Federated Architecture for Personal Mobility Service in Autonomous Transportation System

Junshu He[1]; Zhuolin Deng[2]; Sheng Liu[3]; Linlin You[4]; Meng Zhou[5]; Ming Cai[6]; and Shaowu Cheng[7]

[1]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China
[2]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China
[3]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China
[4]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China (corresponding author). Email: youllin@mail.sysu.edu.cn
[5]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China
[6]School of Intelligent Systems Engineering, Sun Yat-Sen Univ., Guangzhou, China; Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, China
[7]Dept. of Transportation Engineering, Harbin Institute of Technology, Harbin, China

ABSTRACT

Along with the trend towards an autonomous transportation system (ATS), the intelligence of personal mobility service (PMS) can be further lifted by sensing travelers' statuses comprehensively, learning behavior patterns accurately, providing travel options appropriately, and giving service responses timely. Such a process relies on a seamless information flow, which shall address data silos caused by laws and regulations about privacy. This paper proposes a federated architecture for PMS, called FPMS, which adopts federated learning, to provide personalized multi-modal options by aggregating personal data in a privacy-preserving way, and utilizing idle resources of personal devices within the service cluster. In general, by analyzing the physical objects involved, functions required, and data processed, a reference architecture of FPMS is designed to guide its construction in ATS effectively and efficiently. Moreover, a performance evaluation between FPMS and conventional centralized PMS is also presented to reveal the advantages of FPMS in saving service costs.

INTRODUCTION

The development of Intelligent Transportation System (ITS) is impelled by a series of emerging technologies (e.g., cloud computing, big data, and 5G), which enable the transformation towards the next-generation autonomous transportation system (Autonomous Transportation System, ATS). Specifically, in ATS, a novel workflow as of "Sensing-Learning-Rearranging-Reacting" (Qin et al. 2019) is implemented to actively perceive status changes of the system, dynamically learn potential mobility demands, rationally schedule related resources, and rapidly provide nonintrusive services.

As one of the core services of ATS, Personal Mobility Service (PMS) is required to support personalized services with the integration of multi-model travel options under the integrated workflow of ATS. First, PMS needs to collaborate with different participants in the transportation system to comprehensively sense statuses of travelers. Second, it shall have the capabilities to learn macroscopic system situations and microscopic user behaviors accurately to assist the decision making in defining service schedules (Danaf et al. 2019). Finally, it can provide personalized solutions with a maximum system-saving efficiently and effectively.

To achieve such goal, massive and diverse data shall be sensed and processed to support personalization and optimization. Conventionally, centralized approaches are employed in PMS, named centralized PMS (CPMS) (Cui et al. 2018, Logesh et al. 2018, Sun et al. 2019), in which, traveler's information and other related data are mostly sent and processed in the data center (Taivalsaari and Mikkonen 2018). Nowadays, as the service scale expand continuously to cover more travel modes and serve more distinctive users, CPMS may face the following challenges:

1) Data security and privacy: As CPMS needs to continuously collect massive sensitive data (e.g., personal references, trajectories, and activities), the central point failure may cause the leakage of information, e.g., the Tesla case where vehicle location information was exposed (Liu et al. 2020). Moreover, since various countries issue laws and regulations (e.g., General Data Protection Regulation (GDRR)) in protecting personal data to advocate for the construction of trustworthy service platforms, many data silos may incapacitate the application of CPMS.

2) Service performance and quality: CPMS needs to implement most of its functionalities in a centralized cloud, which may become the performance bottleneck to process related requests timely, especially, when the number of service users grow exponentially and rapidly (Ahvar et al. 2019).. In such case, the reliability and availability of PMS may be downgraded in providing a high quality service.

3) User experience and acceptance: The heterogeneity within user preferences requires PMS to react with personalized options. However, due to the fact that user preferences are generally omitted in CPMS, current solutions can mainly generate related user menus based on objective metrics, such as time and cost (Bajaj et al. 2015), which may affect the user experience, resulting in a low acceptance rate.

To tackle these emerging challenges, this paper, first, proposes a novel federated PMS, called FPMS, which incorporates a privacy-preserving decentralized mechanism, named federated learning (FL), to coordinate service participants in PMS (McMahan et al. 2017, Yang et al. 2019). Secondly, the reference architecture of FPMS in functional, logical, and physical views are proposed. Finally, the performance difference between FPMS and CPMS is analyzed through a dedicated evaluation to demonstrate the merit of FPMS in collaborating the service cluster consisting of massive smart devices.

