A Federated Platform Enabling a Systematic Collaboration Among Devices, Data and Functions for Smart Mobility

You, Linlin; Danaf, Mazen; Zhao, Fang; Guan, Jinping; Azevedo, Carlos Lima; Atasoy, Bilge; Ben-Akiva, Moshe

DOI
10.1109/TITS.2023.3236991

Publication date
2023

Document Version
Final published version

Published in
IEEE Transactions on Intelligent Transportation Systems

A Federated Platform Enabling a Systematic Collaboration Among Devices, Data and Functions for Smart Mobility

Linlin You, *Member, IEEE*, Mazen Danaf, Fang Zhao, Jinping Guan, Carlos Lima Azevedo, Bilge Atasoy, *Member, IEEE*, and Moshe Ben-Akiva

***Abstract*—Through the vast adoption and application of emerging technologies, the intelligence and autonomy of smart mobility can be substantially elevated to address more diversified demands and supplies. Along with this trend, a systematic collaboration among three essential elements of smart mobility services, namely devices, data and functions, is being studied to comprehensively break down the intrinsic barriers that existed in current solutions, to support the integration of connectable devices, the fusion of heterogeneous data, the composability of reusable functions, and the flexibility in their cooperations. To enable such a collaboration, this paper proposes a federated platform, called Future Mobility Sensing Advisor (FMSA), which can 1) manage the three elements through standardized interfaces separately and uniformly; 2) create a fully connected knowledge graph to orchestrate the three elements efficiently and effectively; 3) support the client-server interaction in centralized and federated modes to handle service requests and edge resources with various availability and accessibilities jointly and adaptively; and 4) accommodate various mobility services to foster harmonious and sustainable mobility tenderly and invisibly. Moreover, the efficiency and effectiveness of the platform are also tested through a performance evaluation, and a pilot supported at the Great Boston Area, respectively. As a result, it shows that FMSA can 1) achieve high performance by using the two interaction modes selectively, and 2) renovate smart mobility towards sustainability through personalized services that can measure user preferences and system objectives mutually.**

***Index Terms*—Smart mobility, systematic collaboration, federated platform, service orchestration, federated computing.**

I. Introduction

WITH the development of emerging technologies, e.g., IoT (Internet of Things), Big Data, AI (Artificial Intelligence), Cloud, Fog and Edge Computing, etc., smart mobility supported by the Intelligent Transportation System (ITS) can significantly improve the quality of daily lives of the inhabitants, and the livability and sustainability of the city through diverse intelligent and autonomous services [1], [2], [3], [4]. As illustrated in Figure 1, the "Intelligence" of smart mobility services and the "Awareness" of the public can be continuously elevated by implementing a systematic collaboration among three elements, i.e., devices, data, and functions, to bridge mobility demands with system supplies efficiently and effectively [5], [6], [7]. For instance, in shared mobility services, cruising vehicles mounted with personal devices of drivers are connected and managed as a service cluster to better support the actual mobility needs of commuters through collective awareness supported by routing and matching functions fine-tuned by information derived from user preferences, travel histories, real-time traffic conditions, etc.

Moreover, since the scale of modern smart mobility systems and services is growing exponentially together with the number of users and devices to be served and supported [8], the complexity in their full lifecycles (i.e., from value proposition to actual application) will also increase significantly, leading to high development and maintenance costs as well as compromised service quality and user experience. Such an impact may pose a dilemma between functional integrity and budgetary constraints for system operators and service providers, and also prevent the adaptation and deployment of concrete solutions due to their low customizability and elasticity [7], [9]. Regarding diverse systems and services coexisting and working simultaneously to assist each other

Manuscript received 28 January 2022; revised 6 September 2022 and 20 November 2022; accepted 6 January 2023. Date of publication 23 January 2023; date of current version 29 March 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62002398, in part by the Collaborative Innovation Center for Transportation of Guangzhou under Grant 202206010056, in part by the National Research Foundation Singapore through the Future Urban Mobility IRG Research Program of Singapore-MIT Alliance for Research and Technology, and in part by the Tripod (Sustainable Travel Incentives with Prediction, Optimization and Personalization) Project under the U.S. Department of Energy Advanced Research Projects Agency-6 Energy (ARPA-E) Traveler Response Architecture using Novel Signaling for Network Efficiency in Transportation (TRANSNET) program under Award DE-AR0000611. The Associate Editor for this article was A. Bucchiarone. *(Corresponding authors: Fang Zhao; Jinping Guan.)*

Linlin You is with the School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 518107, China, and also with the Intelligent Transportation Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Mazen Danaf and Moshe Ben-Akiva are with the Intelligent Transportation Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Fang Zhao is with the Future Urban Mobility IRG, Singapore-MIT Alliance for Research and Technology Centre, Singapore 138602 (e-mail: fang.zhao@smart.mit.edu).

Jinping Guan is with the School of Architecture, Harbin Institute of Technology, Shenzhen 518055, China (e-mail: jinpingg@mit.edu).

Carlos Lima Azevedo is with the Department of Management Engineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark, and also with the Intelligent Transportation Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Bilge Atasoy is with the Transportation Engineering and Logistics, Delft University of Technology, 2628 Delft, The Netherlands, and also with the Intelligent Transportation Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/TITS.2023.3236991

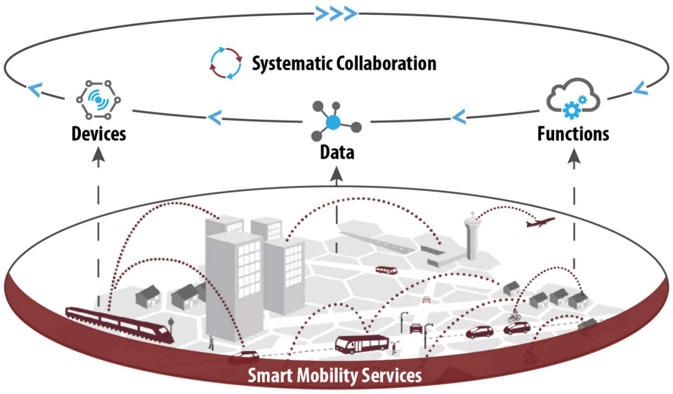


Fig. 1. The systematic collaboration among the three elements (i.e., devices, data, and functions) of smart mobility services.

in the field of transportation, the study of the systematic collaboration among the three elements becomes essential to not only impel the reuse of the well-prepared and tested elements but also to foster the service innovation in smart mobility and further elevate its intelligence and autonomy in supporting various stakeholders (e.g., service end-users, operators, planners, etc.)

Currently, such a collaboration attracts much attention, however, intrinsic barriers between the devices, data, and mobility-related functions can still be observed [10], [11], [12], namely:

- Devices are application-specific without unified protocols to harness diversified sensing and computing capabilities;
- Data are private or proprietary without a common approach to access sensitive and isolated information;
- Functions are coupled without a standardized interface to enable reusable and scalable application composition;
- Cooperations among the three elements are miscellaneous without a flexible process to implement user-oriented and system-optimized services.

To tackle these barriers, several solutions have been proposed. Initially, solutions that harness massive and heterogeneous mobility data are studied to enable analytical services for various modes, e.g., buses, taxis, subways, etc. [13], [14]. Then, solutions with layer-wise architectures (that manage the three elements separately) are explored to provide services for specific users, e.g., commuters, drivers, operators, etc. [15], [16] Even though current solutions achieve the technical capabilities to host the pre-configured elements efficiently and effectively [9], [13], [15], along with the vast application of emerging technologies and the widespread concerns about information integrity, they still face several challenges to support the systematic collaboration [17], [18], namely:

- How to harness ever-diversified devices regardless of the actual restrictions in their specifications (i.e., hardware and software) and availabilities;
- How to utilize fine-grained data preserved at both the cloud and edge and protected by different levels of access (i.e., to be open, proprietary, private, etc.);
- How to refactor coupled but well-tested functions to be customizable for a reduction and improvement of development cost and service quality, respectively.

Accordingly, this paper proposes a federated platform, called Future Mobility Sensing Advisor (FMSA). In general, compared to current solutions, its novelty is in line with the following three aspects:

- It introduces an “as a service” element repository that can manage various sensing devices and resources, multi-modal data, and reusable functions through standardized interfaces (i.e., resource controllers, data endpoints, and function callers) separately and uniformly;
- It enables a loosely coupled orchestration among the three elements based on a fully connected knowledge graph, through which, available elements, together with their correlations, can be identified according to actual service requirements elastically and adaptively;
- It supports two interaction modes (i.e., centralized and federated) to handle general service requests and responses, and, more importantly, harness resources distributed at the edge in terms of local data and computing capabilities jointly and cost-efficiently.

Moreover, FMSA is also evaluated through 1) a performance analysis to reveal its efficiency and effectiveness in optimizing the usage of resources in cloud-edge collaborating clusters, and 2) a case study of one of its applications to demonstrate its capabilities in supporting harmonious and sustainable mobility.

The remainder of this paper is structured as follows. First, related challenges and solutions in supporting the systematic collaboration are summarized in Section II. After that, FMSA is proposed in Section III, and then, evaluated and discussed in Section IV. Finally, Section V concludes the work and sketches the future research directions.

II. Emerging Challenges and Related Solutions

Smart mobility services are usually implemented and operated in a separate manner, which may cause 1) the over-release of devices/resources, e.g., the increase of empty miles in cruising due to the over-competition among car-hailing companies; 2) the under-estimation of data value, e.g., the cross-validation of data on user preference (e.g., stated and revealed) can better measure the intra- and inter-heterogeneity of user behaviors; and 3) the re-development of common functions, e.g., the trip routing and event broadcasting functions are miscellaneous among smart mobility services. To overcome the above issues, a systematic collaboration among the essential elements, i.e., devices, data, and functions, starts attracting much attention to renovate smart mobility services [7]. Along with this trend, related challenges and solutions are emerging and proposed, respectively.

A. Emerging Challenges

In general, four kinds of challenges are emerging while implementing the systematic collaboration, namely:

1) Challenge 1 (C1): Specification-Unrestricted Device Management: With the rapid development of information and communication technologies, the sensing, computing, and communicating capabilities of smart devices have become more and more diverse [19]. However, intrinsic differences in these capabilities can be identified across smart mobility services, as related devices purchased from different vendors or deployed at different network layers are often with distinguishable hardware specifications and software setups [11]. Such differences can create invisible barriers to enable interoperability among devices. Therefore, how to manage plentiful and diverse devices with fewer restrictions becomes critical to forging the collaboration foundation.

