




Article

Building Urban Public Traffic Dynamic Network Based on CPSS: An Integrated Approach of Big Data and AI

Gang Xiong ^{1,2,†} , Zhishuai Li ^{1,3} , Huaiyu Wu ^{1,*}, Shichao Chen ^{1,4} , Xisong Dong ¹, Fenghua Zhu ^{1,2} and Yisheng Lv ¹

¹ The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; gang.xiong@ia.ac.cn (G.X.); lizhishuai2017@ia.ac.cn (Z.L.); shichao.chen@ia.ac.cn (S.C.); xisong.dong@ia.ac.cn (X.D.); fenghua.zhu@ia.ac.cn (F.Z.); yisheng.lv@ia.ac.cn (Y.L.)

² The Cloud Computing Center, Chinese Academy of Sciences, Dongguan 523808, China

³ School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

⁴ The Institute of Systems Engineering and Collaborative Laboratory for Intelligent Science and Systems, Macau University of Science and Technology, Macau 999078, China

* Correspondence: huaiyu.wu@ia.ac.cn

† Current address: No. 95 Zhongguancun East Road, Haidian District, Beijing, China.

Abstract: The extensive proliferation of urban transit cards and smartphones has witnessed the feasibility of the collection of citywide travel behaviors and the estimation of traffic status in real-time. In this paper, an urban public traffic dynamic network based on the cyber-physical-social system (CPSS-UPTDN) is proposed as a universal framework for advanced public transportation systems, which can optimize the urban public transportation based on big data and AI methods. Firstly, we introduce three modules and two loops which composes of the novel framework. Then, the key technologies in CPSS-UPTDN are studied, especially collecting and analyzing traffic information by big data and AI methods, and a particular implementation of CPSS-UPTDN is discussed, namely the artificial system, computational experiments, and parallel execution (ACP) method. Finally, a case study is performed. The data sources include both traffic congestion data from physical space and cellular data from social space, which can improve the prediction performance for traffic status. Furthermore, the service quality of urban public transportation can be promoted by optimizing the bus dispatching based on the parallel execution in our framework.

Keywords: urban public transportation; cyber-physical-social system; advanced public transportation systems; big data



Citation: Xiong, G.; Li, Z.; Wu, H.; Chen, S.; Dong, X.; Zhu, F.; Lv, Y. Building Urban Public Traffic Dynamic Network Based on CPSS: An Integrated Approach of Big Data and AI. *Appl. Sci.* **2021**, *11*, 1109. <https://doi.org/10.3390/app11031109>

Academic Editor: Vassilis Charissis
Received: 1 December 2020
Accepted: 19 January 2021
Published: 26 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Urban public transportation is the main means to meet the increasing travel demand of metropolitan and alleviate traffic jams [1]. However, the sharing rate of public transportation is increasing inconspicuously. On the one hand, the private cars grow rapidly. On the other hand, there are still many problems in urban public transportation, such as insufficient transparency and sharing among traffic information, and inconvenient transition among different public transport means, which makes citizens unwilling to travel with public transportation.

Advanced public transportation systems (APTS) is an important component of intelligent transportation systems (ITS). Cutting-edge technologies of information and communication are applied for APTS to coordinate commands and control vehicles. With the popularity of intelligent terminals, massive data becomes accessible and is analyzed by artificial intelligence (AI), greatly enriching APTS. For example, the taxi software “Uber” provides the personalized real-time service for passengers and drivers in most cities all over the world. The real-time bus software called “Bus is coming” can provide real-time

information about buses in many Chinese cities. The success of these applications confirms that APTS can support traffic scheduling, and makes the services of taxis and buses more qualified.

However, the obtained multiple sources are isolated, thus leading to inefficiency. The arising of cyber-physical-social system (CPSS) provides a comprehensive approach for multi-source data fusion, which integrates physical system and social system into a unified cyber system. To realize real-time management and harness the human factors in public transportation systems, we propose a universal CPSS framework, which includes a small loop for real-time feedback and a big loop for periodical updates. To this end, we illustrate a feasible proposal on how to collect data and utilize advanced models, exploiting human and social factors further. Finally, we demonstrate the proposed framework via a computational case study. Specifically, the proposed framework is driven by big data and AI, in which urban public traffic dynamic network based on the cyber-physical-social system (CPSS-UPTDN) is built to promote the effective management of urban public transportation.

Our contributions are summarized as follows: (1) The framework about CPSS-UPTDN is proposed for APTS, and the basic modules of CPSS in transportation is also researched. (2) The key technologies are provided for our proposed framework. (3) The feasibility of traffic congestion prediction within the framework is proven by a detailed case.

Specifically, we record traffic congestion in physical space by web-crawler. Then, the cellular data is analyzed and incorporated into the traffic-unit level according to the traces of citizen, and the travel demand is fused with the traffic data for analysis, to judge the traffic status accurately, which is in favor of adjusting the public transport operation timely and dynamically.

The rest of this paper is organized as follows. In Section 2, we briefly review the research status and development of urban public transportation, cyber-physical system (CPS) with CPSS, and CPSS for transportation. The framework of CPSS-UPTDN and its research schemes are proposed in Section 3. Next, the key technologies are introduced for big data acquisition, the application of AI methods, and specific implementation (ACP method) in Section 4. Then, a particular case conducted with physical and social data is exemplified in Section 5. Finally, Section 6 concludes this paper.

2. Research and Development Trends

With the gradual expansion of cities and the remarkable increase in population, public transport priority is an inevitable choice for urban traffic mode. In this section, we review the literature on social systems analysis based on human behavior data.

2.1. Development of Urban Public Transportation

In the field of urban public transportation research, Aslam et al. developed taxi congestion bypass program and system based on the real-time perception of the road congestion status, and the simulation results of more than 1000 vehicles can reduce the travel time about 15% [2]. Lin et al. studied the classification characteristics of bus passenger trips based on those historical data from the management system for bus passengers [3]. Kokkinogenis et al. studied the establishment of artificial transportation system considering the social factors and policy elements, and validated the rationality and effectiveness of the various transportation policies on the travelers' behaviors through various experiments [4]. Chen et al. proposed a data-driven method for dynamic people-flow prediction, where the feasibility of people-flow prediction using real-time cellular probe data was investigated [5].

