Guest editorial: Big data-driven analytics for smart cities: technology-based insight

1. Introduction

The smart city is an overarching framework for thinking about the future in which ecological, social and economic factors are all taken into account in the quest to better human life (Khansari *et al.*, 2014; Chu *et al.*, 2021). Smart cities use heterogeneous network infrastructure, pervasive sensing devices, big data processing and smart control systems (Zhang *et al.*, 2017). Prior studies have examined the operation and management of a smart city from various perspectives, such as charging management for electric vehicles (Shuai *et al.*, 2016), car sharing and air pollution (Barnes *et al.*, 2020) and healthcare management (Xu *et al.*, 2018). However, most of them mainly rely on well-established research methodologies, such as surveys, descriptive analysis and linear regression models. Solving smart city issues is undoubtedly costly and complicated and requires us to develop novel research approaches based on empirical data to provide more appropriate solutions to build smart city plans.

In recent years, the expansion of digital infrastructures such as the Internet of Things (IoT) and information and communication technologies (ICT) has enabled the rapid proliferation of city-level big data (Batty, 2013). With the assistance of big data analytics and operations research, our study methodologies and insights can be considerably enhanced by utilizing such spontaneous, objective and vast data. Therefore, there is broad interest in academics and practice in learning how to apply big data analytics and optimization for the operation and management of a smart city from a variety of viewpoints, including urban planning and layout (Wang *et al.*, 2017, 2018), smart energy management (Barbry *et al.*, 2019), tourism (Guo *et al.*, 2014) and healthcare services (Ding *et al.*, 2019).

How do novel approaches powered by big data analytics contribute to developing a smart city with high efficiency? Motivated by this concern and aware of the potential contribution of big data management and application research to the development of smart cities, we dedicate a special issue of Industrial Data and Management Systems to data-driven analytics applied to smart city-related issues. This special issue aims to publish the insights and perspectives of academics regarding solutions to smart city issues and challenges from various perspectives, such as the resource allocation and optimization of power stations, the location and capacity for shared car parking space, transportation strategy and operation management of car sharing. Twelve papers in this special issue tackle the subject of data-driven analytics developments for smart cities from various angles. These studies can be categorized into four primary categories: the enhancement of key performance prediction models, urban planning and layout optimization, smart hospitality and tourism and demand management. In these papers, recent advancements in data-driven prediction and optimization analytics research focusing on various aspects of a smart city are discussed, such as battery storage management, hospital performance prediction, economic development in a smart city, urban rain gauge network and spatial

Generous financial support was provided by the National Natural Science Foundation of China [grant numbers 71872061] and the Key Program of NSFC-FRQSC Joint Project (NSFC No. 72061127002, FRQSC No. 295837). This work is partially supported by Shenzhen Humanities & Social Sciences Key Research Bases.

Industrial Management & Data Systems
Vol. 122 No. 10, 2022
pp. 2145-2150

0263-5577
DOI 10.1108/IMDS-10-2022-813

structure, location-routing problem, sharing networking matching, artificial intelligence enabled robots, tourism destination and bus travel management and parallel batch scheduling.

We believe that this special issue, with its diverse assortment of papers, achieves our goal of presenting studies that ingeniously push the boundaries of this area.

2. Models for predicting emerging smarty city issues

Frequently, addressing issues related to improving citizens' quality of life requires the development of novel approaches that bring fresh insight. A smart city is a potential solution to the issues resulting from the unparalleled rate of urbanization. However, the rising availability of big data poses a barrier to the smart city transformation process. With large data sets, conventional statistical and economical methods may not perform effectively. Therefore, a possible path forward is utilizing powerful machine learning methods for city-wide data analysis.

This first section offers three papers with innovative approaches, the first of which explores models for estimating hospital performance metrics and the second of which combines various ways to estimate economic performance in a smart city. The third paper develops a learned embedding model for crime event prediction.

"Making the hospital smart: using a deep long short-term memory model to predict hospital performance metrics" by Qiong Jia, Ying Zhu, Rui Xu, Yubin Zhang and Yihua Zhao constructs a deep long short-term memory (DLSTM) model to forecast time-series hospital performance measurements, such as daily patient visits. With smaller model performance indicators, such as root mean square percentage errors, the static DLSTM approach outperforms seasonal autoregressive integrated moving averages (SARIMA), single- and multiple-RNN, deep gated recurrent units (DGRU), traditional long short-term memory (LSTM) and dynamic DLSTM. In particular, static DLSTM surpasses all other models in forecasting the number of daily patient visits, medical examinations and prescriptions. Based on the proposed model, hospitals can make more accurate forecasts of outpatient visits, medical examinations, and prescriptions, affecting their construction plans and improving the efficacy with which they manage pertinent data.

