ACM TIST Special Issue on Urban Intelligence

Past decades saw the rapid development of cities and the boom of urban population, but also the rise of many urban issues, such as traffic congestion, energy shortage, and pollution. The situation is continuing: as predicted by the United Nations, the world's urban population will add another 2.5 billion by 2050, an increase of 66% over today's total population. While the pressure on resources is unprecedented, increasing volume and diversity of data is at the same time being generated and collected with the help of new technologies. This calls for the integration of advanced information and computational technologies to develop intelligent solutions for urban issues. Indeed, there has been a variety of research in the AI community, pioneering in applying AI and data science to the practice of urban computing [1, 2]. These include the development of smart communities, smart home automation, intelligent transport systems, smart vehicle sharing systems, data-driven routing systems, smart grids, and smart energy solutions.

One important branch of research on urban intelligence is a model-based approach (or AI-based approach) in which researchers are looking at various decision-making issues (e.g., planning, scheduling) in the urban domain, assuming the knowledge about the model is available or can be estimated. One example is to provide optimal policies for the government in consideration of people's strategic behavior such as optimal pricing of taxi systems [3], optimal placement and management of EV charging stations [4–5], and disaster response [6]. It is often very challenging to solve such decision-making problems due to the existence of many complex interdependent factors, many strategic players (e.g., people) with conflicting preferences, multi-level optimization, uncertainty about the environment, sequential decision making, and large (even infinite) strategy space.

Another important branch of research focuses on urban intelligence—a data-driven approach in which researchers use real-world urban data to discover patterns, model behaviors, and predict the future. For example, patterns detected from human mobility data can be used to understand personal interests [7], discover urban functions [8], infer social relationships [9], and estimate air pollution [10]. The urban data cover a wide range of datasets collected from different sources, e.g., taxi trips, mobile phone locations, geotagged tweets, Points-of-Interest (POIs), household energy consumption, and sensor data from cyber physical systems. While each dataset has its own characteristics, there is an increasing trend on how to fuse the data collected from different sources and discover the patterns in a more holistic way [11], e.g., how to integrate trajectories with surrounding contexts to semantically understand human mobility [7] and how to utilize cross-domain data to estimate air quality [10].

The objective of this special issue is to highlight leading work in urban intelligence, to identify challenges, and to explore future topics in this area. The call for papers of this special issue was perceived very positively with 73 regular submissions, and 15 papers were finally accepted through a rigorous review process. The acceptance rate was therefore 20%. All articles are briefly presented in the following.

AI FOR URBAN INTELLIGENCE

Orienteering Problems (OPs) are used to model many routing and trip planning problems. OPs are a variant of the well-known traveling salesman problem where the goal is to compute the highest

2157-6904/2017/11-ART23
https://doi.org/10.1145/3154942

reward path that includes a subset of vertices and has an overall travel time less than a specified deadline. However, the applicability of OPs is limited due to the assumption of deterministic and static travel times. The article titled "Risk-Sensitive Stochastic Orienteering Problems for Trip Optimization in Urban Environments" extends Stochastic Orienteering Problems (SOPs) to Dynamic SOPs (DSOPs), which allows for time-dependent travel times. The authors introduce a new objective criterion for SOPs and DSOPs to represent a percentile measure of risk and provide non-linear optimization formulations along with their linear equivalents for solving the risk-sensitive SOPs and DSOPs. The authors also provide a local search mechanism for solving the risk-sensitive SOPs and DSOPs as well as results on existing benchmark problems and a real-world theme park trip-planning problem.

Arriving on time and total travel time are two important properties for vehicle routing. Existing route guidance approaches always consider them independently because they may conflict with each other. The paper titled "A Multiagent-Based Approach for Vehicle Routing by Considering Both Arriving on Time and Total Travel Time" develops a semi-decentralized multiagent-based vehicle routing approach where vehicle agents follow the local route guidance by infrastructure agents at each intersection, and infrastructure agents perform the route guidance by solving a route assignment problem. It integrates the two properties by expressing them as two objective terms of the route assignment problem. Regarding arriving on time, it is formulated based on the probability tail model, which aims to maximize the probability of reaching destination before deadline. Regarding total travel time, it is formulated as a weighted quadratic term, which aims to minimize the expected travel time from the current location to the destination based on the potential route assignment. The weight for total travel time is designed to be comparatively large if the deadline is loose. Experimental results on real road networks justify its ability to increase the average probability of arriving on time, reduce total travel time, and enhance the overall routing performance.