In China, about 700 "Smart Cities" are being constructed, and public transportation is an important part [6]. However, there are still prominent problems, such as the insufficient construction of public transportation infrastructure, and the low transportation supply capacity. The travel sharing ratio of the public traffic in Chinese large cities is only about 20% on average, and sharing ratio of the public traffic in a small and medium-sized city is less than 10% [7].

2.2. From CPS to CPSS

The Cyber-Physical System was proposed by the USA's national science foundation (NSF) in 2007, which is a complex system mainly consisting of engineering complexity demand. It is becoming a hot topic all over the world [8]. The international business machines corporation (IBM) puts forward that CPS is a strategic application concept practice of "Smart Earth". The NSF has funded more than 500 research projects in such areas as basic theory, method tools, and platform systems of CPS.

However, the existing researches of complex systems mainly focus on engineering complexity. The social complexity elements like human acting as designers, builders, operators, and terminal users, usually play important roles in these complex systems. Therefore, to achieve secure reliability and efficient management of those systems, the engineering complexity elements and social complexity elements should be studied as an indivisible whole. To address these problems, Wang firstly proposed CPSS in which those complex systems can integrate more even all elements with in-depth wisdom gradually, and then build up more and more complex systems [9]. CPSS is a kind of complex system composed of physical system, social system, and cyber system, in which a cyber system connects the physical system and social system [10,11].

Based on CPS, the social system is added in CPSS which contains humans and social factors. Moreover, the cyber system can be achieved through an artificial system to connect the social and physical system [12,13]. Effective management strategies are generated through the interaction between artificial and real systems. It is an organic combination of human organization and physical entity system through intelligent human-machine interaction, which is the basis of intelligent management for complex systems [14].

At present, with the rapid development of new information technologies like mobile Internet, cloud computing, big data, and AI, the study of CPSS becomes a common demand of the many national strategy industries, and also becomes the frontier of theory, technology, and application of "system science and systems engineering discipline".

2.3. CPSS for Public Transportation

The current study on traffic CPS does not consider the behaviors of travelers, which leads to incomplete research. Therefore, the effective combination of the human society and physical system through CPSS is expected to achieve sufficient management for complex public transportation systems [15,16].

In the field of CPSS for transportation, knowledge-based methods remain popular. Figueiras et al. developed a personalized and interactive transportation service mobile application based on the support of European Union "FP7 MobiS" project [17]. Liu et al. used a knowledge-based method to estimate automatic passenger flow [18]. For the first time, considering social factors, the authors conducted a special analysis and research of the traffic CPSS [19]. Zheng et al. discussed the acquisition, cleaning, and fusion techniques of various big data, and the integration, extraction, and application of real-time transportation conditions based on natural language and text [20]. He et al. used mobile data, bus cards, and other social signals to collect transportation demand, forecast transportation congestion and found the root reasons, and then provided the best transportation path and other guidance [21]. Calastri et al. captured social network structures, lifetime events, short-term travel, and activity planning from multi-source data collections [22]. Vinel et al. addressed tight coupling between the performance of vehicle-to-vehicle communications and the performance of cooperative ITSs' safety applications [23]. Han et al. proposed a CPSS-based parallel driving framework, which consists of connected vehicles with different levels of automation and necessitates a unified approach for future smart and safe driving [24].

From the research overview, CPSS in transportation is becoming the forefront of theoretical research, and is the application basis in urban public transportation. Researching the CPSS in public transportation is necessary for human traveling behaviors, and can effect, or even dominate the performance of an urban public transportation system.

3. The Framework of CPSS-UPTDN

In this section, inspired by the CPS-ITS [11], the CPSS-UPTDN framework are presented. Concretely, the overall framework of CPSS for UPTDN includes three modules (i.e., distributed storage, artificial system, and real-time system), and two loops (i.e., a small loop and a big one) and their relationship is illustrated in Figure 1. Then we will describe the components of Figure 1 in detail and discuss the integration of big data and AI.

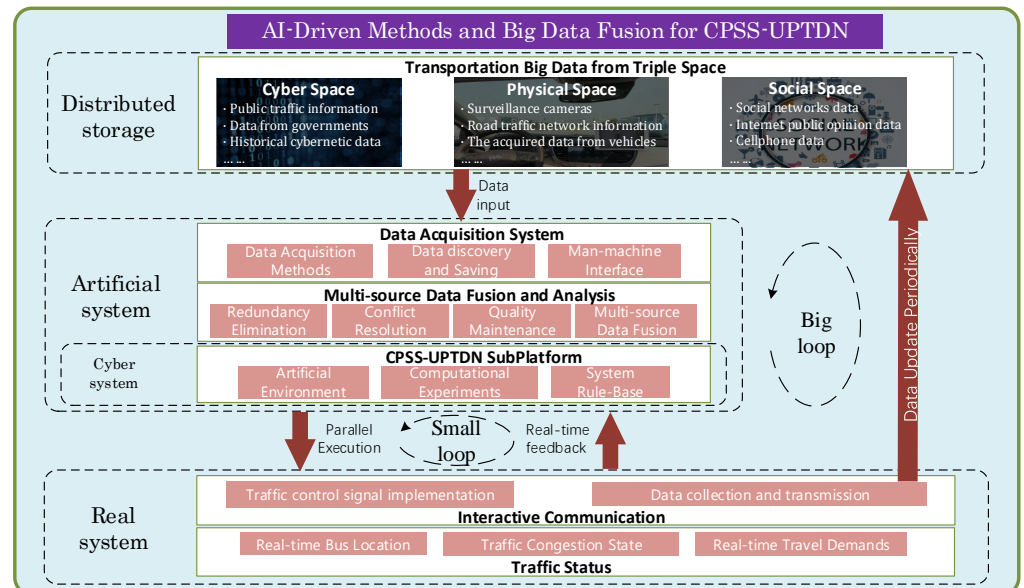


Figure 1. The proposed urban public traffic dynamic network based on the cyber-physical-social system (CPSS-UPTDN) is mainly composed of three modules and two loops.