2) Challenge 2 (C2): Access-Preserving Data Integration: With the ever-growing interests and concerns about data capitalization and privacy protection, the accessibility of data structured in various forms can be preserved in different levels, e.g., to be open, proprietary, consensual, private, etc. [10], [20], [21], [22]. In turn, it can dramatically increase the complexity of heterogeneous data integration, and also may make widely applied single-model solutions less efficient or even invalidated. Hence, how to integrate multi-modal data with different levels of accessibility becomes essential to support collective awareness and intelligence.

3) Challenge 3 (C3): Requirement-Driven Function Customization: To reduce the system development cost while improving the overall system performance, the ability to encapsulate and reuse well-developed and tested functions (mostly deeply embedded in different systems) [23] becomes necessary. Unlike the existing service-orchestration solutions [24], the systematic collaboration advocates functions to be customizable and re-deployable across various platforms and systems. Accordingly, dynamic service development and operation driven by the actual requirement can be implemented to unleash the full potential of each function. Thus, enabling function customization across various systems becomes crucial to implement required service logic in a more efficient and effective manner.

4) Challenge 4 (C4): Context-Adaptive Client-Server Interaction: Due to the spatiotemporal attributes within the field of transportation, mobility users can join, hover or leave smart mobility services freely. In turn, it forms a dynamic service context as represented by the nodes with various availabilities. To ensure the reliability of smart mobility services, the client-server interaction shall be elastic to support both the common mode for general request processing and a novel mode for cloud-edge collaboration. This allows a collaborated utilization of diversified resources located at the edge to ease the workload of the central server and improve the user experience [4], [18], [20], [25]. Hence, the elasticity of client-server interaction needs to be adaptive according to the actual usage context to support a spectrum of smart mobility services.

TABLE I

THE OVERALL EVALUATION OF REVIEWED SOLUTIONS (◯ FULLY SUPPORTED ◐ PARTIALLY SUPPORTED ◯ NOT SUPPORTED)

Solutions	C1	C2	C3	C4
MOBANA (2017)	○	◐	○	○
TRANSense (2018)	◐	◐	○	○
FIWARE (2019)	◐	○	◐	○
PTS (2020)	○	◐	◐	○
SCP (2021)	○	◐	◐	○
FMSA (Proposed)	●	●	●	●

B. Related Solutions

As shown in Table I, the abilities of related solutions to address the four emerging challenges are evaluated. Initially, most scholars aim to harness massive and heterogeneous mobility data for a common data repository/center along with the emergence of big data and artificial intelligence, through which, less biased and more detailed knowledge can be mined to assist the management of various transportation modes. E.g., for buses, MOBANA [13], as a data integration and analysis framework, can monitor real-time transit data to provide value-added services; and for taxis and subways, TRANSense [14], as a collaborative and analytical framework, can detect anomalous transportation events. In general, even though these solutions can enable data-driven services to improve the efficiency and effectiveness of smart mobility, data with various granularities and accessibilities are still not well managed to support collective awareness and intelligence.

Later on, the research community starts to optimize the management of various sensing devices and accelerate the implementation of diverse stakeholder-oriented services. Based on ubiquitous IoT, the ecosystem concept, such as the aforementioned "marketplace", through which, sensing resources can be identified, required data can be extracted, and dedicated services can be provided. E.g., FIWARE [15], as an open-source IoT platform, is utilized to establish a global market where various data resources and city services can be uniformly managed and utilized; and PTS (Parallel Transportation Systems) [16], as an IoT-enabled ITS, can collaborate various traffic data to provide services for operators and end-users. As the pioneer, these solutions illustrate the potential of systematic collaboration in managing sensing devices, data, and functions. However, interoperable control of diverse resources, comprehensive integration of restricted data, and dynamic customization of re-deployable functions are still not discussed to further strengthen the collective awareness and intelligence for smart mobility.

More recently, studies about automated service processes have been conducted. E.g., SCP (Smart City Platform) [9], as a supportive infrastructure, can process an amount of heterogeneous data from multiple sources and support a range of mobility services according to a unified mobility service flow. Even though SCP is superior to conventional solutions, there can not enable controlling devices with various configurations to ensure interoperability and collaborating preserved resources at the edge to implement personalized services. Eventually, such drawbacks can affect the automated process to bridge data silos with fine-grained data distributed at the edge for the collective awareness and intelligence advocated in the systematic collaboration [22], [25].

In summary, as listed in Table I, current solutions focus more on device management, data integration, and service composition, lacking techniques to fully address the

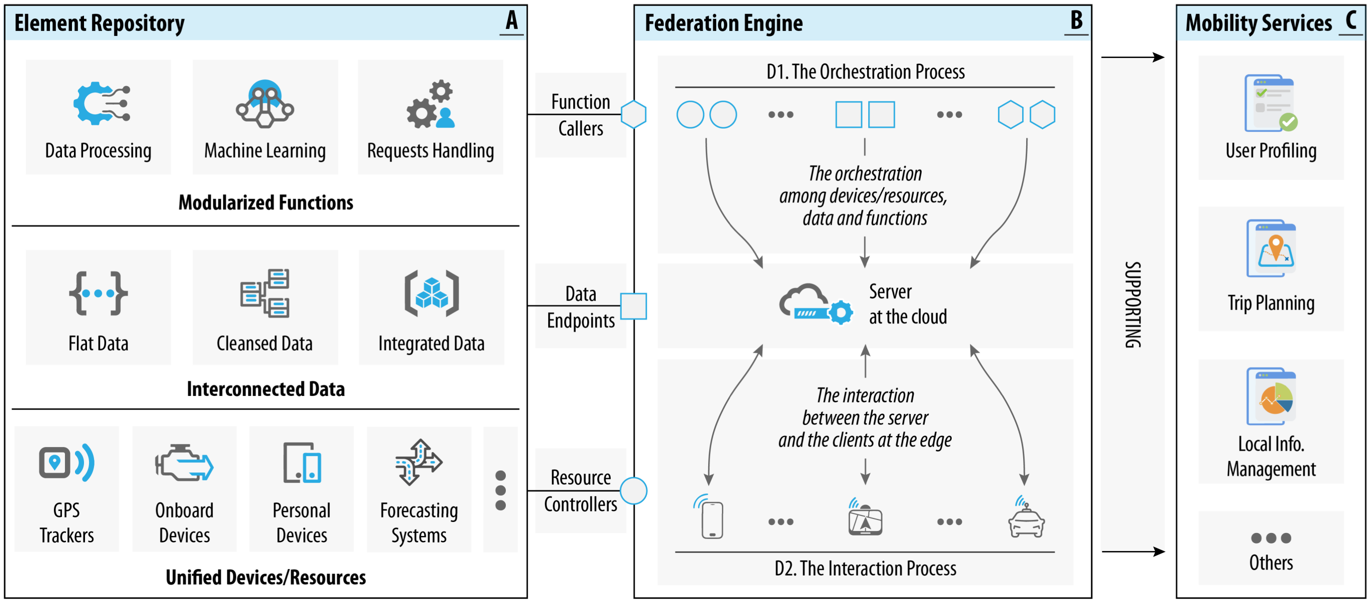


Fig. 2. The overall architecture of FMSA: A) Element Repository to manage unified devices/resources (that sense the running statuses of the system and the users), interconnected data (i.e., flat, cleansed, and integrated data), and modularized functions (i.e., data processing, machine learning, and requests handling modules); B) Federated Engine to orchestrate the three elements at the cloud and serve clients at the edge; and C) Mobility Services to provide mobility-related services (e.g., user profiling, trip planning, etc).

four emerging challenges. Specifically, managed devices are, mostly with the same or similar specifications (in terms of hardware and software) and availabilities (i.e., sharing the same working pace). Such a restriction may render current solutions obsolete for harnessing diversified IoT devices in various smart mobility services. Moreover, even though the integration of massive heterogeneous data has been widely discussed, its efficiency relies on the premise that all the original data needs to be accessible. As a result, collective awareness and intelligence will be degraded, especially when knowledge from private and sensitive data can not be directly extracted, fused, and utilized. Finally, despite the service composition can be automated to implement required functionalities, the scalability and elasticity to manage modularized functions are still limited by its centralized mode. This can not release the full potential contained in the cloud-edge collaborating system.

Hence, this paper proposes a federated platform FMSA to tackle the four challenges by implementing an on-demand element repository to comprehensively manage devices, data, and functions through unified means (i.e., resource controllers, data endpoints, and function callers). Moreover, a hybrid federation engine will be also implemented to orchestrate the three elements according to the actual application requirements, and manage both cloud and edge resources with various availabilities and accessibilities, through centralized and federated interaction modes. Accordingly, value propositions of mobility users demonstrating collective awareness and intelligence can be implemented cost-efficiently.

III. Future Mobility Sensing Advisor (FMSA)

As shown in Figure 2, FMSA contains an element repository, which decouples the devices/resources, data, and functions, and manages them as on-demand and reusable services with unified interfaces, i.e., resource controllers, data endpoints, and function callers. Moreover, based on the available elements, FMSA implements a federation engine, consisting of element orchestration and client-server interaction processes, to implement collective awareness and intelligence in a requirement-driving and resource-collaborating manner. Finally, various mobility services (a.k.a., value propositions of mobility users) can be implemented to support smart mobility with development costs reduced and service quality improved.

A. Element Repository

It forms a pool of reusable elements for FMSA, including 1) unified devices/resources to sense the statuses related to service users and transportation systems, 2) interconnected data to host and provide data in various forms and granularities, and 3) modularized functions to implement required functionalities in an elastic and loosely coupled way. In the following sections, the design details of the three elements are introduced respectively.

1) Diversified Devices/Resources: As shown at the bottom of Figure 2 (A), it consists of a) various smart devices, e.g., GPS trackers, onboard devices (OBD), personal smartphones and tablets, etc., to sense the moving behavior of vehicles and travelers; b) forecasting systems of network conditions to detect traffic changes in response to user travel decisions; and c) others resources, e.g., governmental or open platforms, for static data including vehicle registries, demographic data, traffic network typologies, public transit schedules, point of interests (POIs), etc.