The article titled "Scalable Urban Mobile Crowdsourcing: Handling Uncertainty in Worker Movement" investigates effective ways of utilizing crowdworkers in providing various urban services. The authors design a task recommendation platform to match tasks to crowdworkers based on workers' historical trajectories and time budget limits, thus making recommendations personal and efficient. To address the challenge of handling crowdworker's trajectory uncertainties, this paper explicitly allows multiple routine routes to be probabilistically associated with each worker. This problem was formulated as an integer linear program whose goal is to maximize the expected total utility achieved by all workers. The authors further exploit the separable structures of the formulation and apply the Lagrangian relaxation technique to scale up computation. Numerical experimental results show that the proposed approach can find significantly better solutions than existing approaches.

The article titled "Simulating Urban Pedestrian Crowds of Different Cultures" reports on an agent-based model of urban pedestrian crowds where culture is explicitly modeled. The authors first extend an established agent-based social agent model, inspired by social psychology, to account for individual cultural attributes discussed in social science literature. The authors then embed the model in a simulation of pedestrians, and explore the resulting macro level crowd behaviors, such as pedestrian flow, lane changes rate, etc. The model was validated by quantitatively comparing the simulation results to the pedestrian dynamics in movies of human crowds in five different countries: Iraq, Israel, England, Canada, and France.

The article titled "A Comfort-Based Approach to Smart Heating and Air Conditioning" addresses the interrelated challenges of predicting user comfort and using this to reduce energy consumption in smart Heating, Ventilation and Air Conditioning (HVAC) systems. The authors propose the Bayesian Comfort Model (BCM) which uses a Bayesian network to learn from a user's feedback, allowing it to adapt to the users' individual preferences over time. The authors further propose an alternative to the ASHRAE 7-point scale used to assess user comfort. Using this model, the authors create an optimal HVAC control algorithm that minimizes energy consumption while preserving user comfort. Empirical evaluation shows that the proposed model schieved significantly better results than current models.

DATA SCIENCE FOR URBAN INTELLIGENCE

Understanding human mobility data is one of the key questions to enable urban intelligence. There are three papers in this special issue working on modeling, annotating, and predicting human mobility data. The article titled "Real-Time Human Mobility Modeling with Multi-View Learning" proposes to study different views of human mobility, which can capture different information and complement each other, for example, the cellphone view can describe some private-vehicle passengers, whereas the transportation view can capture some inactive cellphone users. Authors test their method on extremely large datasets in the Chinese city Shenzhen with cellphone data and transportation data including taxis, buses, and subways.

The article titled "Spotting Trip Purposes from Taxi Trajectories: A General Probabilistic Model" focuses on interpreting the purpose of taxi trips given the pick-up and drop-off locations by effectively using POI information of origin and destination regions, and the transitions between regions.

"Personalized Air Travel Prediction: A Multi-Factor Perspective" models and predicts human mobility in air travel. They propose a method to predict where a customer will fly to and which airline carrier to fly with, by considering multiple complex factors of the market situation and individual characteristics of customers.

While mining human mobility data is important for urban intelligence, how to assess privacy risk while we are acquiring large-scale human mobility is another critical issue. "A data mining approach to assess privacy risk in human mobility data" provides a fast and flexible approach to estimate privacy risk by training classifiers to capture the relation between individual mobility patterns and the level of privacy risk of individuals.

Two papers study how to better understand consumers' behaviors. The article titled "Using Online Geotagged and Crowdsourced Data to Understand Human Offline Behavior in the City: An Economic Perspective" studies how the features extracted from the digitized and crowdsourced user behavior are informative in inferring local demand. They use a dataset of restaurant bookings from Open Table (local demand) and analyze it by correlating with several online sources including local neighborhood information, street events and construction information, Foursquare check-ins data, and geotagged tweets.