3.1. Basic Modules in CPSS-UPTDN

CPSS-UPTDN consists of three modules and two loops. The distributed storage module is responsible for data storage, the interaction of artificial system module and real system module implements parallel control strategies. At the same time, considering the tremendous traffic data, our framework contains a “Small Loop for real-time status feedback” and a “Big Loop for periodical data update”. In the small loop, it can timely obtain the travel needs of users and give feedback in the small loop. For city-wide analysis, under the low requirement for the real-time situation, the big loop can be used to update data periodically, which requires the prediction methods to timely track global changes.

Distributed storage: The data set can be acquired from triple space, namely cyber, physical and social space. Traffic data from cyberspace includes public traffic information (bus routes, ticketing, bus stations, routes radiation diagram, transportation hubs, and so on), data from governments, etc. Data in physical space contains surveillance cameras, road traffic network information (road, length and charge information, etc.), and the acquired data for vehicles includes speed, vehicle information, and congestion of roads, etc. Social space data in various forms is mainly from Internet public opinion data and social networks, such as search engines, news portals, Weibo, and cellular signals, from which travel traces can be extracted. The mass data can be logically stored in Hbase which is developed on hadoop distributed file system (HDFS), which can realize the distributed and parallel computing.

Artificial system: (1) *Data Acquisition System.* The distributed data is calculated by MapReduce. Data verification and preliminary screening are implemented through man-machine interfaces, which realizes data preprocessing and correlation. (2) *Multi-source Data Fusion and Analysis.* This module is responsible for data cleaning and fusing the cross-space data. To facilitate data processing, the redundant data needs to be eliminated concerning task requirements. Then, the fusion process is performed, which consists of

two steps. One is called feature-level fusion, which is to predict the traffic parameters based on the feature of existing data. The other is state-level fusion, which judges the traffic state according to the current public transportation status information. The basic pipeline of fusion and analysis includes multi-source information extraction, preprocessing, fusion, target parameter acquisition, and state estimation. (3) *Cyber System (CPSS-UPTDN SubPlatform)*. In CPSS, the cyber system can be constructed as a part of the artificial system to connect the physical system (i.e., objects and environment) with the social system (i.e., humans and social factors). Thus, the construction of the sub-platform is the most important module.

Real system: (1) *Interactive communication*. CPSS-UPTDN has stringent requirements for rapid and reliable dissemination of traffic-related information. The real-time travel demand generated in the real system is fed back to the artificial system, and relevant computational experiments are designed to predict effective schemes for the real system. With the improvement of Vehicular-communication technologies, real-time capability can be improved further. (2) *Traffic status monitor*. The traffic system can be enhanced by mass qualities distributed sensors, such as on-board units, roadside units, cellphones, public opinion, and GPS. With the various sensors, multi-grained data are collected and updated to the distributed storage module periodically. Based on various transportation scenarios with computational experiments, the behaviors of a real urban public transportation system can be monitored and analyzed. The elements and rules in a real system are applied to the artificial system, including various learning strategies, optimization algorithms, and transportation scenes.

Small Loop for real-time status feedback: The small loop handles the local information concerned by managers. When real urban public transportation feeds back traffic status, the cyber system outputs the predicted traffic status and dispatching scheme according to relevant computational experiments. Then the effective scheme will be transmitted to the real system through parallel execution in real-time.

Big Loop for periodical data update: The big loop deals with the global information in urban public transportation, and updates all the models in the artificial system using multi-source data periodically. The real system generates new data about traffic status constantly. It is unrealistic that feedback these data to distributed storage module real-time. So a big loop is designed to periodically update these massive data and then used to optimize the models in the artificial system. Through real-time feedback and periodical update, this kind of interaction between the real and artificial systems can effectively manage urban public transportation.

3.2. The Integration of AI and Big Data in CPSS-UPTDN

The development of AI and big data has greatly promoted the progress of urban public transportation. Plenty of traffic control strategies and AI algorithms have developed regarding big data and the growing need for real-time traffic information in ITS [25,26]. CPSS-UPTDN provides a novel approach to merge heterogeneous data and AI models for traffic prediction and resource dispatching.

The integration of AI and big data is embodied for data preprocessing and artificial environment construction in the platform. For the annotated data such as cellular networks and social media, unsupervised learning (e.g., text feature extraction, clustering) is used to filter out raw data, and implement dimension reduction and extract features. The travel demand under different spatio-temporal granularity can be conducted and summarized. For traffic status estimation, prediction models based on deep learning can be leveraged. For some decision-making problems such as traffic signal and bus dispatching, schemes generated by reinforcement learning can be examined and tested in an artificial environment, which can improve the traffic capacity of the road network and service quality.

4. Key Technologies of CPSS-UPTDN

In this section, we will mainly introduce the key technologies in CPSS-UPTDN, which include the acquisition and transmission of big data, application of AI methods, and the specific implementation method.

4.1. Collection and Transmission for Big Data in UPTDN

Data acquisition technology is a crucial stage and the foundation of the follow-up. On the basis of the acquisition mode of big data center designed by [27], we introduced the multi-source data acquisition module. As Figure 2 shows, the multi-source big data in UPTDN can be derived from the bus terminals, rail transit, real-time road condition, travel demand of residents, traffic data in the social network, and so on. They correspond to a data collection device. The collected data will be transmitted to the data center by wire or wireless. The CPSS-UPTDN sub-platform will interact with the real system in time.