Moreover, resource controllers (RCs) are also designed to provide a standardized means of controlling devices and resources existing in various smart mobility services. As shown by the abstract class diagram in Figure 3 (A),

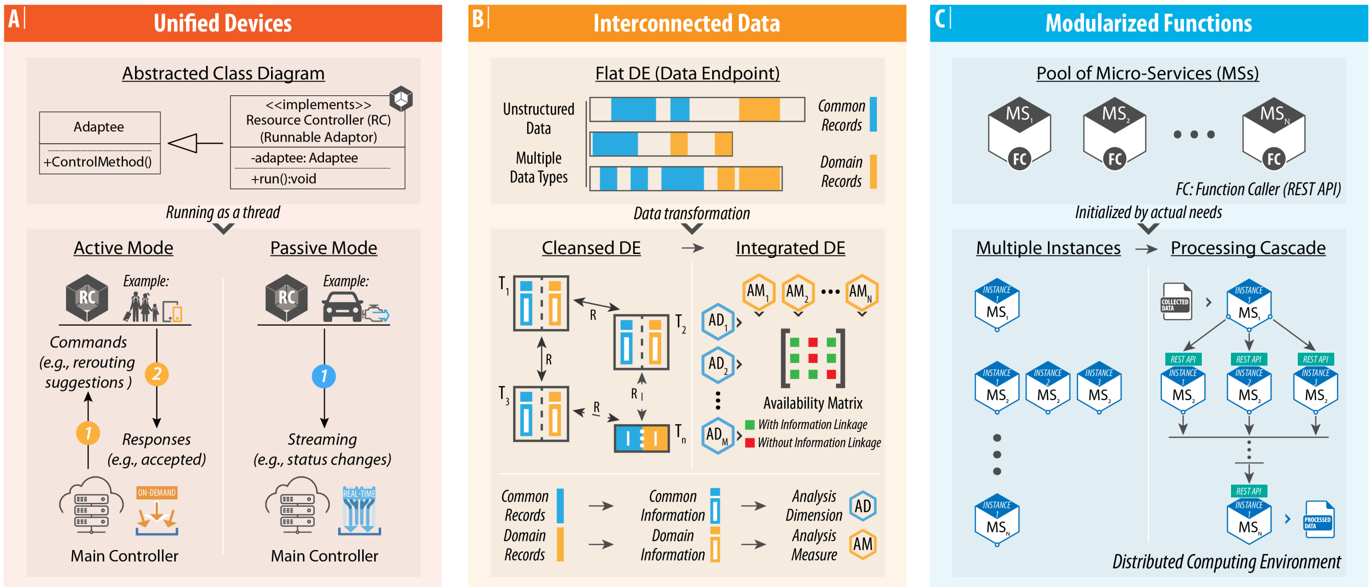


Fig. 3. The elements of FMSA and their adaptive interfaces in terms of RC (Resource Controller), DE (Data Endpoint), and FC (Function Caller). (A) Unified Devices, in which, RC is designed according to the "Adaptor" pattern and can be used in active and passive modes, (B) Interconnected Data, in which, DEs are provided to manage flat, cleansed, and integrated data, respectively, and (C) Modularized Functions, in which, micro-services can be deployed and integrated through FCs (REST APIs) to implement high-performance processing cascades according to the actual needs in a distributed computing environment.

the adaptor pattern is utilized. In general, the "Adaptee" is the device/resource providing distinguishable control methods, and the "Adaptor" implements a runnable interface (i.e., the "run" method) to encapsulate the control method of the "Adaptee". According to such a design, RCs can ensure the scalability of FMSA to manage diversified devices and resources (e.g., with different hardware and software specifications).

As for the control mode of RCs, as shown in Figure 3 (A), FMSA can control related devices/resources actively or passively. In particular, RCs in the active mode can be invoked on-demand. E.g., while users are traveling, the main controller of FMSA can push rerouting suggestions to their personal devices and receive corresponding responses (i.e., accepted or not). On the contrary, the passive RCs are bounded with a channel listener to run consistently. E.g., the vehicles' statuses can be streamed to the main controller in real-time. Generally, these two kinds of RCs can be used adaptively based on the specification of related devices/resources, e.g., smart devices can be manipulated either by active RCs if they require interactions with the main controller periodically, or by passive RCs if they need to remain proactive.

2) Interconnected Data: As shown in the middle of Figure 2 (A), it manages heterogeneous multi-source data in three kinds, namely a) flat data containing original data sensed; b) cleansed data storing normalized data with noises removed (e.g., POI records without coordinates, drifting GPS points, etc.) and useful information extracted (e.g., travel modes and stops detected); and c) integrated data managing fused data with pre-defined analysis dimensions (e.g., people, place, time, etc.) and measures (e.g., travel modes, stop activities, etc.).

Moreover, a semi-semantic data model, called interconnected data model (IDM) [12], is implemented to ensure data scalability. As shown in Figure 3 (B), in flat data, records are unstructured with multiple types/groups. In each group, useful records are identified and categorized into common records that are shared among data groups, and domain records that merely appeared in certain data groups. Second, in cleansed data, relational tables (T) and relationships (R) among them are designed according to the data types presented in the flat data. In each table, unstructured common and domain records are transformed into standardized common information (CI) and domain information (DI). Finally, in the integrated data, analysis dimensions (ADs) and analysis measures (AMs) are created according to CI and DI, respectively, and information linkages (ILs) are defined according to table relationships.

Specifically, to identify the connectivity between ADs and AMs, an IL availability matrix Γ is created as defined in formula 1, where Γ is a $m \times n$ matrix, m is the total number of ADs, n is the total number of AMs; and the coexistence of AD_i and AM_j is determined by the condition that whether CI of AD_i and DI of AM_j are in a unique record generated through the natural join [$\bowtie$] of related tables T.

$$\Gamma = \begin{bmatrix} IL_{11} & IL_{12} & \cdots & IL_{1n} \\ IL_{21} & IL_{22} & \cdots & IL_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ IL_{m1} & IL_{m2} & \cdots & IL_{mn} \end{bmatrix}$$

$$s.t. \begin{cases} IL_{ij}, \; i = \; 1 \; to \; m, \;\; j = \; 1 \; to \; n \\ IL_{ij} = \begin{cases} 1, & AD_i \circledast AM_j \; is \; True \\ 0, & otherwise \end{cases} \\ AD_i \circledast AM_j = \begin{cases} True, & \bowtie_{CI_i, DI_j} [T] \neq NULL \\ False, & otherwise \end{cases} \end{cases} \tag{1}$$

TABLE II
THE MAPPING BETWEEN CRUD AND HTTP(S) METHODS

Operations	Methods	Description
Create	POST	To create or save data to a dataset
Read	GET	To retrieve data from a dataset
Update	PUT	To update records in a dataset
Delete	DELETE	To remove records from a dataset

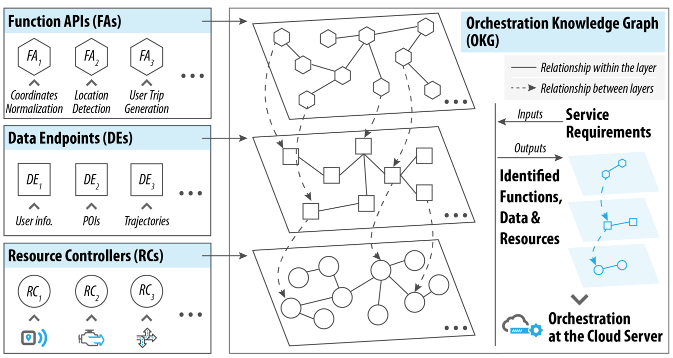


Fig. 4. The orchestration among devices/resources, data, and functions. Note that: 1) Available devices/resources, data, and functions are represented by RCs, DEs, and FCs, respectively; 2) The orchestration knowledge graph (OKG) manages RCs, DEs, and FCs together with their in-tier and between-tier relationships; and 3) The orchestration process consumes service requirements as inputs and provides identified devices/resources, data, and functions as outputs.

Through IDM, data can be sequentially transformed (as shown at the bottom of Figure 3 (B)) from the common and domain records in flat data (e.g., POI, user-specified origins and destinations, etc.), to the common and domain information in cleansed data (e.g., all converted to places with unified geotags), and finally to the analysis dimensions (ADs) (i.e., virtualized place indicators) and measures (AMs) (e.g., visited number) in integrated data. Accordingly, input data can be uniformly processed and knowledge can be rapidly mined by configuring the ADs and AMs. For instance, to answer how often people in different age groups will visit the places recommended by the system, a query with AD as the age groups, and AM as recommended places can be configured and applied to the integrated data for the results.

Finally, to build a common data management portal, data endpoints (DEs) are provided to assist the CRUD (Create, Read, Update and Delete) operations. According to the mapping given in Table II, DEs can be implemented by first encoding related data queries and results in JSON format and then encapsulating and transiting the JSON message together with authorization tokens in HTTP(s) requests. Such that, data preservation and usage can be decoupled and supported efficiently and effectively.

3) Modularized Functions: As shown on the top of Figure 2 (A), there are three kinds of modularized functions, namely a) data processing modules, which support the data transformation from flat data to integrated data as defined in IDM; b) machine learning modules, which implement centralized or federated algorithms to detect stops of sensed trajectories (based on centralized learning) [18], train personalized models for trip recommendations (based on federated learning) [22], etc.; and c) requests handling modules providing corresponding responses for service requests, e.g., user registration, trip planning, reward management, etc.

Moreover, as shown in Figure 3 (C), modularized functions can form a pool of micro-services (MSs). In general, MSs of FMSA can be deployed dynamically and utilized through Function Callers (FCs), which are defined as RESTFul APIs. Hence, multiple instances of MSs can be initialized and integrated as processing cascades to implement specific service logics according to actual needs, e.g., to cleanse, reconcile and integrate heterogeneous multi-source data, or to federate multiple devices to build a user behavior analysis model in a privacy-preserving manner. Owing to the elasticity enabled by MSs, the number of instances can be adjusted based on the actual workloads in FMSA to remove the potential performance bottlenecks.

B. Federation Engine

As shown in Figure 2 (D), it consists of two processes, namely an orchestration process among the three elements to support collective awareness and intelligence by implementing required service logic agilely, and an interactive process to support the communication between the server in the cloud and clients at the edge adaptively.

1) Orchestration Process: As shown in Figure 4, it intends to implement the service logic according to the given requirements automatically. To achieve that, first, a unified orchestration knowledge graph (OKG) is created to a) store RCs, DEs and FCs separately in three tiers; b) define the relationships of two nodes within each of RC, DE and FC tiers; and c) specify relationships between two adjacent tiers, including tiers of RCs and DEs, and tiers of DEs and FCs.