OVERVIEW OF PMS

Personal mobility service (PMS), as one of the core services of ATS, provides personal travel options integrating multiple transportation modes to assist individual mobility. As shown in Figure 1 (a), when starting a trip, users can access the personal trip menu timely via various channels (e.g., smartphone app, tablet, and computer), to view and decide on route, mode, time, cost, etc. for personalized services. Such a process is enabled by the collaboration between the physical and virtual worlds as shown in Figure 1 (b). In general, users can receive personalized menus upon on individual requests, which are tackled through a global optimization between user preferences and system objectives to elevate the system saving as well as user experience. In general, through an inifite loop, which can constantly sense data of users and system, and actively provide personalized mobility services, PMS can continuously improve the quality of services, and the efficiency of the system.

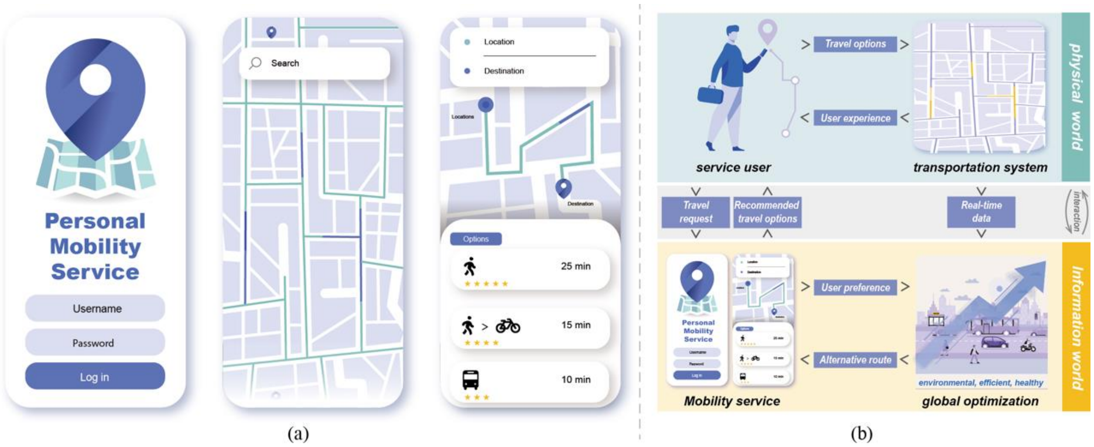

Figure 1. (a) A classic PMS trip menu UI for smartphone; (b) A general PMS collaborating the physical and cyber spaces;

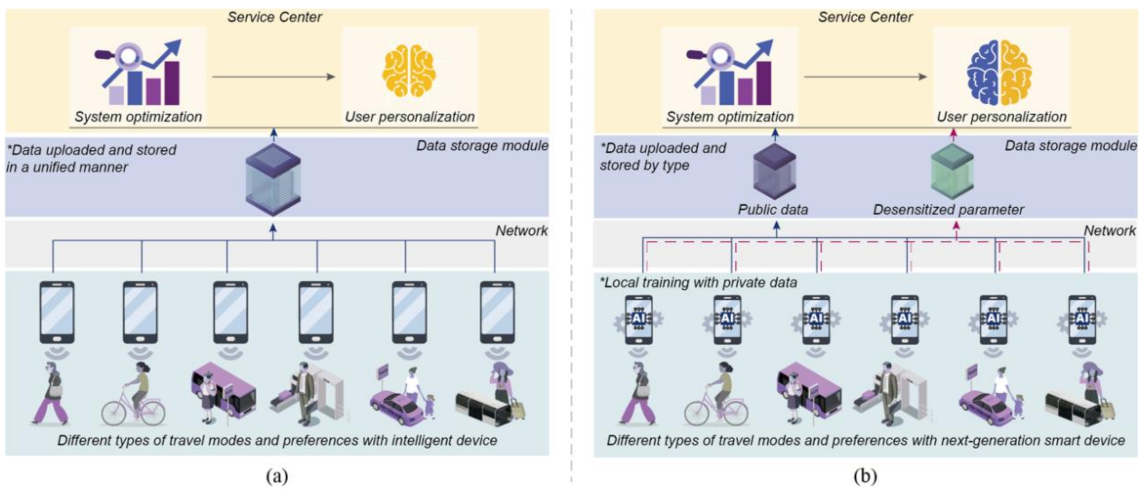

Figure 2. (a) The workflow of CPMS; (b) The workflow of FPMS.