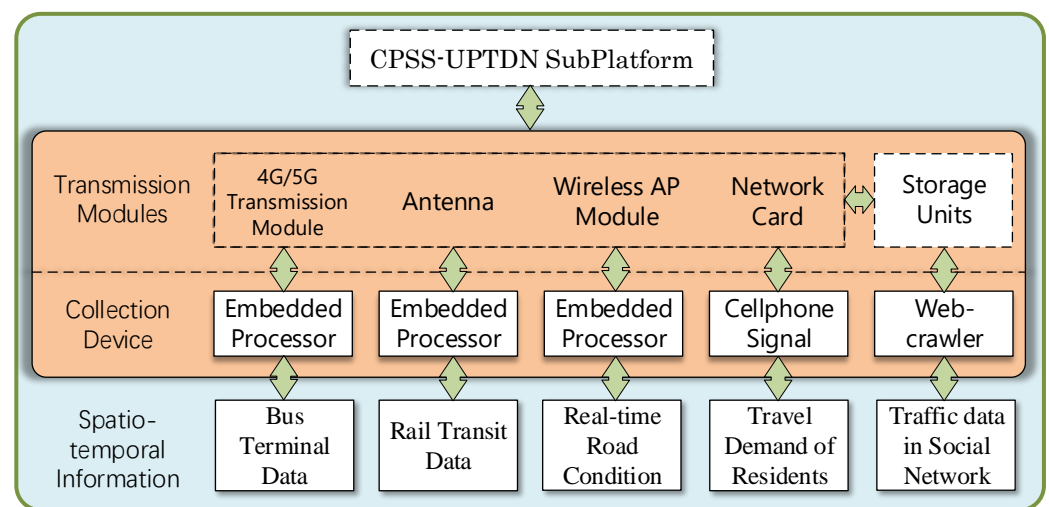


Figure 2. Data collection and transmission to CPSS-UPTDN Subplatform.

4.2. AI Methods in UPTDN

AI methods are used to collect data related to urban transportation, then traffic information elements are extracted, and semantic analysis of these data is conducted with the help of natural language processing and other models. Clustering and deep learning algorithms can be used to explore the travel patterns of urban residents and the fluctuation of traffic status. Meanwhile, reinforcement learning and other related algorithms are used to determine the dispatching schemes of real urban public transportation.

The origin data can be structured (data in database) or semi-structured (text, graphics, image data), or even heterogeneous data distributed over the network. So the methods for data acquisition are various, and could include data mining, statistical machine learning, natural language processing, and so on. The methods for data completion and multi-modal information association mainly involve subspace learning, sparse representation, and pattern recognition. The construction of sub-platform requires the regression methods, including deep learning, reinforcement learning, etc., to predict the supply and demand of real systems and then develop dynamic dispatching schemes.

4.3. Implementation of CPSS-UPTDN Based on ACP Method

The proposed CPSS-UPTDN is a kind of APTS framework, which can be implemented through the artificial system, computational experiments, parallel execution (ACP) method [28]. These modules for AI and big data processing are integrated according to the principles as Figure 3 to achieve the connection and interaction between the real and artificial systems.

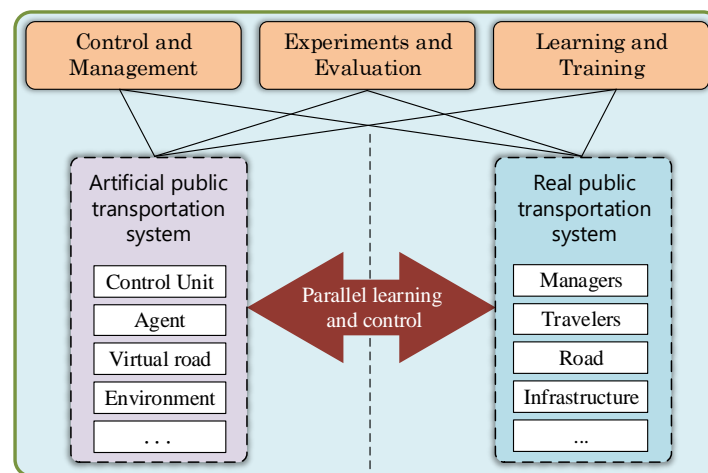


Figure 3. The ACP based Artificial System construction.

Specifically, the complex transportation system is modeled and constructed as an artificial system which is the simulation of the real system. Then, numerous computational experiments are performed on the system in the artificial society, so that the artificial system can face scenes that are less frequent or rarely appear in the real world. Aiming at the difficulty in modeling and analyzing traditional social methods for complex social objects, the combination of the artificial environment and computational experiments has changed the solution from “modeling-control” to “artificial scenario generation-control scenario response”.

The CPSS-UPTDN can be executed when the artificial system is built. The overall pipeline is described in Algorithm 1. The hierarchy of a typical computational procedure in CPSS-UPTDN framework can adopt the service patterns of cloud computing, as Figure 4 shows. Firstly, the physical resource management layer (Infrastructure as a Service, IaaS) is constructed as a foundation, which is used to access and visualize the original data. Then, the logic resource processing layer (Platform as a Service, PaaS) is used to process and analyze data. Finally, the data center operation and maintenance layer (Software as a Service, SaaS) is carried out to optimize various tasks. There are also operation management and security interfaces. The whole platform is composed of several subsystems, among which each one provides specific services of information, management and surveillance.

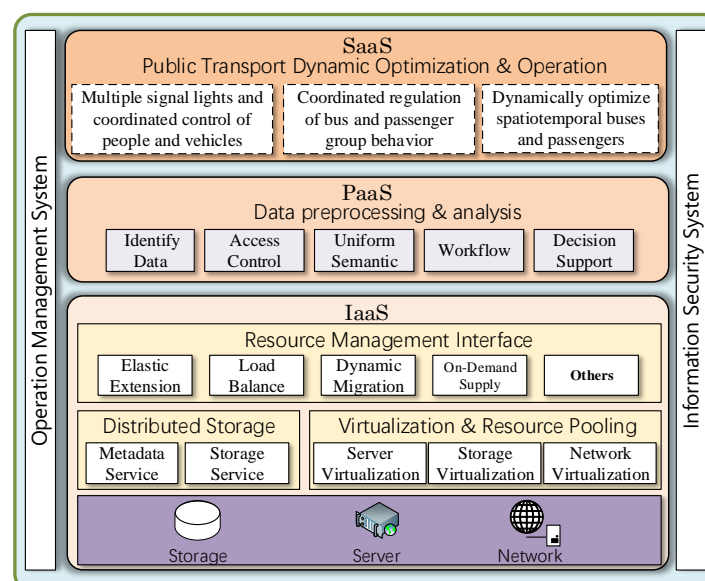


Figure 4. Hierarchy of CPSS-UPTDN computational procedure.

