7, no. 3, pp. 235:1–15, 2017.

[9] R. Kitchin, "The real-time city? Big data and smart urbanism," *GeoJournal*, vol. 79, no. 1, pp. 1–14, Feb. 2014.

[10] A. B. Sharma, F. Ivančić, A. Niculescu-Mizil, H. Chen, and G. Jiang, "Modeling and analytics for cyber-physical systems in the age of big data," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 41, no. 4, pp. 74–77, Apr. 2014.

[11] J. Wan, D. Zhang, Y. Sun, K. Lin, C. Zou, and H. Cai, "VCMIA: A Novel Architecture for Integrating Vehicular Cyber-Physical Systems and Mobile Cloud Computing," *Mob. Networks Appl.*, vol. 19, no. 2, pp. 153–160, Apr. 2014.

[12] H. O. Al-sakran and H. O. Al-sakran, "Intelligent Traffic Information System Based on Integration of Internet of Things and Agent Technology," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 2, pp. 37–43, 2015.

[13] M. Lippi, M. Bertini, and P. Frasconi, "Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 871–882, Jun. 2013.

[14] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transp. Res. Part C Emerg. Technol.*, vol. 43, pp. 50–64, Jun. 2014.

[15] H.-F. Yang, T. S. Dillon, and Y.-P. P. Chen, "Optimized Structure of the Traffic Flow Forecasting Model With a Deep Learning Approach," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 10, pp. 2371–2381, Oct. 2017.

[16] L. Qu, W. Li, W. Li, D. Ma, and Y. Wang, "Daily long-term traffic flow forecasting based on a deep neural network," *Expert Syst. Appl.*, vol. 121, pp. 304–312, May 2019.

[17] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, "Dynamic route planning with real-time traffic predictions," *Inf. Syst.*, vol. 64, pp. 258–265, 2017.

[18] J. Wang and H. Niu, "A distributed dynamic route guidance approach based on short-term forecasts in cooperative infrastructure-vehicle systems," *Transp. Res. Part D Transp. Environ.*, May 2018.

[19] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transp. Res. Part C Emerg. Technol.*, vol. 43, pp. 3–19, 2014.

[20] B. L. Smith, B. M. Williams, and R. Keith Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. Part C Emerg. Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.

[21] Y. Xie, Y. Zhang, and Z. Ye, "Short-Term Traffic Volume Forecasting Using Kalman Filter with Discrete Wavelet Decomposition," *Comput. Civ. Infrastruct. Eng.*, vol. 22, no. 5, pp. 326–334, Jul. 2007.

[22] H. Sun, H. Liu, H. Xiao, R. He, and B. Ran, "Use of Local Linear Regression Model for Short-Term Traffic Forecasting," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1836, pp. 143–150, Jan. 2003.

[23] S. Sun, C. Zhang, and G. Yu, "A Bayesian Network Approach to Traffic Flow Forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 124–132, Mar. 2006.

[24] P. Cortez, M. Rio, M. Rocha, and P. Sousa, "Multi-scale Internet traffic forecasting using neural networks and time series methods," *Expert Syst.*, vol. 29, no. 2, p. no–no, Dec. 2010.

[25] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic Flow Prediction With Big Data: A Deep Learning Approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 1–9, 2014.

[26] D. Wei and H. Liu, "An Adaptive-Margin Support Vector Regression for Short-Term Traffic Flow Forecast," *J. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 317–327, Oct. 2013.

[27] K. Y. Chan, T. S. Dillon, and E. Chang, "An Intelligent Particle Swarm Optimization for Short-Term Traffic Flow Forecasting Using on-Road Sensor Systems," *IEEE Trans. Ind. Electron.*, vol. 60, no. 10, pp. 4714–4725, Oct. 2013.

[28] Y. Zhang and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transp. Res. Part C Emerg. Technol.*, vol. 43, pp. 65–78, 2014.

[29] F. Moretti, S. Pizzuti, S. Panzieri, and M. Annunziato, "Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling," *Neurocomputing*, vol. 167, pp. 3–7, Nov. 2015.

[30] S. Krichen, S. Faiz, T. Tlili, and K. Tej, "Tabu-based GIS for solving the vehicle routing problem," *Expert Syst. Appl.*, vol. 41, no. 14, pp. 6483–6493, 2014.

[31] T. Cheong and C. C. White, "Dynamic traveling salesman problem: Value of real-time traffic information," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 619–630, 2012.

[32] M. Gendreau, G. Ghiani, and E. Guerriero, "Time-dependent routing problems: A review," *Comput. Oper. Res.*, vol. 64, pp. 189–197, 2015.