The second paper, "Mixed-frequency data-driven forecasting the important economies' performance in a smart city: a novel RUMIDAS-SVR model", authored by Weiqing Wang, Zengbin Zhang, Liukai Wang, Xiaobo Zhang and Zhenyu Zhang, examines the predictive power of mixed-data sampling for important indicators in the real world. This research applies reverse unrestricted mixed-data sampling (RUMIDAS) to support vector regression (SVR) to construct a novel RUMIDAS-SVR model. Using the mixed-frequency consumer price index (CPI), producer price index (PPI) and consumer confidence index (CCI) as predictors, the authors then apply the novel RUMIDAS-SVR model to forecast the development performance of all high-tech listed companies, a significant sector of the economy that reflects the potential and dynamism of urban economic development. Empirical findings indicate that the developed RUMIDAS-SVR is superior to competing models in terms of mean absolute error (MAE) and root-mean-squared error (RMSE). The study provides theoretical and practical insights to help smart city policymakers create a favorable macroeconomic environment, such as controlling inflation or stabilizing prices.

The third paper in the area of prediction models, "A multi-dimensional city data embedding model for improving predictive analytics and urban operations", is written by Zhe Jing, Yan Luo, Xiaotong Li and Xin Xu."

The research presents a model for learning region embeddings. By encoding discrete variables as continuous vectors that encode the meaning of a location, the learned embedding can be used for more precise prediction. They employ random walk and skip-gram to learn

embedding and to update the initial embedding generated by Graph Convolutional Network (GCN). The researchers then apply the suggested model to a real-world data set from Manhattan, New York, and utilize the learned embedding to forecast crime events. Their findings demonstrate that the suggested model may learn more aspects of city data, lower the amount of computation and leverage distributed computing to develop and transform smart cities.

3. Urban planning and layout optimization

A key development area in smart city development is the application of data and technology to improve urban planning and layout. In this special issue, we have three papers examining this topic from the different perspectives of city spatial structure innovation, integrated reverse logistics planning and rain gauge network design.

The first paper in this area, "City spatial structure and smart city innovation: the case of China", is authored by Yongtai Chen, Rui Li, Enyu Zeng and Pengfei Li. This study explores the relevance of city spatial layout for smart city innovation from the perspective of agglomeration externalities, as well as the diversity in innovation across different geographic regions and population scales of cities. Centralization and concentration indexes are created to comprehend the spatial structure of 286 (prefecture-level) Chinese cities. The relationship between the geographic structure and innovation capability of smart cities is explored using panel data. The more concentrated and uniformly dispersed the urban population is, i.e. the more the city's spatial organization leans toward weak-monocentricity, the higher the level of creativity in smart cities, according to their research. This result is more applicable to eastern, central and smaller cities and those in the Midwest. In contrast to the bulk of past research, which analyzed city spatial structure using a single dimension of "the proportion of people in sub-centers to the population of all central sites," this study develops geographic centralization and spatial concentration. We hope this study can assist the construction of smart cities from the perspective of the development model of city spatial structure.

The second paper in this area, "A multi-cycle and multi-echelon location-routing problem for integrated reverse logistics", is written by Xiaofeng Xu, Wenzhi Liu, Mingyue Jiang and Ziru Lin. The rapid growth of smart cities and green logistics has encouraged a great deal of research on reverse logistics. The diversity of data provides the opportunity for a novel study on the location-routing problem (LRP) within reverse logistics. There is a paradox between the total cost of the firm and the negative social utility, which indicates that reducing the negative social utility costs a certain amount of money. The goal of this research is to use panel data to aid in the investigation of the multi-cycle and multi-echelon location-routing problem in reverse logistics network (MCME-LRP-RLN), hence decreasing the cost of enterprise facility location. Initially, a negative utility objective function is derived from panel data and integrated into a multi-cycle and multi-echelon location-routing model incorporating reverse logistics. Then, the particle swarm optimization (PSO) method was added to the crossover of the multi-objective immune genetic algorithm (MOIGA) to tackle multi-objective, large-scale problems, hence enhancing the convergence speed and performance of the algorithm. This paper is a ground-breaking investigation of the MCME-LRP-RLN problem and blends data analytic approaches with operations research modeling. The study concludes with helpful logistics planning management advice based on its findings.

The third paper in this area, "Optimal rain gauge network to reduce rainfall impacts on urban mobility – a spatial sensitivity analysis", is written by Felipe de Oliveira Simoyama, Lívia Rodrigues Tomás, Felipe Matheus Pinto, Luiz Leduino Salles-Neto and Leonardo Bacelar Lima Santos.

A sustainable, intelligent transportation system must not only reduce the environmental impact of transportation but also mitigate the effects of natural disasters on transportation.