Algorithm 1 The pipeline of CPSS-UPTDN**Input:** Multi-source big data generated by real urban public transportation**output:** The optimal dispatching service and personalized recommendation

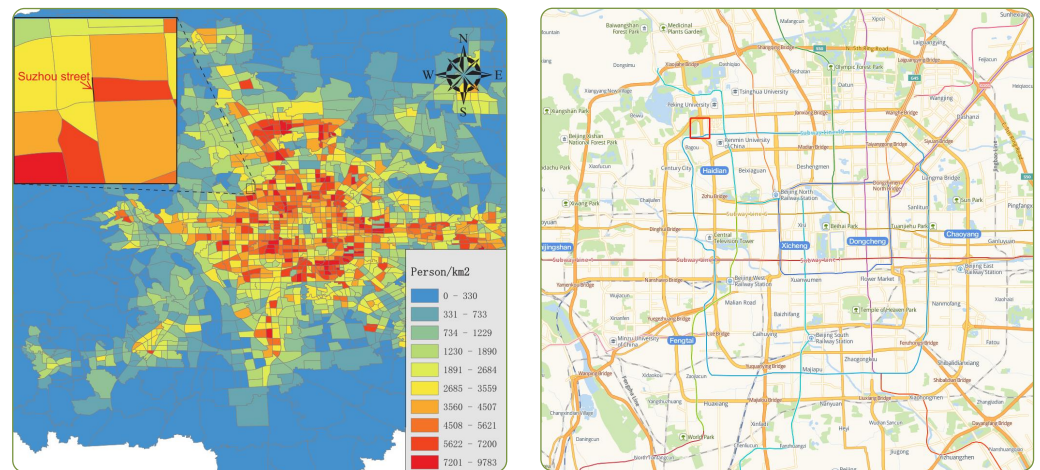
- 1: Data acquisition systems capture and store the big data.
- 2: The basic models of the urban public transportation system are established by the fusion and analysis of multi-source data.
- 3: The artificial public transport system is established by the basic model and artificial social method, then the CPSS-UPTDN platform is built.
- 4: **while** The real system generates real-time travel information **do**
- 5: Travel demand is fed back into the artificial system.
- 6: According to the computational experiments, the effective bus dispatching scheme is evaluated and verified in an artificial system.
- 7: **yeild** Output effective scheme and provide personalized service for travelers.
- 8: The real transportation system generates new multi-source traffic big data.
- 9: **if** Periodical time is up **then**
- 10: Data acquisition systems collect new traffic data.
- 11: Update the basic traffic models.
- 12: Update the artificial traffic system and platform.
- 13: **end if**
- 14: **end while**

5. A Detailed Case of CPSS-UPTDN

There are various potential applications of CPSS-UPTDN based on AI and big-data technologies. We discuss a typical case for bus dispatching based on our proposed framework in detail. By combining cellular data, a deep learning algorithm is designed to predict the road congestion extent and facilitate the follow-up dispatching in public transportation. The big data comes from physical network resources and social systems. Next, we will describe the detailed processes of dynamic dispatching for urban buses.

5.1. Acquiring Useful Information from Triple Space

We acquire real-time judgment on users' current travel demand and road conditions based on cellular data and physical traffic status data as multi-source traffic data. The cellular data comes from one of the top three mobile carriers in China. The data is stored and accessed via HDFS, in which MapReduce algorithm and "HIVE" commend are carried out for data preprocessing. To protect users' privacy, the pseudonym mechanism has been adopted during our experiments. Note that it is also worth using blurry location and encryption to avoid ethical problems in future research. In detail, the selected traffic analysis zone (TAZ) is Suzhou Street in Hai-Dian district in Beijing, China. The travel demand is extracted from the historical origin and destination of users. Urban residents living distribution is obtained from cellular data as shown in Figure 5. By using the application programming interface provided by "Amap" (<https://lbs.amap.com/>), traffic status in this region every 3 mininutes was requested by web crawlers for a period (08:00–20:00 × 10 workdays from 1 January 2019 to 23 January 2019) to make as datasets. Figure 6 shows the congestion data we acquired by web crawlers from 8:00 to 20:00 in 1 January 2019. It can be clearly seen that at around 10:00 am and 18:00, the patency rate of the road is decreased, which means that the congestion of this intersection is higher at these morning-evening rush hours.



(a) Distribution of residents from cellular data (b) Traffic map for Urban area in Beijing

Figure 5. Location and residential population distribution of the TAZ in Beijing.

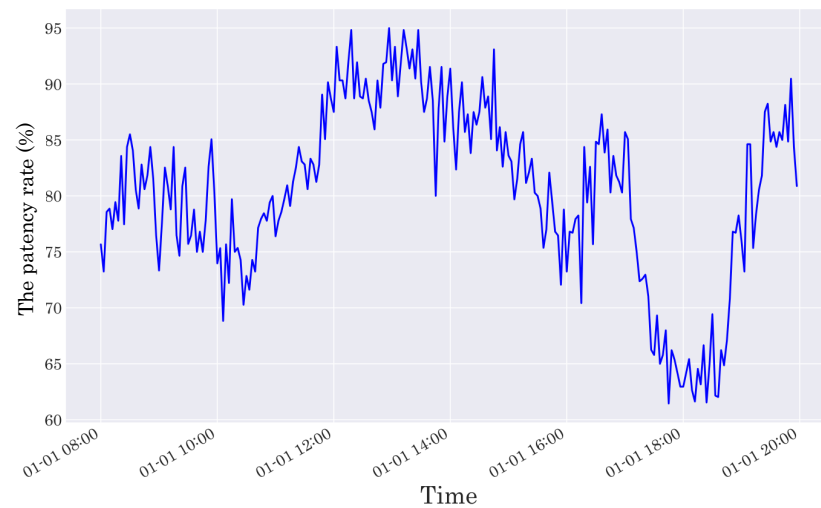


Figure 6. The patency rate in Suzhou street at 8:00–20:00, 1 January 2019.