Since such a continuous improvement mainly relies on the sensing and learning capabilities of PMS, two distinctive pradigms of PMS are discussed: 1) Centralized PMS (CPMS), which requires data to be gathered and processed in the service center as shown in Figure 2(a); and 2) Federated PMS (FPMS), which can process individual mobility data in a privacy-preserving way by uploading desensitized parameters to the service center as shown in Figure 2(b).

Comparing to CPMS, FPMS can not only address the data silos caused by rules and regulations related to data security and privacy, but also optimize the resource utilization with in the service cluster formed by heterogeneous devices. In short, FPMS is a step ahead of the current solution as illustrated by CPMS to enable a novel way in optimizing the mobility demand and supply with a multi-modal information integration by using both open and private data.

PROPOSED ARCHITECTURE OF FPMS

The construction of services in ATS relies on the analysis of physical objects participation, the business logic implementation and the service functions involved. Hence, to assist the design of FPMS in ATS, the reference architecture of FPMS in physical, functional and logical views are proposed and discussed in this section.

Physical Object in FPMS. As show in Figure 3, physical objects (PO) of FPMS can be categorized into three groups, namely entity, module, and system:

Entity: As the basic PO in ATS, the entity is the primary information source, which interact with the internal functions of the system. In general, entities in FPMS mainly contain a) traveler (e.g., commuters, tourists, etc.), which consumes the service; b) vehicle (e.g., private cars, commercial vehicles, etc.), which supports the service for the movement of travelers; and c) environment (e.g., temperature, humidity, etc.), which may influence the service indirectly.

Module: As the intermediate PO between entity and system, the module provides simple functions embedded and utilized at the edges. In general, it consists of a) Environmental Monitoring Equipment (EME), which collects current surface weather conditions from sensors on-board; b) Roadside Monitoring Equipment (RME), which collects current road condition by querying fixed sensors on or near the roadway; c) Personal Information Device (PID), which provides the capability for travelers to receive personal travel menu. Specifically, PID can serve as a common interface for travelers to interact with the system through personal travel choice models (also called travel utility function (Azevedo et al. 2018)), which is learnt based on FL to generate personalized travel options.

System: As the compound PO of the service, the system implements complex functions required by the service. In general, it mainly contains a) Weather System (WS), which provides weather information, e.g., weather forecasts, warnings of hazardous weather, etc.; b) Transportation Information Center (TIC), which acts as a data hub for information extraction and dissemination; c) Transportation Management Center (TMC), which manages multi-mode vehicles on the road, and diverse services under the request; d) Personalized Recommendation Center (PRC), which maintains the person travel choice model learnt from FL, and optimizes the running of system to not only maximize the overall savings but also improve the user experience by measuring system objectives and traveler preferences in the same time.

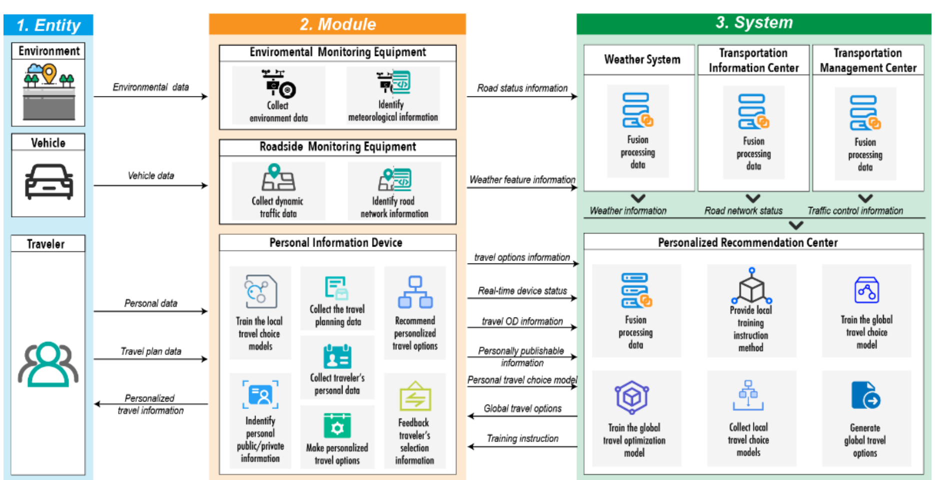

Figure 3. The physical view of FPMS with physical objects involved.