Moreover, a sub-graph containing related orchestration elements and their relationships can be extracted according to service requirements based on three steps:

Step 1 (Node Extraction): It will first match the service requirements with related function sets, denoted as M_{FC}, then identify data M_{DE} utilized and generated by the M_{FC}, and finally select devices/resources M_{RC} associated with the M_{DE};

Step 2 (Edge Extraction): Based on M_{FC}, M_{DE}, and M_{RC}, it will extract directly linked in-tier and between-tier relationships, denoted as E_{In} and E_{Bt}, respectively;

Step 3 (Subgraph Creation): By combining the identified nodes and edges, the subgraph $S_{OKG} = ((V : M_{FC}, M_{DE}, M_{RC}), (E : E_{In}, E_{Bt}))$ can be created. As an example, the subgraph to support a personalized trip planning service is given in Table III.

Finally, based on the subgraph S_{OKG}, a loosely coupled orchestration process can be implemented. Specifically, a) according to the extracted nodes, the orchestration skeleton can be created with related functions, data, and devices/resources, and then, b) according to the identified edges, the orchestration logic can be implemented. After that, the process can be deployed in a distributed running environment to support the given service. It is worth noting that the

TABLE III
THE ORCHESTRATION SUBGRAPH FOR A PERSONALIZED TRIP PLANNING SERVICE

Items in S_{OKG}	Description
Node: FC_1	Trip Planner to plan a trip according to user-specified origin and destination
Node: FC_2	Trip Personalizer to rank trip alternatives according to user preferences
Node: DE_1	User Choices with records about selected trips
Node: DE_2	User Travel Histories with records about user performed travel trajectories
Node: RC_1	Personal Devices providing user choices and travel histories
Edge: E_{In1}	FC_1 with FC_2, trips are recommended after they are generated
Edge: E_{In2}	DE_1 with DE_2, a trip is selected before execution
Edge: E_{Bt1}	FC_2 with DE_1, trips are recommended based on user choices
Edge: E_{Bt2}	FC_2 with DE_2, trips are recommended based on user travel histories
Edge: E_{Bt3}	DE_1 with RC_1, trips are selected on user devices
Edge: E_{Bt4}	DE_2 with RC_1, trips are executed and monitored by user devices

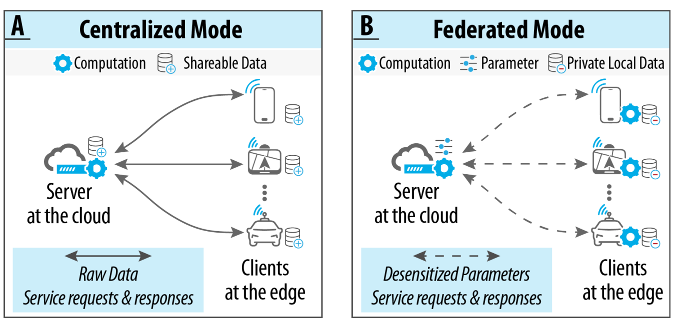


Fig. 5. The interaction between the server at the cloud and the clients at the edge: A) the centralized mode to support the interaction with the clients, whose data can be uploaded and processed at the server; and B) the federated mode to support the interaction with the clients, whose data needs to be preserved and processed locally.

process can be optimized to keep the workload balanced by creating and deploying multiple nodes (instances of related interfaces) when performance droppings are experienced.

2) Interaction Process: Besides the common client-server interaction in handling service requests and responses, there is another interaction to support the usage of data sensed and resources maintained at the edge.

Since the local data at the edge can be categorized into a) public or consented data that are shareable through secured channels, and b) sensitive and private data that are preserved and only used by the data owner, as shown in Figure 5, the centralized as well as federated modes are introduced to support the usage of local data, respectively:

Mode 1 Centralized mode (CM): It has been widely utilized to support various smart mobility services. In general, as shown in Figure 5 (A), it requires that a) all the local data from the edges shall be uploaded to the server, and b) all the local data will be managed and processed by the server. Based on such a mode, a high-performance computing (HPC) server with a high bandwidth is generally necessary to harness massive data uploaded by the clients.

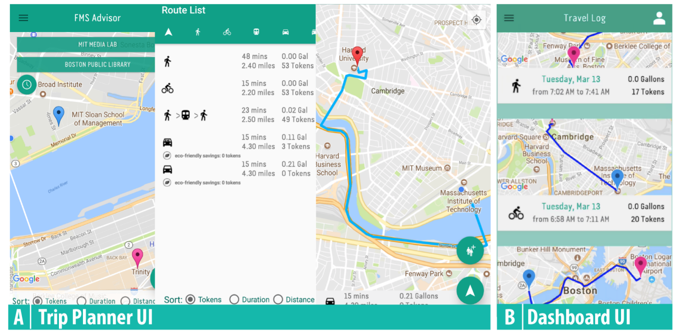


Fig. 6. (A) Trip Planner UI, which assists users to plan, view, and execute a trip; (B) Dashboard UI, which assists users in checking and reviewing their travel histories.

Mode 2 Federated mode (FM): Contrary to the centralized mode, FM can handle data silos caused by regulations about data protection and user privacy (i.e., data are owned by the end users and can not be shared or exchanged). Specifically, instead of exchanging raw data directly, the edges will first process their data locally according to the requirements of the server (i.e., learning a personalized trip planner based on a discrete choice model), and then share the desensitized parameters (i.e., model parameters) to the server. After the local parameters are received, the server will aggregate them for the global parameters that will be returned to the edges for their local usage. By using such a mode, the workload and the communication cost of the server can be reduced.

In general, through the federation engine, the service orchestration can be implemented based on the fully connected knowledge graph rapidly, and the client-server interaction can be supported based on the two modes adaptively. Since the orchestration process can be load balanced, the actual usage of the interaction process can affect the final performance of FMSA. Hence, the advantages and disadvantages of the two interaction modes will be evaluated in Section IV-A.

C. Mobility Services

As shown in Figure 2 (E), it consists of various mobility-related services, including those for user registration, trip planning, local information management, etc. To better serve the users, dedicated UIs (User Interfaces) can be designed and implemented. For example, as shown in Figure 6 (A), through the trip planner UI, users can plan, view, and execute a trip. Moreover, through the dashboard UI as shown in Figure 6 (B), users can view their travel histories with a summary of travel modes, durations, and tokens rewarded by the system.

In summary, as a novel federated platform with high adaptivity, scalability, and elasticity, FMSA can enable a systematic collaboration among the three elements to renovate smart mobility services. Specifically, first, sensing devices/resources, mobility data, and reusable functions can be decoupled and managed through unified interfaces (i.e., RCs, DEs, and FCs), respectively. Second, the hybrid cooperation process can be implemented in the federation engine to support both the orchestration among the three elements and the interaction between the server at the cloud and the clients at the edge. Finally, various stakeholder-oriented services can be supported by FMSA, e.g., to implement personalized mobility.

IV. Evaluation and Discussion

As a federated platform, FMSA can be deployed in a distributed environment consisting of central servers in the cloud, and smart devices at the edge. Unlike the centralized solutions, whose computational tasks are mostly executed at the server without utilizing idle but plentiful resources at the edge, the potential improvement in system performance enabled by FMSA is directly related to the proposed federated mode. Therefore, the two client-server interaction modes, i.e., CM and FM, are first evaluated to reveal their advantages and disadvantages. Next, the capability of FMSA as a smart mobility service to foster harmonious and sustainable mobility is demonstrated through a case study.

TABLE IV
The Variables Used to Configure PDs and Local Data in Each Learning Round

Variable	Value	Description
NP	1 to 10%	Percentage of new PDs per round
IP	1 to 10%	Initial proportion of the data
ND	1 to 2%	Percentage of new data per round
DS	180	Maximum data size

TABLE V
List of Abbreviations Used in the Evaluation Measures

Abbr.	Description	Abbr.	Description
CM	Centralized Mode	FM	Federated Mode
CPT	Computation	COM	Communication
CS	Central Server	PD	Personal Devices
SN	Simple number	NS	New Samples
T	Time	C	Cost
TT	Training time	TC	Training cost

A. Performance Evaluation of FMSA

Based on the decoupled design of devices, data, and functions, as well as the scalable orchestration process among them, the performance of FMSA mainly relays on the type of interaction process to be utilized. Such that, this section will analyze the advantages and disadvantages of CM and FM through a common learning task.

1) Learning Process: In general, CM and FM are with different learning processes:

- **The learning process of CM**: It consists of two steps, namely a) a data uploading step, in which, local data of the edges are uploaded; and b) a model training step, in which, the model is trained at the server iteratively;
- **The learning process of FM**: It works with a) a local training step, in which, local models are trained at the edges and then uploaded to the server; and b) a model aggregation step, in which, the global model is updated at the server and afterwards broadcasted to the edges to start a new learning iteration.

Even though the learning processes of CM and FM are different, the same model with similar performance can be trained [25], [26]. Such that, the evaluation will be conducted to analyze their differences in training time and cost.

2) Data and Model Configuration: Without loss of generality, a standard dataset Swissmetro [27] is utilized, which contains over ten thousand samples about travel choices made by more than one thousand respondents on travel options of car, metro, and train. Since the dataset may contain records with missing or misleading information, it is cleansed by removing samples of respondents 1) without choices on the three options or 2) with inconsistencies in the choices (e.g., the time and cost of travel are not proportional).

After that, a common evaluation environment is created with one central server (CS) and one hundred personal devices (PD). In the meanwhile, to reproduce the real-world scenario that the number of users and user data grows over time, PDs and their local data will be controlled according to the configuration listed in Table IV. Specifically, in each learning round, the number of PDs grows at a rate of NP. Besides that, the size of the local data of a PD starts with an initial proportion IP, and grows with a rate of ND in each learning round until reaching DS.

Finally, Gibbs sampling is applied in CS and PDs for parameter fitting to building a personalization model (based on the mixed logit model).

3) Data Transmission Process: In accordance with the learning processes, CM and FM have different data transmission processes to build the personalized model as well:

- **The data transmission process of CM**: Since CM needs raw data to build the data, along with the growth of PDs and their local data, the newly added data of the existing clients, as well as the initial data of the newly joined clients, will be uploaded to the server;
- **The data transmission process of FM**: Instead of uploading the raw data, the model parameters of each client are uploaded to the server for the update of the global parameters.