5.2. Data Extraction and Fusion

The travel demand data and traffic status data can be unified into the same and characteristic space to realize data fusion by metric learning methods. Specifically, we divide the population into 10 categories according to the number of people with travel demand per square kilometers, so a 10-bit one-hot vector is used to encode the input of travel demand. Then, we binarize the input so that they can be thought as a vector from the Euclidean space. Then, we round up the time-stamp in the trip of users. For example, the trip at 08:20 is rounded to 08:00, and the trip in 08:40 is rounded to 09:00. Then, the travel demand at 08:00–20:00 about 12 traffic units (marked by green at the top left corner of (b)) around Suzhou street is analyzed (Figure 7). The travel demand at morning-evening rush hours is significantly increasing and high, and it is gentle and low at night.

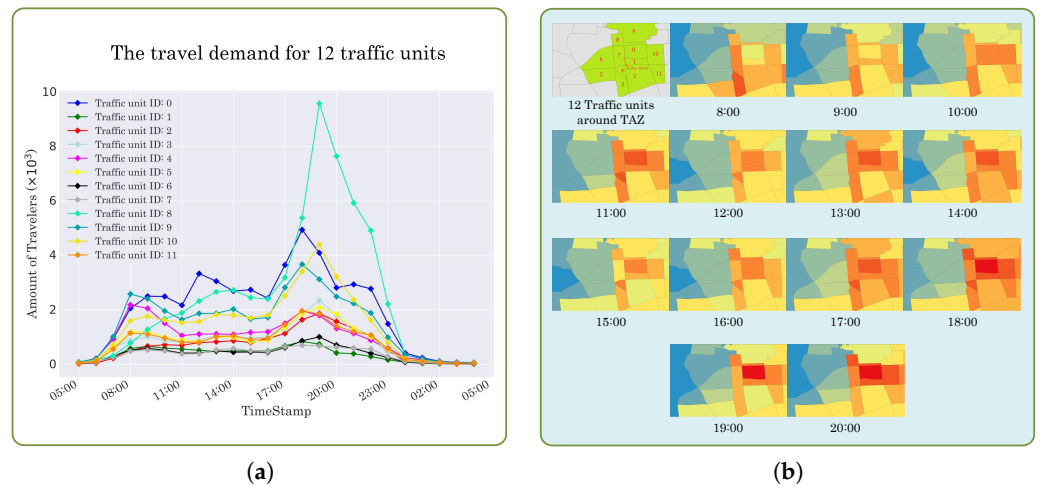


Figure 7. (a) The travel demand obtained from cellular data of 12 traffic units around TAZ and (b) divide it into 10 categories from 8:00 to 20:00.

5.3. Prediction Model Construction and Analysis

According to the processed data, we leverage the deep learning model named stacked autoencoder (SAE) to predict the traffic status and analyze the travel demand of residents, providing data support for decision-making. The model we used and its hyper-parameters are shown in Figure 8. It is composed of a multi-layer sparse autoencoder. The bottom layers are responsible for extracting the features of the data and pre-training the network. The top layer is a predictor that performs regression to predict results.

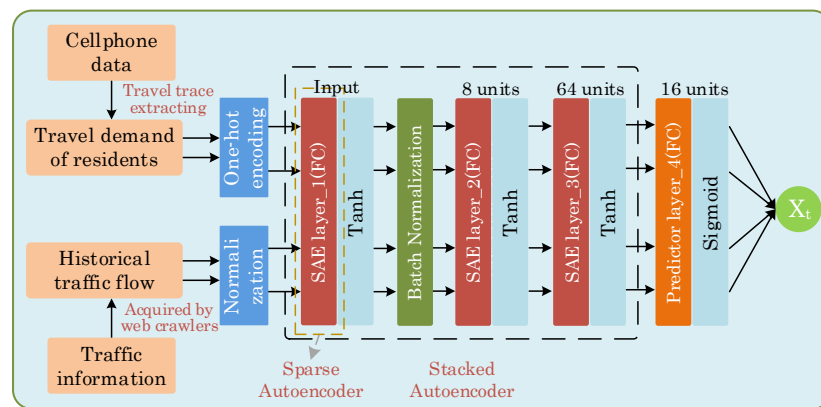


Figure 8. The structure of the SAE network.

The inputs of the model are divided into two parts, one of which receives one-hot labeled cellular data while the other receives the historical patency rate after normalization. The output of the network is the patency rate we predict. By training the network with the dataset, we can obtain an effective prediction model. Using the model with current inputs, we can analyze the crowd distribution in bus routes and predict traffic patency and residents' travel demand more accurately, and then provide the shortest and fastest traffic routes for travelers. Besides, a comparative model that does not consider travel demand is conducted, and then the impact of multi-modal data on model is analyzed.

Two commonly used performance metrics are adopted to evaluate the results:

- Mean absolute error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \tag{1}$$

- Mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}, \tag{2}$$

where n represents the number of samples, y_i is the observed values and \hat{y}_i is the predicted results.

The historical step of patency rate is set as 5 to predict the status in the next period, and the performance of the two models is demonstrated in Table 1. We conduct the experiments using Keras on a desktop with CPU (Intel i7-8700) and GPU (Nvidia GTX-1060). The results are the average of 10 independent experiments. It can be seen that the time cost is acceptable when hourly travel demand is introduced. The predicting results are shown in Figure 9.

Table 1. Performance of models and Computational cost.