A well-distributed collection of rain gauges is required for monitoring services in smart cities. Based on a database of actual rainfall events in Brazil, the authors of this paper examine ways to improve the rain gauge network to lessen rainfall's influence on urban mobility. They propose an assessment of resilience that includes geographical location disturbances and provide a formulation for the maximal coverage location. In this study, the robustness of the objective function is 99.99%. The robustness of the number of covered demand points is 88.93%, and the frequency connected with each candidate ranges from 11.71 to 69.48%. The suggested method can help policymakers create an optimal and robust network of rain gauges capable of addressing uncertainties by selecting ideal sites without requiring additional instruments. This result is consistent with, i.e. that for certain regions, the configuration of the rain gauge network is more critical than its size or density.

4. Smart hospitality and tourism

Understanding and managing hospitality and tourism issues is another important aspect of building a more sustainable smart city. This special issue has three papers in this area: identifying more efficient timetables for city buses and understanding the psychological demands and interplay of different mobile social applications for users.

The first paper, "Data-driven bus timetabling with spatial-temporal travel time", is written Xiang Li, Ming Yang, Hongguang Ma and Kaitao (Stella) Yu.

With the advancement of digital technology and big data analytics capabilities in the bus industry, practitioners prefer to generate deterministic travel time based on the on-board GPS data under the maximum probability rule and mean value rule, which simplifies the optimization procedure but performs poorly in the timetabling practice due to the loss of uncertainty on travel time. This study proposes a GPS-data-driven approach to bus scheduling that considers the spatial-temporal characteristic of travel time. A scenario-based robust scheduling approach is presented to maximize the predicted profit of the bus carrier. The authors provide a set of binary variables to convert the resilient model into an integer linear programming model and speed up the solving process by compressing the solution space so that CPLEX can effectively solve the optimal timetable. Case studies based on bus line 628 in Beijing are provided to demonstrate the effectiveness of the suggested methodology. The results indicate that: (1) the scenario-based robust model could increase expected profits by 15.80% when compared to the maximum probability model; (2) the scenario-based robust model could increase expected profits by 30.74% when compared to the mean value model; and (3) the solution space compression strategy could effectively reduce computing time by 97%. This study enhances the practicability and efficiency of timetables and demonstrates the significance of big data analytics in enhancing the administration of public transportation operations.

This second paper is titled "Strategic technological determinant in smart destinations: obtaining an automatic classification of the destination's quality." It was written by Sergio Díaz-González, Jesus M. Torres, Eduardo Parra-López and Rosa M. Aguilar. This study designs and applies a methodology for identifying and quantifying the important elements of a municipal tourist attraction. The authors construct numerous key performance indicators (KPIs) for a municipal STD based on focus groups with stakeholders from the Spanish tourist municipality of Puerto de la Cruz tourism business. Similarly, the authors identify the technology required to collect, handle and represent the data and the mechanism for measuring the quality of the STD using the analytic hierarchy process (AHP). This research outlines a theoretical strategy for enhancing STD management.

The third paper in this area, "Artificial intelligence enabled robots for stay experience in the hospitality industry in a smart city", is written by Shivam Gupta, Sachin Modgil, Choong-Ki Lee, Minsook Cho and Yaena Park.

This study examines and investigates artificial intelligence (AI)-enabled robots in the hospitality industry to improve the guest experience in smart cities. Through semi-structured interviews, this study shows that deploying AI-enabled robots in the hospitality business of a smart city facilitates automation, information collecting, personalization and seamless service. In addition, a back-and-forth mapping mechanism based on epistemological principles leads to formulating four propositions that constitute a research framework. In a smart city, the management of the hotel industry can deploy AI-enabled robots to improve and automate procedures, as well as provide enhanced personalization to improve the guest experience.

5. Supply and demand management

The fourth topic covered in this special issue is supply and demand management, which includes a paper on sharing networks that ensures sustainable resource allocation, a study on analytics for improving intelligent scheduling, and a paper on the energy supply system.

The first paper, "A branch-and-price algorithm for robust parallel batch scheduling problem with uncertain size," by Ting Wang, Xiaoling Shao and Xue Yan, investigates a parallel batch scheduling problem (PBSP) with uncertain job size, which is crucial for realizing the flexibility of product production and mass customization of personalized goods. For the purpose of describing this issue, they suggest a robust formulation in which the work size is determined by budget-constrained assistance. The authors then present an exact algorithm based on the branch-and-price framework for achieving the robust solution to PBSP, where the pricing subproblem can be reduced to a robust shortest path problem with resource limitations (R-SPPRC). The robust subproblem is transformed into a deterministic mixed integer programming (MIP) by duality. From the MIP, they generate a series of deterministic shortest route problems with resource constraints (SPPRCs), for which they provide an effective label-setting algorithm with a strong dominance rule. Their findings demonstrate the efficacy of our algorithm and the significance of incorporating uncertainty into the PBSP.