4) Evaluation Measures: Before starting, the abbreviations and the hyperparameters used in the evaluation measures are listed in Table V and Table VI, respectively.

In general, the training time and cost (i.e., TT and TC) are directly correlated to the ones of learning participants in each round. Hence, initially, the time and cost of computation and communication of a participant in a learning round shall be measured. By adopting the methods used in [28] and [29], they can be measured separately according to Formulas 2, 3, 4, and 5. It is worth noting that the energy consumption of

TABLE VI
LIST OF HYPERPARAMETERS USED IN THE EVALUATION MEASURES

Param.	Value	Description
SM	0.01kb	The size of model parameters
SS	0.07kb	The size of a sample
f_{CS}	3.0 GHz	CPU frequency of CS
f_{PD}	0.5 GHz	CPU frequency of PD
Q	10 cycles/bit	Computation workload per sample
B	1 MHz	Bandwidth of the network
r_j	100-1000 kb/s	Uploading rate of the j_{th} PD
ε_{cpt}	1	Computation coefficient
ε_{com}	50	Communication coefficient

computation and communication is used to represent the training cost in the proposed measures.

$$T_{i,k}^{CPT} = \frac{Q_{i,k}}{f_i} \tag{2}$$

$$C_{i,k}^{CPT} = \varepsilon_{cpt} \times Q_{i,k} \times {f_i}^2 \tag{3}$$

$$T_{i,k}^{COM} = \frac{S_{i,k}}{r_i} \tag{4}$$

$$C_{i,k}^{COM} = \frac{2 \times \varepsilon_{com} \times S_{i,k}}{r_i} \times 2^{\frac{r_i}{B}-1} \tag{5}$$

where i indicates a participant, which can be CS or PD; and k represents a learning round. Moreover, for the i_{th} participant in the k_{th} learning round, $T_{i,k}^{CPT}$ and $C_{i,k}^{CPT}$ stand for the computation time and cost respectively, as well as $T_{i,k}^{COM}$ and C_k^{COM} for the communication time and cost; $Q_{i,k}$ represents the overall workload of model training; f_i donates the CPU frequency; $S_{i,k}$ represents the size of data to be uploaded; r_i defines the uploading rate; and finally, ε_{cpt}, ε_{com} and B are the computation coefficient, communication coefficient and bandwidth respectively.

Since CM and FM work differently, their corresponding $T_{i,k}^{CPT}$, $C_{i,k}^{CPT}$, $T_{i,k}^{COM}$, $C_{i,k}^{COM}$ shall be calculated separately:

a) The calculation for CM: Since the model training of CM only happens at the server, the computation time and cost of CM are equivalent to the ones of CS.

Hence, by applying $Q_{CS,k}$ and f_{CS}, Formulas 2 and 3 can be rewritten to Formulas 6 and 7, respectively.

$$T_k^{CPT,CM} = \frac{Q_{CS,k}}{f_{CS}} \tag{6}$$

$$C_k^{CPT,CM} = \varepsilon_{cpt} \times Q_{CS,k} \times {f_{CS}}^2 \tag{7}$$

where $T_k^{CPT,CM}$ and $C_k^{CPT,CM}$ are the computation time and cost of CM in the k_{th} learning round; and $Q_{CS,k}$ is the workload of CS that can be measured based on Formula 8.

$$Q_{CS,k} = Q \times \sum_j^N (S_{j,k})$$
$$s.t.: \ S_{j,k} = SS \times SN_{j,k} \tag{8}$$

where Q is the average workload per sample; SS is the unified size of a data sample; and for the j_{th} PD at the k_{th} round, $S_{j,k}$ is its total sample size (in KB), and $SN_{j,k}$ is its sample number (in counts).

Moreover, since CM only requires new data (including the local data of newly added PDs, and the new data added to the existing PDs) to be uploaded in parallel, by applying the size of new samples $NS_{j,k} = (S_{j,k} - S_{j,k-1})$ and related uploading rate r_j into Formulas 4 and 5, the communication time and cost of CM can be computed by Formulas 9 and 10, respectively.

$$T_k^{COM,CM} = \max_{j \in N}(\frac{NS_{j,k}}{r_j}) \tag{9}$$

$$C_k^{COM,CM} = \sum_j^N (\frac{2 \times \varepsilon_{com} \times NS_{j,k}}{r_j} \times 2^{\frac{r_j}{B}-1}) \tag{10}$$

where $T_k^{COM,CM}$ and $C_k^{COM,CM}$ are the communication time and cost of CM in the k_{th} learning round.

Finally, the training time and cost of CM in a learning round, noted as TT_k^{CM} and TC_k^{CM}, can be calculated according to Formulas 11 and 12, respectively.

$$TT_k^{CM} = T_k^{CPT,CM} + T_k^{COM,CM} \tag{11}$$

$$TC_k^{CM} = C_k^{CPT,CM} + C_k^{COM,CM} \tag{12}$$

b) The calculation for FM: Since the complexity of the model aggregation step of FM at CS can be $O(1)$ [30], and each PD can train the local model concurrently, it needs to measure the computation cost and time of each PD based on Formulas 13 and 14, respectively.

$$T_{j,k}^{CPT,FM} = \frac{Q_{j,k}}{f_{PD}} \tag{13}$$

$$C_{j,k}^{CPT,FM} = \varepsilon_{cpt} \times Q_{j,k} \times {f_{PD}}^2 \tag{14}$$

where $T_{j,k}^{CPT,FM}$ and $C_{j,k}^{CPT,FM}$ are the computation time and cost of the j_{th} PD in the k_{th} learning round of FM; and $Q_{j,k}$ is the result of Formula 15.

$$Q_{j,k} = Q \times S_{j,k} \tag{15}$$

Moreover, after the local training, PDs will send their local model parameters to CS in parallel. Therefore, the communication time and cost of each PD in FM can be computed according to Formulas 16 and 17 respectively.

$$T_{j,k}^{COM,FM} = \frac{SM_j}{r_j} \tag{16}$$

$$C_{j,k}^{COM,FM} = \frac{2 \times \varepsilon_{com} \times SM_j}{r_j} \times 2^{\frac{r_j}{B}-1} \tag{17}$$

where SM_j is the size of local parameters of the j_{th} PD. It is worth noting that since all the PDs train the same model, they will have the same SM.

Finally, due to the concurrency of PDs, the training time and cost of FM in learning round, marked as TT_k^{FM} and TC_k^{FM}, can be calculated according to Formulas 18 and 19, respectively.

$$TT_k^{FM} = \max_{j \in N}(T_{j,k}^{CPT,FM} + T_{j,k}^{COM,FM}) \tag{18}$$

$$TC_k^{FM} = \sum_j^N (C_{j,k}^{CPT,FM} + C_{j,k}^{COM,FM}) \tag{19}$$

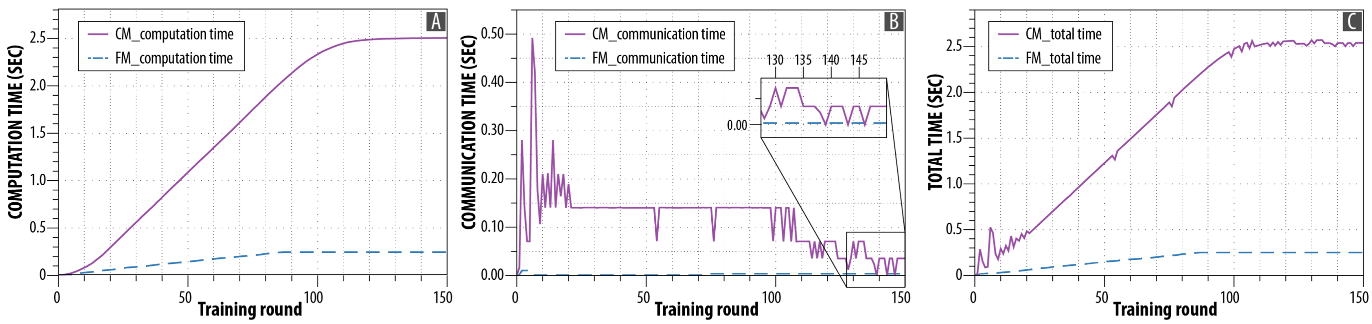


Fig. 7. The evaluation on interaction time. A) The analysis of computation time; B) the analysis of communication time; and C) the analysis of the total time consumed in CM and FM, respectively.

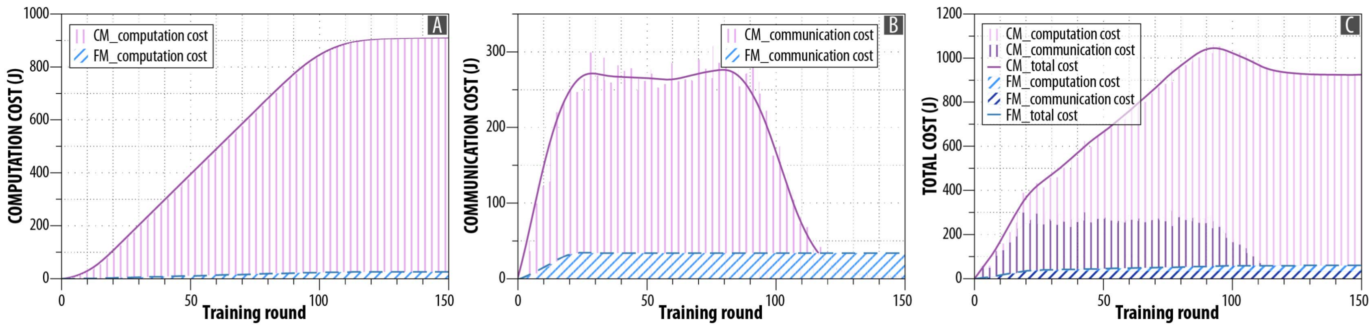


Fig. 8. The evaluation on interaction cost. A) The analysis of computation cost; B) the analysis of communication cost; and C) the analysis of the total cost required by CM and FM, respectively.

5) Evaluation Results: According to the evaluation setting and measures described above, the performance in terms of interaction time and cost of CM and FM are evaluated.

a) Interaction time: First, as shown in Figure 7 (A), FM is about 10 times faster than CM in computation, as parallelization is implemented among edges in FM. It shows that FM can boost the performance by utilizing the idle computing power dispersed at the edges, instead of fully relying on the central server as of CM.