Model Name	Inference Time (10^{-4} s)	Training Time (epoch/s)	MAE	MAPE%
SAE with travel demand	3.581	0.243	3.07	3.74
SAE	3.526	0.238	3.27	3.97

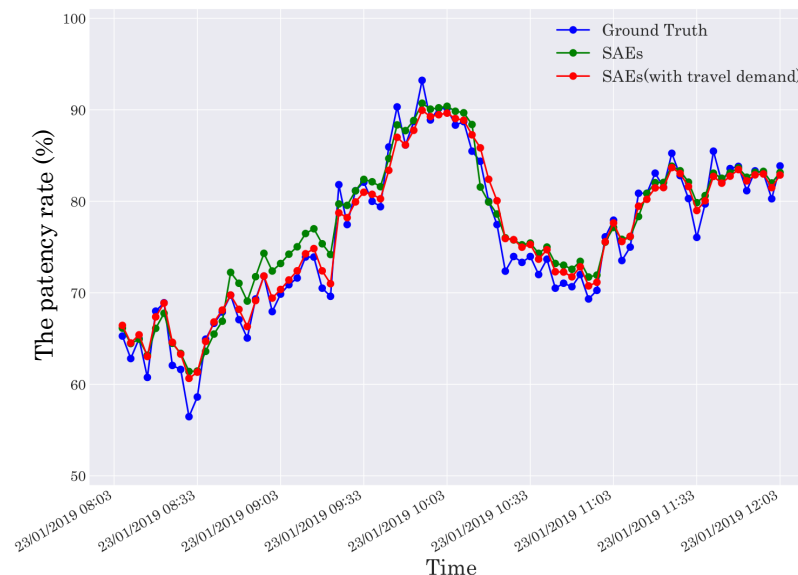


Figure 9. The comparison between SAE model with and without travel demand obtained from cellular data.

The performance of prediction improves by 6.12% and 5.79% after appending the population travel demand feature, which reflects the fluctuation of traffic patency around Suzhou street more accurately.

5.4. Parallel Execution and Output

By putting the predicted results into the artificial system and designing relevant computational experiments, we can simulate and demonstrate the bus dispatching scheme and generate possible execution rules. Then the effective dispatching scheme can be adopted as output to carry on the dynamic allocation to the bus operation. It can also help to obtain the decision-making support to transportation infrastructure construction, such as the placement of bus stations and the route planning of buses. The artificial system can be updated periodically and optimized by the feedback of the real urban public transportation system. By leveraging triple space information, parallel execution can simulate the complex status which may arise, and evaluate the effects of solutions dynamically in the artificial environment cost-effectively and safely. Thus, it can give an intuitive analysis conclusion, and provide auxiliary decision-support for the reasonable adjustment of the bus route network and, finally, other traffic infrastructure.

6. Conclusions

The urban public transportation system is not aimed at a single user, but a group of people. The volume of big data is an order of magnitude difference compared with individual travel features. After obtaining traffic big data, we still need to develop relevant AI algorithms to harness the full power of the data, and model the volatile environment, thus predicting the travel demand of metropolitan in advance.

In this paper, a universal CPSS-UPTDN framework for APTS is introduced to study the evolution behavior of public transportation, and provide support for the theoretical research of CPSS. We demonstrate key technologies for big data acquisition, the application of AI methods, and its specific implementation. Finally, the feasibility of traffic status prediction is proved in our framework through detailed experiments. However, the public system is extremely cumbersome due to its volatile dynamics and human factors. The execution on a real system needs the higher-authority permission of urban public transportation, so this module has been tested in our Labs, but has not been used in practice. As a next step, we will adopt a parallel test scheme [29] to continue our study and prepare for application. We intend to build a small simulation system and design real-world scenarios that correspond exactly to the artificial system, comparing their outputs to update the artificial system from time to time. Simultaneously, various methods can be developed to mix real scene data and virtual scene data to provide different scenarios to test our framework. With the improvement of computing power and algorithm, we believe that CPSS-UPTDN can be a novel thought to combine multi-source big data and AI methods. It still needs some time for the across-the-board implementation of a real-time and intelligent urban public transportation system.

Author Contributions: Conceptualization, G.X. and Z.L.; methodology, Z.L.; software, S.C.; validation, Z.L., H.W. and X.D.; formal analysis, H.W.; investigation, S.C. and F.Z.; resources, Y.L.; data curation, Y.L.; writing—original draft preparation, G.X. and Z.L.; writing—review and editing, Z.L.; visualization, Z.L.; supervision, G.X.; project administration, G.X.; funding acquisition, G.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by National Natural Science Foundation of China under Grants 61773381, U1909204, U1811463 & 61872365; Chinese Guangdong's S&T project (2019B1515120030); Dongguan's Innovation Talents Project (Gang Xiong).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CPSS	Cyber-physical-social system
ITS	intelligent transportation systems
UPTDN	Urban public traffic dynamic network
APTS	Advanced public transportation systems
AI	Artificial Intelligent
ACP	Artificial system, Computational experiments, Parallel execution
TAZ	Traffic analysis zone
HDFS	Hadoop Distributed File System
API	Application Programming Interface
SAE	Stacked AutoEncoder
MAE	Mean Absolute Error
MAPE	Mean Absolute Percent Error