"Sharing network features analysis and dispatching strategy design" is the second paper in this field, written by Tong Lv, Lefeng, Shi and Weijun He. A crucial task for a sharing business is dynamically distributing shared items to balance demand and supply at various sharing locations in a sharing network. This research provides a new algorithm for constructing a highly efficient dispatch strategy based on the findings of sharing network features. They discover that the sharing stations of a common sharing network may be divided into six unique types based on their profit dynamics; a sharing network comprising different combinations of sharing stations will exhibit varied profit features. Taking into account the features, a system is devised and proven to direct the dynamic dispatch of shared items. Due to the fact that the proposed technique takes into account the interaction characteristics between sharing locations in a sharing network, its computation speeds and convergence efficacy to the global optimal scheme are superior to those of comparable studies. It is more suitable for sharing businesses seeking greater time efficiency.

"An investigation of battery storage operating strategies in the context of smart cities" by Xing Yao, Shao-Chao Ma, Ying Fan, Lei Zhu and Bin Su develops a simulation-based economic power dispatch model capable of quantifying the effects of various profit orientations of energy storage on a city-level power system. They then apply the model to a case study of Jiangsu, China. Simulated and discussed are the cost and emissions of the power system under two energy storage operation techniques. Their findings have major significance for policymakers and investors in building the energy system and assessing the potential cost and emission impact of energy storage on power systems with various capacity mixes.

6. Conclusion

We appreciate the support of the editors and reviewers for this special issue. Without them, this special issue would not be possible. Finally, we thank all the contributors to this special issue. The issue was a memorable experience due to the overwhelming response to our call. We hope you enjoy reading the included papers.

Yue Guo
Southern University of Science and Technology, Shenzhen, China

Lean Yu
School of Economics and Management, University of the Chinese Academy of Sciences, Beijing, China

Yichuan Ding
Desautels Faculty of Management, McGill University, Montreal, Canada, and

Leandro C. Coelho
Université Laval, Québec, Canada

References

Barbry, A., Anjos, M.F., Delage, E. and Schell, K.R. (2019), "Robust self-scheduling of a price-maker energy storage facility in the New York electricity market", *Energy Economics*, Vol. 78, pp. 629-646.

Barnes, S.J., Guo, Y. and Borgo, R. (2020), "Sharing the air: transient impacts of ride-hailing introduction on pollution in China", *Transportation Research Part D: Transport and Environment*, Vol. 86, 102434.

Batty, M. (2013), "Big data, smart cities and city planning", *Dialogues in Human Geography*, Vol. 3 No. 3, pp. 274-279.

Chu, Z., Cheng, M. and Yu, N.N. (2021), "A smart city is a less polluted city", *Technological Forecasting and Social Change*, Vol. 172, 121037.

Ding, Y., Park, E., Nagarajan, M. and Grafstein, E. (2019), "Patient prioritization in emergency department triage systems: an empirical study of the Canadian triage and acuity scale (CTAS)", *Manufacturing and Service Operations Management*, Vol. 21 No. 4, pp. 723-741.

Guo, Y., Liu, H. and Chai, Y. (2014), "The embedding convergence of smart cities and tourism internet of things in China: an advance perspective", *Advances in Hospitality and Tourism Research (AHTR)*, Vol. 2 No. 1, pp. 54-69.

Khansari, N., Mostashari, A. and Mansouri, M. (2014), "Impacting sustainable behavior and planning in smart city", *International Journal of Sustainable Land Use and Urban Planning*, Vol. 1 No. 2.

Shuai, W., Maillé, P. and Pelov, A. (2016), "Charging electric vehicles in the smart city: a survey of economy-driven approaches", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17 No. 8, pp. 2089-2106.

Wang, K., Xia, W. and Zhang, A. (2017), "Should China further expand its high-speed rail network? Consider the low-cost carrier factor", *Transportation Research Part A: Policy and Practice*, Vol. 100, pp. 105-120.

Wang, K., Xia, W., Zhang, A. and Zhang, Q. (2018), "Effects of train speed on airline demand and price: theory and empirical evidence from a natural experiment", *Transportation Research Part B: Methodological*, Vol. 114, pp. 99-130.

Xu, B., Li, L., Hu, D., Wu, B., Ye, C. and Cai, H. (2018), "Healthcare data analysis system for regional medical union in smart city", *Journal of Management Analytics*, Vol. 5 No. 4, pp. 334-349.

Zhang, K., Ni, J., Yang, K., Liang, X., Ren, J. and Shen, X.S. (2017), "Security and privacy in smart city applications: challenges and solutions", *IEEE Communications Magazine*, Vol. 55 No. 1, pp. 122-129.