Second, as shown in Figure 7 (B), the curve of communication time of CM vibrates heavily at the beginning and then decreases gradually to the bottom, where FM stays consistently. It shows that after a service cluster gets stabilized (i.e., to learn a global model) with a periodic update, both CM and FM can maintain a low bandwidth consumption.

Finally, the overall time curves presented in Figure 7 (C) share similar shapes and trends to the ones illustrated in Figure 7 (A). It shows that for both CM and FM, the computation time is more decisive than the one of communication in supporting a rapid interaction between the server and clients.

b) Interaction cost: First, CM needs to accumulate all the data and process the "big" data once at a time, while FM enables edges to process their "small" data in parallel. Accordingly, the computation cost (measured by the energy consumption) of FM is significantly lower than the one of CM as shown in Figure 8 (A). It shows that even though edges are generally with a relatively low computing power compared to CS (e.g., f_{PD} is 6 times lower than f_{CS}), the concurrency at the edge can still significantly reduce the burden of CS for a performance boost.

TABLE VII
THE ADVANTAGES AND DISADVANTAGES OF CM AND FM

Mode	Advantages		Disadvantages	
CM	1)	Comprehensive for data integration	1)	Waste of computing resources at the edge
	2)	Less computation intensive at the edge	2)	Performance bottleneck at the peak
	3)	Easy to control and manage the cluster	3)	Single-point failure and data leakage
FM	1)	Workload balanced service cluster	1)	Incapable for high-resolution data
	2)	Protection of data security and privacy	2)	Intrusive to use the local resources
	3)	Multi-model knowledge fusion	3)	Complex to create information linkage

Second, as shown in Figure 8 (B), similar to the trend in communication time, the communication cost of CM rapidly reaches and consistently stays at the peak in the early stage, and then gradually drops to the bottom. While, expect the ignition stage, the cost of FM remains the same over time, as regular interaction between the edges and CS is required. It shows that CM may overcome FM after services get stabilized when fewer data exchanges are required.

Finally, since the cost of computation is much higher than the communication, FM can remain an obvious advantage in supporting the service cluster as illustrated in Figure 8 (C). It shows that by comparing CM and FM, the utilization of edge resources can form a balanced and outperforming cluster

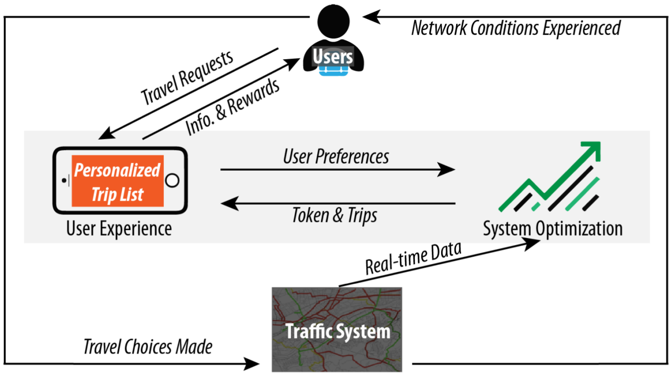


Fig. 9. The personalization workflow of Tripod supported by FMSA.

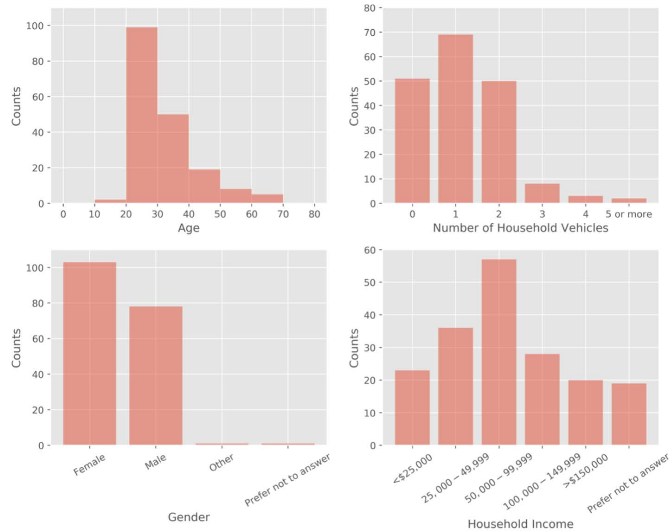


Fig. 10. Respondent Characteristics: (A) Age, (B) Number of Household vehicles, (C) Gender, and (D) Household income.

to support various smart mobility services with reduced overall cost.

In summary, Table VII lists the advantages and disadvantages of the two interaction modes in FMSA. In general, although CM may cause a waste of idle resources at the edge and information leakage due to the single point failure, it is beneficiary for data integration to preserve integrated data in high resolution and is rational to control and manage services with less computing capabilities at the edge. In contrast, FM may impede the creation of a full-resolution dataset, since their representative information is extracted and uploaded for data integration instead of exchanging sensitive data directly. Nevertheless, it can better manage distributed resources to remove performance bottlenecks experienced during the peak of CM, and also connect more devices with a stable and moderate workload for the central server. Hence, since FMSA implements both CM and FM to manage the interaction between the server and the clients, its scalability and robustness in resource utilization can be ensured for high service quality and friendly user experience.

B. Case Study of FMSA

As an application of FMSA, the Tripod (Sustainable Travel Incentives with Prediction, Optimization and Personalization) project [31] is conducted to orient users towards more sustainable travel alternatives by offering them real-time information and incentives. To achieve such a goal, FMSA is used to 1) personalize user trips; 2) support a pilot at the Greater Boston Area (GBA); and finally, 3) analyze user acceptance rates and behavioral changes to reveal the effects of Tripod.

1) Trip Personalization: It generates recommended trips by considering not only the preferences of users but also the real-time conditions of the traffic system through an interaction between user experience and system optimization as illustrated in Figure 9. Specifically, before a journey starts, travel requests will be fulfilled through a list of personalized trips. Moreover, after a recommended trip is chosen and executed, user incentives in terms of expected travel experience, real-time travel information, and reward tokens will be provided based on the savings gained by the system. Through such a flow, it is expected to see that users will be more supportive of a system with predefined objectives fulfilled, and as the benefits, the forecasted travel experience can be provided, and system-wide savings can be accumulated.

2) Pilot Description: It is conducted from April to July 2018. Since sufficient user travel data is needed to support the impact analysis, a total of 603 users were recruited. They are required to install the app on their smartphones (i.e., Android or iOS), and to provide their anonymous travel histories for a period of 14 consecutive days. The studied data includes revealed preferences, i.e., stops and their purposes, travel paths and corresponding travel modes, and stated preferences, i.e., options on choosing other modes for a trip performed previously. To better summarize the data distribution, the characteristics of qualified participants (who have completed the 14 days requirements) are outlined in Figure 10. We can see that the sample is more towards young and middle-aged respondents reflected by their age and household income. This is as expected since the study is smartphone-based, which is more attractive and user-friendly for these people.

Moreover, to integrate the data generated during the study (e.g., sensing data from GPS and accelerators, choice data about travel modes, screened user profiles, etc.), FMSA is customized by orchestrating three key functions, namely, a) a travel chain detection function, which detects activities at stops and modes of travels (for the detail, please refer to [32] and [33]), b) a trip planning function, which generates candidate trips according to user inputs, e.g., origin, destination, etc., and c) a trip validation function, which verifies user actual trajectories with selected trips. Apart from the above three functions, FMSA also consists of other essential functions, e.g., user profiling, travel dashboard, etc.

Finally, based on the collaboration of devices, data, and functions, FMSA provides the fundamental services as shown in Figure 6, and more dedicated services as shown in Figure 11. Specifically, a travel diary is provided for users to review and correct their travel trajectories on the map as shown in Figure 11(A), and also related modes in the timeline as shown in Figure 11 (B). It is worth noting that the studied

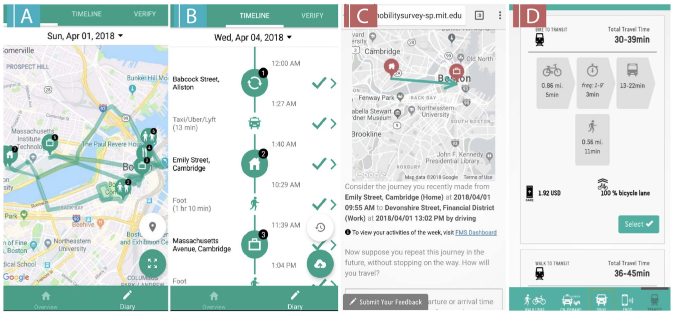


Fig. 11. (A) Travel Trajectory UI, which shows the user travel trajectory of one day on the map; (B) Travel Dairy UI, which shows the user travel diary of one day in the timeline; (C) Summary of Past Trip UI, which provides the details of a trip performed previously; (D) List of Trip Alternatives UI, which presents the details of each trip alternative.

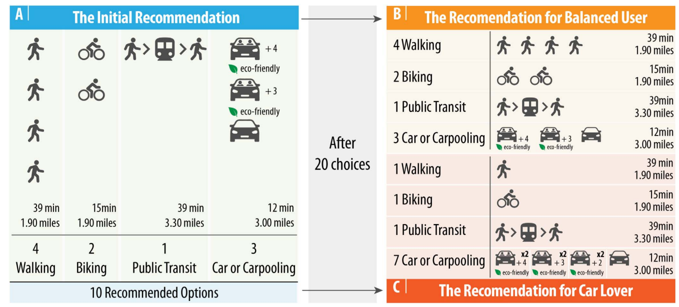


Fig. 12. The analysis of behavioral changes: A) The initial recommendation; (B) The recommendation of "balanced user" after 20 choices; (C) The recommendation for "car lover" after 20 choices.

TABLE VIII
NON-PERSONALIZED (NP) VS. PERSONALIZED (P) HIT-RATES

	Hit-rate of NP	**Hit-rate of P**
Day_0	75.9% *	75.9% *
Day_5	77.1% ↑	81.0% ↑
Day_{10}	77.2% |	82.0% ↑
OR (Observed Rate)	**77.2%**	**84.2%**

Note that: 1) [*] Since no observed choices are made on day 0, the hit-rates of NP and P are the same; and 2) OR (Observed Rate) is the hit-rates of actual participants.

travel modes include non-motorized modes (i.e. walking, biking, and bike-sharing), private motorized modes (i.e., car and carpooling), and on-demand modes (e.g., Uber/UberPool, Lyft/Lyft Line, car sharing, and taxi). In addition, to support the collection of user choices on preferred trip alternatives, which are generated according to a trip previously performed, two more UIs are implemented to a) present the details of the past trip with the origin, destination, start time, and end time as shown in Figure 11 (C); and b) the preferred alternatives to be selected with a summary of related travel cost, modes and durations as shown in Figure 11 (D).