Functional Orchestration in FPMS. All functions encapsulated in physical objects can be categorized into four domains, namely autonomous-sensing, autonomous-learning, autonomous-rearranging, and autonomous-reacting, as illustrated in Figure 4 (a).

Autonomous-sensing: It includes the sensing functions for active information acquisition. To protect user privacy, the sensing function can work in two modes, namely a) public mode, in which, data, such as travel OD (Origin and Destination), dynamic traffic data, etc., with user consents or open accessibilities can be obtained, and b) private mode, in which, data are stored and processed locally at the edge, and desensitized aggregation parameters are exchanged between the edge and cloud.

Autonomous-learning: It includes the learning functions, such as data fusion, portrait extraction, and option optimization. Similar to the data sensing modes, the learning function can also work under two distinctive modes. First, in private mode, as illustrated in Figure 4 (b), through FL, each edge can train the travel choice model locally and upload the local model to the cloud for a sharable global model, which can also be further customized at the edge. Second, in public mode, through the fusion of the multi-source heterogeneous data, the cloud can comprehensively obtain the running status of the system.

Autonomous-rearranging: It includes the rearranging functions to generate personalized travel options automatically. In general, two steps are required for the option generation, namely a) general option generation, which creates global options according to the system objectives at the cloud, and b) personal option generation, which customized the global options according to user preferences at the edge.

Autonomous-reacting: It includes the reacting functions to support the execution of selected travel options. In general, a feedback loop is formed by linking the reacting functions with the sensing functions. Hence, through a continuous iteration, the service quality and user experience, which measure the service level of PMS, can be consistently improved.

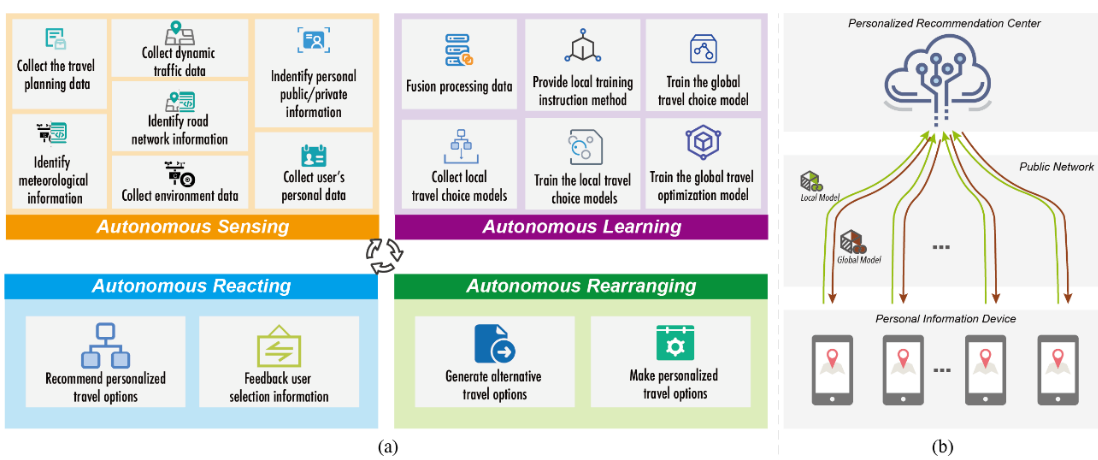

Figure 4. (a) The function composition of FPMS; (b) Schematic diagram of the implementation of the federated learning in autonomous learning domain.

Logical Workflow in FPMS. The logical workflow is shown in Figure 5, which illustrates the relationship between two connected functions through interaction flows exchanging data in four kinds. They are: a) open data, which is obtained from open sources, e.g., the road network from OpenStreetMap; b) consent data, which includes data with user consents, e.g., travel OD, user feedbacks, etc.; c) desensitized data, which is desensitized from private data, such as user preferences; and d) operation data, which includes data generated during the operation of FPMS, such as fused information, global model, etc.