The article titled "Social Bridges in Urban Purchase Behavior" argues that people who live in different communities but work at close-by locations could act as "social bridges" between the respective communities, and they are correlated with similarity in community purchase bahavior. They provide empirical evidence by studying millions of credit card transaction records for tens of thousands of individuals and show that community is a much stronger indicator of similarity in their purchase behavior than traditionally considered factors such as income and socio-demographic variables.

Efficient anomaly detection from temporal urban data can report events in a timely fashion. The article "GeoBurst+: Effective and Real-Time Local Event Detection in Geotagged Tweet Streams" effectively integrates three different data types—location, time, and text—by considering the different representations of those data types and the complicated cross-modal interactions among them. And the article "illiad: Intelligent Invariant and Anomaly Detection in Cyber Physical Systems" combines model-based and data-driven approaches to analyze the time series data generated from cyber physical systems.

Understanding energy consumption is important for energy conservation in the city. "Energy Usage Behavior Modeling in Energy Disaggregation via Hawkes Processes" studies the energy disaggregation problem, which takes a whole home electricity signal and decomposes it into its component appliances. Authors propose to model the relationship between the energy usages of various appliances by householders across different time slots.

Lastly, an interesting article in this special issue studies how to effectively collect spatial data from the crowd. "A Real-Time Framework for Task Assignment in Hyperlocal Spatial Crowdsourcing" proposes a novel spatial crowdsourcing platform that engages individuals in the act of collecting various types of spatial data such as weather conditions, air pollution, and noise levels. They deal with the challenge of how to maximize the number of assigned tasks under a budget constraint, despite the dynamic arrivals of workers and tasks.

Bo An
Nanyang Technological University, boan@ntu.edu.sg

Nick Jennings
Imperial College, n.jennings@imperial.ac.uk

Zhenhui Jessie Li
Pennsylvania State University, jessieli@ist.psu.edu

Guest Editors

REFERENCES

[1] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. 2014. Urban computing: Concepts, methodologies, and applications. *ACM Trans. Int. Syst. Technol.* 5, 3, 38.

[2] Urban Computing. https://www.microsoft.com/en-us/research/project/urban-computing/.

[3] Jiarui Gan and Bo An. 2017. Game theoretic considerations for optimizing efficiency of taxi systems. *IEEE Int. Syst.* 32, 3, 46–52, 2017.

[4] Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Soh Yeng Chai. Optimal pricing for efficient electric vehicle charging station management. In *Proceedings of the 15th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'16)*. 749–757.

[5] Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Ana Bazzan. 2015. Optimal electric vehicle charging station placement. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*. 2662–2668.

[6] Sarvapali D. Ramchurn, Trung Dong Huynh, Feng Wu, Yuki Ikuno, Jack Flann, Luc Moreau, Joel E. Fischer, Wenchao Jiang, Tom Rodden, Edwin Simpson, Steven Reece, Stephen J. Roberts, and Nicholas R. Jennings. 2016. A disaster response system based on human-agent collectives. *J. Artif. Int. Res.* 57 (2016), 661–708.

[7] Fei Wu and Zhenhui Li. 2016. Where did you go: Personalized annotation of mobility records. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management.* ACM, 589–598.

[8] Jing Yuan, Yu Zheng, and Xing Xie. 2012. Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 186–194.

[9] Hongjian Wang, Zhenhui Li, and Wang-Chien Lee. 2014. "PGT: Measuring mobility relationship using personal, global and temporal factors. In *Proceedings of the 2014 IEEE International Conference on Data Mining (ICDM'14)*. IEEE, 570–579.

[10] Yu Zheng, Furui Liu, and Hsun-Ping Hsieh. 2013. U-Air: When urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1436–1444.

[11] Yu Zheng. 2015. Methodologies for cross-domain data fusion: An overview. *IEEE Trans. Big Data* 1, 1 (2015), 16–34.