3) Result Analysis: Through the pilot, the effects of Tripod are analyzed to answer two main questions: a) whether users are willing to accept personalized trips that are adjusted to recede potential impacts on the overall system; and b) whether user behaviors are altered towards the overall objective of the system (to be sustainable in terms of low-carbon and eco-friendly). Accordingly, the acceptance rate of recommended trips and the behavioral change of users are analyzed.

a) The analysis of acceptance rate: It is measured by the hit-rates between the selections of personalized and non-personalized trips. Since a large number of choices are needed to make a rational comparison, 10 thousand synthetic agents, which represent the behaviors of actual participants in choosing recommended trips (specified by the pairs of origin and destination observed in the travel histories of participants), are created, and accordingly, the changes of hit-rates are analyzed. For the technical details to set up this simulation, please refer to our previous work [34].

As listed in Table VIII, the result shows that on days 5 to 10, the average hit-rate of non-personalized trips increases slightly from 75.9% to 77.2%, while that of personalized trips increases substantially to 82% on day 10. Moreover, for users with a sufficiently long choice history, the average hit-rate of personalized trips will reach the observed rate (OR) of 84.2% (calculated based on the data of actual participants).

Since personalized trips have a higher possibility of being chosen, it can be observed that mobility users are not withstanders unwilling to cooperate with the system. Instead, they can be incentivized to perform trips with both their preferences and the system objective addressed.

b) The analysis of behavioral change: It is analyzed based on the data of a "balanced user", who makes a rational choice among all of the alternatives recommended, and a "car lover", who prefers car-related alternatives instead of others. Moreover, to avoid potential disturbances, two types of users are controlled to make independent choices on a trip with the same origin and destination for 20 times. Note that, most of the respondents are young and middle-aged, and some of them might prefer long walking for the purpose of exercise and relaxation after school or work during the summer.

As shown in Figure 12 (A), in the first round, these two types of users receive the same recommendation with 10 options consisting of 4 walking, 2 biking, 1 public transit, and 3 car or car-pooling options. It is worth noting that the initial recommendation has been optimized toward the overall system objective. After 20 rounds, as shown in Figure 12 (B), the "balanced user" can still obtain a tender recommendation. In contrast, as shown in Figure 12 (C), the "car lover" can obtain one option for each of walking, biking, and public transit, but 7 options for car or carpooling, which are associated with tokens rewarded by the system for eco-friendly trips.

It can be inferred that even though options could be incentivized towards the overall objective of the system (low-carbon and eco-friendly), it can still serve relevant users to a limited degree, instead of apparently altering their behaviors.

In summary, the pilot of Tripod demonstrates the capabilities of FMSA to support a personalized and sustainable mobility system. Through the dedicated services of FMSA, users can make their own contributions to the system invisibly, and the system can transfer itself towards the predefined objective gradually.

C. Discussion

As a novel solution, FMSA can be applied to study and support individual mobility by integrating both user preferences and system objectives. Accordingly, there can be two kinds of operation modes, namely:

- **Operated by private companies**: In such a mode, FMSA can be released as an on-demand service for both inhabitants and local authorities. Specifically, as for the inhabitants, their daily commutes can be supported and incentivized with the expected travel experience and accumulated rewards, respectively. Moreover, as for the local authorities, the system-wise optimization towards the pre-defined objectives (e.g., to be eco-friendly) can be achieved by gradually cultivating users' travel behaviors, and more importantly, related changes can also be tracked and analyzed in a long-running and privacy-preserving manner for informed decision-making;
- **Operated by government authorities**: In such a mode, FMSA can be deployed as a common platform to test and analyze the influence of certain technologies and policies (e.g., the application of autonomous vehicles, and low-carbon restrictions) even before launching. Moreover, instead of spending efforts on maintaining the platform, the local authorities may contract with a third party to actually manage the platform and recruit participants. Therefore, they can focus more on analyzing the sensed data for valuable insights to optimize or upgrade the system eventually.

With the above two operation modes, the technological advancement of FMSA can bring multiple benefits for related stakeholders, namely:

- **Collective awareness through multi-model data integration**: From the perspective of end-users, collective awareness requires a cost-efficient means to serve users with inter- and intra-heterogeneities that are represented or embedded in both open and private data. In FMSA, besides the integration of open data that has been widely discussed and applied, a privacy-preserving aggregation of user data is also supported by a federated interaction between the clients and the server. Such that, collaborative awareness in smart mobility services can be implemented based on intelligent cores built by fusing knowledge of multi-modal data;
- **Unified management of the three elements**: From the perspective of system administrators, unified management can significantly reduce the burden of monitoring and detecting the running statuses and the performance bottlenecks of the system, respectively. In FMSA, the three elements, i.e., devices, data, and functions, are encapsulated by standard interfaces and can be used "as a service". Hence, while a wide range of smart mobility services adopts such a mechanism, an element-level collaboration can be enabled through a smart mobility ecosystem that can not only accelerate the service deployment process but also reduce related maintenance costs.
- **Adaptive orchestration for various value propositions**: From the perspective of service developers, adaptive orchestration becomes essential in the field of transportation, as it can rapidly deliver user-oriented services that can fertilize the innovation of smart mobility services and broaden their spectrum. In FMSA, a requirement-driven collaboration among the three elements is implemented based on a semantic OKG, through which, elements and their relations can be identified. Accordingly, by orchestrating identified elements, dedicated service logic can be implemented dynamically and adaptively with service development complexity and cost reduced.

V. Conclusion

This paper presents a federated platform, called FMSA, which enables collective awareness and intelligence through a systematic collaboration among the three elements of smart mobility services. Compared to current solutions, first, FMSA can manage a common element repository to decouple and refactor the three elements (i.e., devices/resources, data, and functions in and across smart mobility services) through three standard interfaces (i.e., RCs, DEs, and FCs), respectively. Second, it also can provide a federation engine to support the orchestration among the three elements and the interaction between the server and clients in a requirement-driving and resource-collaborating manner. Finally, it can implement various mobility services cost-efficiently to impel the application of systematic collaboration, in turn fostering the development of the ecosystem for smart mobility.

Moreover, the efficiency and effectiveness of FMSA are demonstrated through a performance evaluation and a case study. Specifically, in the evaluation, the performance of the platform is analyzed through the testing of FM and CM. Compared to CM, FM can not only locally process sensitive data but also better utilize distributed computing resources to ease the burden of the central server with computation time accelerated by about 10x and cost reduced by about 6x. Moreover, in the case study, the capability of the platform is further analyzed through one of its applications named Tripod to support an individual-oriented and system-wise optimized mobility service. In general, it reveals that FMSA can not only be easily customized and enhanced to manage service clusters with high performance but also support various mobility services with high quality. In Tripod, the balance between the users and the overall system can be achieved, as both individual preferences and system-wise objectives are mutually measured to support smart mobility.

In the future, the systematic collaboration supported by FMSA should be further enhanced by investigating a visualized and computational representation of the three elements. Such that, a common foundation can be established to explore intelligent and autonomous transportation systems that can reconcile the physical and cyber spaces and make smart mobility to be more human-independent. Second, an asynchronous mechanism for harnessing both public and private data will be studied to enable a system-wide knowledge fusion and transformation. Finally, the service categories of FMSA will be enriched and deployed to foster the development of the smart mobility ecosystem.

REFERENCES

[1] J. Mulerikkal, S. Thandassery, M. D. K. Deepa, V. Rejathalal, and B. Ayyappan, "JP-DAP: An intelligent data analytics platform for metro rail transport systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9146–9156, Jul. 2022.

[2] R. Kazhamiakin, E. Loria, A. Marconi, and M. Scanagatta, "A gamification platform to analyze and influence citizens' daily transportation choices," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2153–2167, Apr. 2021.

[3] I. Al Ridhawi, M. Aloqaily, A. Boukerche, and Y. Jararweh, "Enabling intelligent IoCV services at the edge for 5G networks and beyond," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5190–5200, Aug. 2021.

[4] L. You, J. He, W. Wang, and M. Cai, "Autonomous transportation systems and services enabled by the next-generation network," *IEEE Netw.*, vol. 36, no. 3, pp. 66–72, May 2022.

[5] D. Thekke Kanapram et al., "Collective awareness for abnormality detection in connected autonomous vehicles," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3774–3789, May 2020.

[6] H. Tran-Dang, N. Krommenacker, P. Charpentier, and D.-S. Kim, "Toward the Internet of Things for physical internet: Perspectives and challenges," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4711–4736, Jun. 2020.

[7] L. You, B. Tuncer, R. Zhu, H. Xing, and C. Yuen, "A synergetic orchestration of objects, data, and services to enable smart cities," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10496–10507, Dec. 2019.

[8] A. Bucchiarone, S. Battisti, A. Marconi, R. Maldacea, and D. C. Ponce, "Autonomous shuttle-as-a-service (ASaaS): Challenges, opportunities, and social implications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3790–3799, Jun. 2021.

[9] A. Al-Rahamneh et al., "Enabling customizable services for multimodal smart mobility with city-platforms," *IEEE Access*, vol. 9, pp. 41628–41646, 2021.

[10] D. Bi, S. Kadry, and P. M. Kumar, "Internet of Things assisted public security management platform for urban transportation using hybridised cryptographic-integrated steganography," *IET Intell. Transp. Syst.*, vol. 14, no. 11, pp. 1497–1506, 2020.

[11] M. A. A. da Cruz, J. J. P. C. Rodrigues, J. Al-Muhtadi, V. V. Korotaev, and V. H. C. de Albuquerque, "A reference model for Internet of Things middleware," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 871–883, Apr. 2018.

[12] L. You, B. Tunçer, and H. Xing, "Harnessing multi-source data about public sentiments and activities for informed design," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 2, pp. 343–356, Feb. 2019.

[13] M. Aloqaily, I. A. Ridhawi, and M. Guizani, "Energy-aware blockchain and federated learning-supported vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 22641–22652, Nov. 2022.

[14] Y. Lu, A. Misra, W. Sun, and H. Wu, "Smartphone sensing meets transport data: A collaborative framework for transportation service analytics," *IEEE Trans. Mobile Comput.*, vol. 17, no. 4, pp. 945–960, Apr. 2018.