According to the ATS operation logic of "sensing-learning-rearranging-reacting", FPMS defines eight steps in the logical flow, namely:

Data collection: It describe the data acquisition process in sensors, roadside cameras, and other related sensing devices.

Data extraction: It describes the data extraction process for each kind of data sensed for key information required.

Data fusion: It describes the data fusion process to aggregate data in different modes, and dispatch aggregated data to related functions.

Data analysis: It describes the data analysis process to train a travel choice model and a global travel optimization model.

Option generation: It describes the option generation process by using the global travel choice model and travel optimization model at the cloud.

Option personalization: It describes the option personalization process to customize global options according to user preferences at the edge.

Option execution: It describes the option execution process to assist travelers during their trips through interactive user interfaces.

Option feedback: It describes the option feedback process to collect user experience information for continuous improvement.

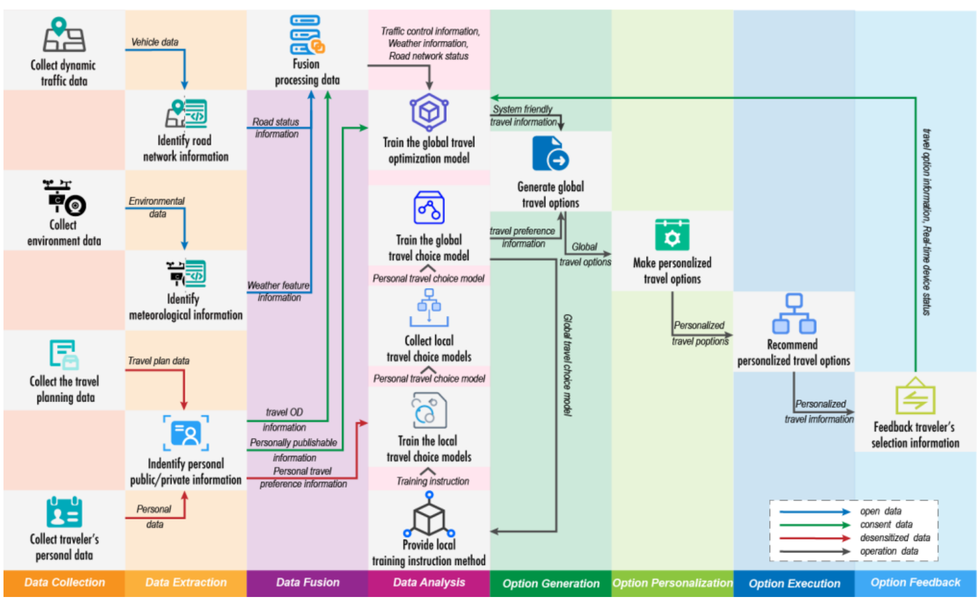

Figure 5. The general work flow of FPMS in a logical view.

THE PERFORMANCE EVALUATION

The variation between CPMS and FPMS is mainly reflected by the learning functions, as they are implemented according to centralized and decentralized approaches, respectively. Hence, a common training setting is defined to reveal the performance differences in model training time and resource consumption.

Evaluation Setting. Firstly, the Swissmetro dataset (Bierlaire et al. 2001) is utilized as the common training data, which is collected in Switzerland in 1998 with 10,729 samples reflecting the travel choice on options of private car, Swwissmetro, and train with three attributes, i.e., travel cost, time, and distance.

Secondly, as for the training procedure, one PRC (Personalized Recommendation Center) and 100 PIDs (Personal Information Devices) are visualized. Moreover, the number of PIDs that participate and related data processed in a learning round will increase gradually as defined in Table 1. Specifically, the initial number of PID is 1, and it increases at a rate of P. Besides, the local dataset of PID contains an initial proportion N about 1% to 5%, and grows with a defined rate I from 1% to 2%.

Finally, the Gibbs sampling (Danaf et al. 2019) is applied in the cloud server and each PID for parameter fitting to train the model.

Table 1. The common constraints used in assigning data.