References

- Brons, M.; Nijkamp, P.; Pels, E.; Rietveld, P. Efficiency of urban public transit: A meta analysis. *Transportation* **2005**, *32*, 1–21. [\[CrossRef\]](#)
- Aslam, J.; Lim, S.; Rus, D. Congestion-aware Traffic Routing System using sensor data. In Proceedings of the International IEEE Conference on Intelligent Transportation Systems, Anchorage, AK, USA, 16–19 September 2012.
- Lin, Y.; Wan, H.; Rui, J.; Wu, Z.; Jia, X. Inferring the Travel Purposes of Passenger Groups for Better Understanding of Passengers. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 235–243. [\[CrossRef\]](#)
- Kokkinogenis, Z.; Monteiro, N.; Rossetti, R.J.F.; Bazzan, A.L.C.; Campos, P. Policy and incentive designs evaluation: A social-oriented framework for Artificial Transportation Systems. In Proceedings of the IEEE International Conference on Intelligent Transportation Systems, Qingdao, China, 8–11 October 2014.
- Chen, X.; Wan, X.; Ding, F.; Li, Q.; McCarthy, C.; Cheng, Y.; Ran, B. Data-Driven Prediction System of Dynamic People-Flow in Large Urban Network Using Cellular Probe Data. *J. Adv. Transp.* **2019**, *2019*, 9401630. [\[CrossRef\]](#)
- Li, C.L. Better, Quicker, Together: Enabling Public Transport Service Quality Co-Monitoring through a Smartphone-Based Platform. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2016.
- Ming, Z.; Liu, X.Y.; Qiu, M.; Shen, R.; Wei, S.; Wu, M.Y. Traffic big data based path planning strategy in public vehicle systems. In Proceedings of the IEEE/ACM International Symposium on Quality of Service, Beijing, China, 20–21 June 2016.
- Derler, P.; Lee, E.A. Modeling Cyber-Physical Systems. *Proc. IEEE* **2011**, *100*, 13–28. [\[CrossRef\]](#)
- Wang, F.Y. The emergence of intelligent enterprises: From CPS to CPSS. *IEEE Intell. Syst.* **2010**, *25*, 85–88. [\[CrossRef\]](#)
- Chen, S.; Zhong, L.; Shen, D. Modeling Social Influence on Activity-Travel Behaviors Using Artificial Transportation Systems. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 1576–1581. [\[CrossRef\]](#)
- Guo, W.; Zhang, Y.; Li, L. The integration of CPS, CPSS, and ITS: A focus on data. *Tsinghua Sci. Technol.* **2015**, *20*, 327–335. [\[CrossRef\]](#)
- Hussein, D.; Park, S.; Han, S.N.; Crespi, N. Dynamic Social Structure of Things: A Contextual Approach in CPSS. *IEEE Internet Comput.* **2015**, *19*, 12–20. [\[CrossRef\]](#)
- Hu, Y.; Wang, F.Y.; Liu, X. A CPSS Approach for Emergency Evacuation in Building Fires. *IEEE Intell. Syst.* **2014**, *29*, 48–52. [\[CrossRef\]](#)
- Zhang, J.J.; Wang, F.Y.; Wang, X.; Xiong, G.; Lee, Y. Cyber-Physical-Social Systems: The State of the Art and Perspectives. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 829–840. [\[CrossRef\]](#)
- Wang, F.Y.; Zheng, N.N.; Cao, D.; Martinez, C.M.; Li, L.; Liu, T. Parallel Driving in CPSS: A Unified Approach for Transport Automation and Vehicle Intelligence. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 577–587. [\[CrossRef\]](#)
- Wang, F.Y.; Zhang, J.J. Transportation 5.0 in CPSS: Towards ACP-based society-centered intelligent transportation. In Proceedings of the IEEE International Conference on Intelligent Transportation Systems, Yokohama, Japan, 16–19 October 2018.
- Figueiras, P.; Costa, R.; Malo, P.; Bradesko, L.; Jermol, M. Knowledge base approach for developing a mobile personalized travel companion. In Proceedings of the International Conference on ITS Telecommunications, Tampere, Finland, 5–7 November 2013.
- Liu, G.; Yin, Z.; Jia, Y.; Xie, Y. Passenger flow estimation based on convolutional neural network in public transportation system. *Knowl.-Based Syst.* **2017**, *123*, 102–115. [\[CrossRef\]](#)
- Xiong, G.; Zhu, F.; Liu, X.; Dong, X.; Huang, W.; Chen, S.; Zhao, K. Cyber-physical-social System in Intelligent Transportation. *IEEE/CAA J. Autom. Sin.* **2015**, *2*, 320–333.
- Zheng, X.; Chen, W.; Wang, P.; Shen, D.; Chen, S.; Wang, X.; Zhang, Q.; Yang, L. Big Data for Social Transportation. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 620–630. [\[CrossRef\]](#)
- He, K.; Xu, Z.; Pu, W.; Deng, L.; Lai, T. Congestion Avoidance Routing Based on Large-Scale Social Signals. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2613–2626. [\[CrossRef\]](#)

22. Calastri, C.; Crastes Dit Sourd, R.; Hess, S. We want it all: Experiences from a survey seeking to capture social network structures, lifetime events and short-term travel and activity planning. *Transportation* **2020**, *47*, 175–201. [[CrossRef](#)]
23. Vinel, A.; Lyamin, N.; Isachenkov, P. Modeling of V2V Communications for C-ITS Safety Applications: A CPS Perspective. *IEEE Commun. Lett.* **2018**, *22*, 1600–1603. [[CrossRef](#)]
24. Han, S.; Cao, D.; Li, L.; Li, L.; Li, S.E.; Zheng, N.N.; Wang, F.Y. From Software-Defined Vehicles to Self-Driving Vehicles: A Report on CPSS-Based Parallel Driving. *IEEE Intell. Transp. Syst. Mag.* **2019**, *11*, 6–14. [[CrossRef](#)]
25. Alam, M.; Schiller, E.; Shu, L.; Wu, X.; Jayo, U.H. Editorial: Dependable and Real-time Vehicular Communication for Intelligent Transportation Systems (ITS). *Mob. Netw. Appl.* **2018**, *23*, 1129–1131. [[CrossRef](#)]
26. Wang, P.; Yang, L.T.; Li, J. An Edge Cloud-Assisted CPSS Framework for Smart City. *IEEE Cloud Comput.* **2018**, *5*, 37–46. [[CrossRef](#)]
27. Gang, X.; Zhu, F.; Dong, X.; Fan, H.; Teng, T. A Kind of Novel ITS based on Space-Air-Ground Big-data. *IEEE Intell. Transp. Syst. Mag.* **2016**, *8*, 10–22. [[CrossRef](#)]
28. Wang, F.Y. Parallel system methods for management and control of complex systems. *Control Desicion* **2004**, *19*, 485–489.
29. Li, L.; Wang, X.; Wang, K.; Lin, Y.; Xin, J.; Chen, L.; Xu, L.; Tian, B.; Ai, Y.; Wang, J.; et al. Parallel testing of vehicle intelligence via virtual-real interaction. *Sci. Robot.* **2019**, *4*, 1–4. [[CrossRef](#)] [[PubMed](#)]