[15] F. Cirillo, G. Solmaz, E. L. Berz, M. Bauer, B. Cheng, and E. Kovacs, "A standard-based open source IoT platform: FIWARE," *IEEE Internet Things Mag.*, vol. 2, no. 3, pp. 12–18, Sep. 2019.

[16] F. Zhu, Y. Lv, Y. Chen, X. Wang, and F. Wang, "Parallel transportation systems: Toward IoT-enabled smart urban traffic control and management," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4063–4071, Oct. 2020.

[17] S. Liu, Q. Chen, and L. You, "Fed2A: Federated learning mechanism in asynchronous and adaptive modes," *Electronics*, vol. 11, no. 9, p. 1393, Apr. 2022.

[18] L. You, F. Zhao, L. Cheah, K. Jeong, P. C. Zegras, and M. Ben-Akiva, "A generic future mobility sensing system for travel data collection, management, fusion, and visualization," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4149–4160, Oct. 2020.

[19] X. Li and L. Da Xu, "A review of Internet of Things—Resource allocation," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 8657–8666, Jun. 2021.

[20] G. Chen et al., "Federation in cloud data management: Challenges and opportunities," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 7, pp. 1670–1678, Jul. 2014.

[21] D. Kondor, B. Hashemian, Y.-A. de Montjoye, and C. Ratti, "Towards matching user mobility traces in large-scale datasets," *IEEE Trans. Big Data*, vol. 6, no. 4, pp. 714–726, Dec. 2020.

[22] L. You, J. He, J. Zhao, and J. Xie, "A federated mixed logit model for personal mobility service in autonomous transportation systems," *Systems*, vol. 10, no. 4, p. 117, Aug. 2022.

[23] G. Castellano, F. Esposito, and F. Risso, "A service-defined approach for orchestration of heterogeneous applications in cloud/edge platforms," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 4, pp. 1404–1418, Dec. 2019.

[24] C. Paniagua, J. Eliasson, and J. Delsing, "Efficient device-to-device service invocation using arrowhead orchestration," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 429–439, Jan. 2020.

[25] L. You, S. Liu, Y. Chang, and C. Yuen, "A triple-step asynchronous federated learning mechanism for client activation, interaction optimization, and aggregation enhancement," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 24199–24211, Dec. 2022.

[26] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, and M. Jirstrand, "A performance evaluation of federated learning algorithms," in *Proc. 2nd Workshop Distrib. Infrastruct. Deep Learn.*, Dec. 2018, pp. 1–8.

[27] M. Bierlaire, K. Axhausen, and G. Abay, "The acceptance of modal innovation: The case of Swissmetro," in *Proc. Swiss Transp. Res. Conf.*, 2001, pp. 1–17.

[28] C. T. Dinh et al., "Federated learning over wireless networks: Convergence analysis and resource allocation," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 398–409, Feb. 2021.

[29] R. Yu and P. Li, "Toward resource-efficient federated learning in mobile edge computing," *IEEE Netw.*, vol. 35, no. 1, pp. 148–155, Jan./Feb. 2021.

[30] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.* (Proceedings of Machine Learning Research), vol. 54, A. Singh and J. Zhu, Eds. USA: PMLR, Apr. 2017, pp. 1273–1282.

[31] C. L. Azevedo et al., "TRIPOD: Sustainable travel incentives with prediction, optimization, and personalization," in *Proc. Transp. Res. Rec. 97th Annu. Meeting*, 2018, pp. 1–19.

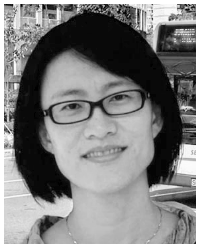
[32] Y. Kim, A. Ghorpade, F. Zhao, F. C. Pereira, P. C. Zegras, and M. Ben-Akiva, "Activity recognition for a smartphone and web-based human mobility sensing system," *IEEE Intell. Syst.*, vol. 33, no. 4, pp. 5–23, Jul./Aug. 2018.

[33] A. Ghorpade, F. C. Pereira, F. Zhao, C. Zegras, and M. Ben-Akiva, "An integrated stop-mode detection algorithm for real world smartphone-based travel survey," in *Proc. Transp. Res. Board 94th Annu. Meeting*, vol. 15, 2015, p. 6021.

[34] X. Song, M. Danaf, B. Atasoy, and M. Ben-Akiva, "Personalized menu optimization with preference updater: A Boston case study," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 8, pp. 599–607, Dec. 2018.

Linlin You (Member, IEEE) received the B.Sc. degree in software engineering from the Harbin Institute of Technology, the dual master's degrees (Hons.) in software engineering and computer science from the Harbin Institute of Technology and the University of Pavia, respectively, and the Ph.D. degree in computer science from the University of Pavia. He is an Associate Professor at the School of Intelligent System Engineering, Sun Yat-sen University, and also a Research Affiliate at the Intelligent Transportation System Laboratory, Massachusetts Institute of Technology. He was a Senior Post-Doc at the Singapore-MIT Alliance for Research and Technology, and a Research Fellow at the Architecture and the Sustainable Design Pillar, Singapore University of Technology and Design. He published more than 50 journals and conference papers in research fields, such as smart cities, autonomous transportation systems, service orchestration, heterogeneous multi-source data fusion, and federated learning.

Mazen Danaf received the B.Eng. and M.Eng. degrees from the American University of Beirut and the Ph.D. degree in transportation from MIT. He is a Research Affiliate with the Intelligent Transportation Systems Laboratory, MIT, and a Market Research Scientist at Aurora. His research focuses on transportation demand modeling, and its applications for self-driving cars and trucks. He published more than 20 journals and conference papers in transportation, demand modeling, simulation, and machine learning.

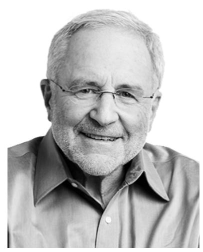


Fang Zhao received the B.Eng. and M.Eng. degrees from the National University of Singapore in 2000 and 2001, respectively, and the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology in 2010. She is a Research Scientist at the Singapore-MIT Alliance for Research and Technology (SMART), Future Urban Mobility (FM) Group. She has an interdisciplinary background with expertise in machine learning, communication networks, and travel surveys. She has led a Research Team in SMART to develop the future mobility sensing technology.

Jinping Guan received the B.S. degree in economic geography and urban planning from Sun Yat-sen University and the joint-training Ph.D. degree in transportation planning and management from Tongji University, Shanghai, China, and the University of California, Berkeley. She is an Associate Professor at the Harbin Institute of Technology. She was a Researcher at the Massachusetts Institute of Technology and a Research Affiliate at Harvard University. She is the author of three books and more than 40 papers. Her research interests include travel behavior and policy, transportation big data and cloud computing platform, the impacts of mobility-as-a-service and autonomous vehicles, mindfulness, and urban spatial structure and transportation. She was a recipient of the National Huaxia Award for Excellence, the Chinese ITS Conference Best Paper Award, the Chinese National Scholarship for Outstanding Ph.D., and the National Urban Planning Best Report Award.

Carlos Lima Azevedo received the B.Sc. degree in structural engineering and the M.Sc. and Ph.D. degrees from the University of Lisbon in 2004, 2008, and 2014, respectively. He is an Associate Professor at the Technical University of Denmark (DTU) and a Research Affiliate at the ITS Laboratory, Massachusetts Institute of Technology (MIT). Prior to joining DTU, he held a position of a Research Scientist and the Executive Director of the Transportation Education Committee, MIT, a Post-Doc at the Singapore-MIT Alliance for Research and Technology, and a Research Scholar at the Portuguese National Laboratory for Civil Engineering. He is an expert in the development of new models and simulation techniques for smart mobility design and assessment. Previous research includes the development and application of large-scale agent-based urban simulation, its application for the design and evaluation of shared automated vehicles on-demand, the development of personalized real-time incentive systems, individual behavior experiments and preferences modeling toward new mobility solutions, methods for the calibration and validation of simulation tools, and innovative road safety assessment models without accident data.

Bilge Atasoy (Member, IEEE) received the B.Sc. and M.Sc. degrees in industrial engineering from Bogazici University, Istanbul, in 2007 and 2009, respectively, and the Ph.D. degree in integrated supply and demand models in transportation problems from EPFL, in November 2013. She has been an Associate Professor at the Department of Maritime and Transport Technology, Transport Engineering and Logistics Section, TU Delft, since April 2018. Her main research interests lie at the intersection of optimization and behavioral models. She applies scientific methodologies in the field of various transport and logistics problems. Prior to joining TU Delft, she was a Research Scientist at the Intelligent Transportation Systems (ITS) Laboratory, MIT, where she is currently a Research Affiliate. At MIT, she led several research projects in the areas of real-time optimization, travel behavior and choice-based optimization for which she also worked on funding applications. She received the Best Ph.D. Thesis Award from the Swiss Operations Research Society.



Moshe Ben-Akiva received the Ph.D. degree in transportation systems from MIT and the Honorary degrees from the University of the Aegean, the Université Lumiére Lyon, the KTH Royal Institute of Technology, and the University of Antwerp. He is the Edmund K. Turner Professor of civil and environmental engineering at the Massachusetts Institute of Technology (MIT), the Director of the MIT Intelligent Transportation Systems Laboratory, and the Principal Investigator at the Singapore-MIT Alliance for Research and Technology. He has worked as a consultant in industries such as transportation, energy, telecommunications, financial services, and marketing, for a number of private and public organizations, including Hague Consulting Group, RAND Europe, and Cambridge Systematics, where he was previously a Senior Principal and a member of the Board of Directors. He was an Advisor at Memetrics and ChoiceStream, provided litigation support to Analysis Group and Brattle Group and is the Chief Scientific Advisor to Mobile Market Monitor. He has coauthored two books, including the textbook *Discrete Choice Analysis* (MIT Press), and nearly 400 papers in refereed journals or refereed conferences. He was recently a member of the Future Interstate Highway System Committee of the National Academies of Sciences, Engineering, and Medicine. His awards include the Robert Herman Lifetime Achievement Award in Transportation Science from the Institute for Operations Research and the Management Sciences, the Lifetime Achievement Award of the International Association for Travel Behavior Research, the Jules Dupuit prize from the World Conference on Transport Research Society, and the Institute of Electrical and Electronics Engineers ITS Society Outstanding Application Award for DynaMIT, a system for dynamic network management.