[33] H. Bast, D. Delling, A. Goldberg, M. Müller-Hannemann, T. Pajor, P. Sanders, D. Wagner, and R. F. Werneck, "Route Planning in Transportation Networks," in *Algorithm Engineering*, Springer, Cham, 2016, pp. 19–80.

[34] M. W. Levin, S. D. Boyles, J. Duthie, and C. M. Pool, "Demand Profiling for Dynamic Traffic Assignment by Integrating Departure Time Choice and Trip Distribution," *Comput. Civ. Infrastruct. Eng.*, vol. 31, no. 2, pp. 86–99, Feb. 2016.

[35] Y.-C. Chiu, J. Bottom, M. Mahut, A. Paz, R. Balakrishna, T. Waller, and J. Hicks, "Dynamic Traffic Assignment: A Primer," *Transp. Res. E-Circular*, no. E-C153, 2011.

[36] K. Abdelghany, A. Hassan, A. Alnawaiseh, and H. Hashemi, "Flow-Based and Density-Based Time-Dependent Demand Estimation for Congested Urban Transportation Networks," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2498, pp. 27–36, Jun. 2015.

[37] E. Mitsakis, J. M. S. Grau, I. Stamos, and G. Aifadopoulou, "An integrated framework for embedding large-scale dynamic traffic assignment models in advanced traveler information systems," *Transp. Plan. Technol.*, vol. 38, no. 8, pp. 866–877, Nov. 2015.

[38] X. Li, W. H. K. Lam, H. Shao, and Z. Gao, "Dynamic modelling of traffic incident impacts on network reliability," *Transp. A Transp. Sci.*, vol. 11, no. 9, pp. 856–872, Oct. 2015.

[39] Wu He, Gongjun Yan, and Li Da Xu, "Developing Vehicular Data Cloud Services in the IoT Environment," *IEEE Trans. Ind. Informatics*, vol. 10, no. 2, pp. 1587–1595, May 2014.

[40] J. Kang, R. Yu, X. Huang, M. Jonsson, H. Bogucka, S. Gjessing, and Y. Zhang, "Location privacy attacks and defenses in cloud-enabled internet of vehicles," *IEEE Wirel. Commun.*, vol. 23, no. 5, pp. 52–59, Oct. 2016.

[41] J. Akbari Torkestani, "Mobility prediction in mobile wireless networks," *J. Netw. Comput. Appl.*, vol. 35, no. 5, pp. 1633–1645, 2012.

[42] J. Cheng, J. Cheng, M. Zhou, F. Liu, S. Gao, and C. Liu, "Routing in Internet of Vehicles: A Review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2339–2352, Oct. 2015.

[43] M. R. Rahimi, J. Ren, C. H. Liu, A. V. Vasilakos, and N. Venkatasubramanian, "Mobile Cloud Computing: A Survey, State of Art and Future Directions," *Mob. Networks Appl.*, vol. 19, no. 2, pp. 133–143, Apr. 2014.

[44] M. Gerla, "Vehicular Cloud Computing," in *2012 The 11th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, 2012, pp. 152–155.

[45] M. Shiraz, A. Gani, R. H. Khokhar, and R. Buyya, "A Review on Distributed Application Processing Frameworks in Smart Mobile Devices for Mobile Cloud Computing," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 3, pp. 1294–1313, 2013.

[46] P. Basanta-Val, N. C. Audsley, A. J. Wellings, I. Gray, and N. Fernandez-Garcia, "Architecting Time-Critical Big-Data Systems," *IEEE Trans. Big Data*, vol. 2, no. 4, pp. 310–324, Dec. 2016.

[47] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, "Next-Generation Big Data Analytics: State of the Art, Challenges, and Future Research Topics," *IEEE Trans. Ind. Informatics*, vol. 13, no. 4, pp. 1891 – 1899, 2017.

[48] Q. Li, T. Zhang, and Y. Yu, "Using cloud computing to process intensive floating car data for urban traffic surveillance," *Int. J. Geogr. Inf. Sci.*, vol. 25, no. 8, pp. 1303–1322, Aug. 2011.

[49] L. Zhou, N. Chen, and Z. Chen, "Efficient Streaming Mass Spatio-Temporal Vehicle Data Access in Urban Sensor Networks Based on Apache Storm," *Sensors*, vol. 17, no. 4, pp. 815:1–14, Apr. 2017.

[50] P. Basanta-Val, N. Fernández-García, A. J. Wellings, and N. C. Audsley, "Improving the predictability of distributed stream processors," *Futur. Gener. Comput. Syst.*, vol. 52, pp. 22–36, 2015.

[51] I. Stoica, "Trends and challenges in big data processing," in *Proceedings of the VLDB Endowment*, 2016, vol. 9, no. 13, pp. 1619–1619.

[52] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker, and I. Stoica, "Apache Spark: A Unified Engine for Big Data Processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, Nov. 2016.

[53] M. A. Alsheikh, D. Niyato, S. Lin, H. Tan, and Z. Han, "Mobile big data analytics using deep learning and apache spark," *IEEE Netw.*, vol. 30, no. 3, pp. 22–29, May 2016.

[54] P. Basanta-Val and M. Garcia-Valls, "A Distributed Real-Time Java-Centric Architecture for Industrial Systems," *IEEE Trans. Ind. Informatics*, vol. 10, no. 1, pp. 27–34, Feb. 2014.

[55] I.-G. Chong and C.-H. Jun, "Performance of some variable selection methods when multicollinearity is present," *Chemom. Intell. Lab. Syst.*, vol. 78, no. 1, pp. 103–112, 2005.

[56] D. Heckerman, "A Tutorial on Learning with Bayesian Networks," in *Learning in Graphical Models*, Dordrecht: Springer Netherlands, 1998, pp. 301–354.

[57] Y. Wang, D. Zhang, Y. Liu, B. Dai, and L. H. Lee, "Enhancing transportation systems via deep learning: A survey," *Transportation Research Part C: Emerging Technologies*, vol. 99. Elsevier Ltd, pp. 144–163, 01-Feb-2019.

[58] M. Veres and M. Moussa, "Deep Learning for Intelligent Transportation Systems: A Survey of Emerging Trends," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–17, Jul. 2019.

[59] S. K. Lakshmanaprabu, K. Shankar, S. Sheeba Rani, E. Abdulhay, N. Arunkumar, G. Ramirez, and J. Uthayakumar, "An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: Towards smart cities," *J. Clean. Prod.*, vol. 217, pp. 584–593, Apr. 2019.

[60] X. Niu, Y. Zhu, Q. Cao, X. Zhang, W. Xie, and K. Zheng, "An Online-Traffic-Prediction Based Route Finding Mechanism for Smart City," *Int. J. Distrib. Sens. Networks*, vol. 11, no. 8, p. 970256, Aug. 2015.

[61] T. Liebig, N. Piatkowski, … C. B.-E., and U. 2014, "Predictive Trip Planning-Smart Routing in Smart Cities," in *Workshop Pro- ceedings of the EDBT/ICDT 2014 Joint Conference*, 2014, pp. 331–338.
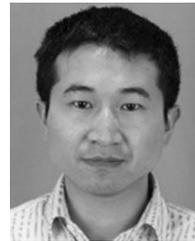
Zhengang Guo received the B.Eng. degree in mechatronic engineering from Northwestern Polytechnical University, Xi'an, China, in 2015, where he is currently working toward the Ph.D. degree. He is also a joint Ph.D. student with the Department of Electrical and Electronic Engineering, Cardiff University, Cardiff, UK, sponsored by the China Scholarship Council.

His research interests include smart manufacturing, intelligent transportation, complex network modeling and optimization.


Yingfeng Zhang (M'17) received the B.Eng., M.S., and Ph.D. degrees in mechanical engineering all from Xi'an Jiaotong University, Xi'an, China, in 1999, 2002, and 2005, respectively.

He is currently a Professor with the Department of Industrial Engineering, Northwestern Polytechnical University, Xi'an, China. His research interests include smart manufacturing, digital twin and core models of future intelligent industrial systems.

Jingxiang Lv received the Ph.D degree in mechatronic engineering from Zhejiang University, Hangzhou, China, in 2014.

He is currently a postdoctoral fellow with the Department of Industrial Engineering, Northwestern Polytechnical University, Xi'an, China. His research interests include green manufacturing, energy consumption modeling and optimization.
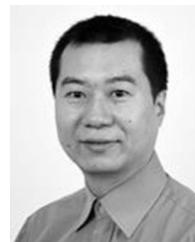
Yang Liu received his M.Sc. (Tech.) degree in Telecommunications Engineering and D.Sc. (Tech.) degree in Industrial Management both from University of Vaasa, Finland.

He is currently an Associate Professor with the Department of Management and Engineering (IEI), Linköping University, Linköping, Sweden. His research interests include global operations strategy, decision support tools, competitive advantage, sustainable operations, advanced manufacturing; control systems, autonomous robots, signal processing and pattern recognition.
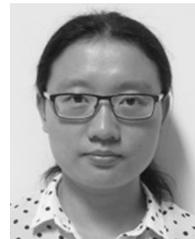
Ying Liu received the B.Eng. degree in industrial engineering from Zhejiang University, Hangzhou, China, in 2008, and the Ph.D. degree in manufacturing engineering and operations management from the University of Nottingham, Nottingham, UK, in 2014.

She is currently a Lecturer with the School of Engineering, University of Glasgow, Glasgow, UK. Her research interests include sustainable manufacturing, intelligent manufacturing and multi-objective optimisation.