Variable	Description	Value
$D_{\min}$	Minimum amount of data	50
$D_{\max}$	Maximum amount of data	180
$P_{\min}$	Minimum number of newly added PID per round	1%
$P_{\max}$	Maximum number of newly added PID per round	10%
$N_{\min}$	Minimum portion of initialized data	1%
$N_{\max}$	Maximum portion of initialized data	5%
$I_{\min}$	Minimum portion of new data per round	1%
$I_{\max}$	Maximum portion of new data per round	2%

Evaluation Indicators. The training of the model requires multiple iterations with the increase of PID and related data until a steady state is reached with a desired accuracy. Such that, the training time and resource consumption of each training round in the first 150^{th} rounds are utilized as the evaluation indications.

In general, the computation and communication cost are directly associated with the size of the data sample processed in each learning round. Hence, let i and k index the PID and the round of iteration, respectively. Then, in the k^{th} learning round, the size of the sample contained by the i^{th} PID is denoted as $V_{i,k}$, and the size of the sample utilized by the PRC of CPMS is marked as $V_{sum,k}$, which is the sum of $V_{i,k}$. It is noting that PRC of CPMS will not process any data sample from PIDs.

Computation: As for the computation phase, the aggregation process in FPMS is ignored, as FedAvg (McMahan et al. 2017) is utilized, whose complexity is $O(1)$. Let f_i and f_c denote the CPU frequency of the i^{th} PID and the PRC respectively, and Q represents the training workload per sample. According to the measurement proposed in (Dinh et al. 2020), the computation time and energy consumption of PRC under CPMS, and the i^{th} PID under FMPS in the k^{th} round can be expressed as shown in Equations 1 to 4.

$$T_{k,PRC}^{cmp,CPMS} = \frac{QV_{sum,k}}{f_c} \tag{1}$$

$$E_{k,PRC}^{cmp,CPMS} = \varepsilon_c QV_{sum,k} f_c^2 \tag{2}$$

$$T_{i,k,PID}^{cmp,FPMS} = \frac{QV_{i,k}}{f_i} \tag{3}$$

$$E_{i,k,PID}^{cmp,FPMS} = \varepsilon_c QV_{i,k} f_i^2 \tag{4}$$

Where, ε_c is the computation coefficient correlated to the hardware configuration.

Communication: As for the communication phase, the cost in the downlink is ignored, as it is negligible compared to the cost in the uplink (Dinh et al. 2020). Assume that r_i defines the uploading rate of the i^{th} PID, S_m represents the size of the model parameter to be uploaded, and S_s illustrates the record size of a sample, according to the measurement proposed in (Yu and Li 2021), the communication time T and energy consumption E of CPMS and FPMS in the k^{th} round can be expressed according to following formulas shown in Equations 5 to 8.

$$T_{i,k,PID}^{com,CPMS} = \frac{S_s(V_{i,k} - V_{i,k-1})}{r_i} \tag{5}$$

$$E_{i,k,PID}^{com,CPMS} = \frac{\varepsilon_t S_s(V_{i,k} - V_{i,k-1})}{r_i} 2^{\frac{r_i}{B}-1} \tag{6}$$

$$T_{i,k,PID}^{com,FPMS} = \frac{S_m}{r_i} \tag{7}$$

$$E_{i,k,PID}^{com,FPMS} = \frac{\varepsilon_t S_m}{r_i} 2^{\frac{r_i}{B}-1} \tag{8}$$

Where, ε_t is the communication coefficient, and B is the wireless bandwidth.

As a result, the total time and energy consumption per round is expressed as presented in Equations 9 to 12.

$$T_k^{sum,CPMS} = T_{k,PRC}^{cmp,CPMS} + \max_{i \in n} T_{i,k,PID}^{com,CPMS} \tag{9}$$

$$E_k^{sum,CPMS} = E_{k,PRC}^{cmp,CPMS} + \sum_i^n E_{i,k,PID}^{com,CPMS} \tag{10}$$

$$T_k^{sum,FPMS} = \max_{i \in n} \left\{ (T_{i,k,PID}^{cmp,FPMS} + T_{i,k,PID}^{com,FPMS}) \right\} \tag{11}$$

$$E_k^{sum,FPMS} = \sum_i^n \left\{ (E_{i,k,PID}^{cmp,FPMS} + E_{i,k,PID}^{com,FPMS}) \right\} \tag{12}$$

Finally, the value of the related hyperparameters are defined in Table 2 to simulate a dynamic environment with PIDs, whose computation and communication capabilities vary among each other.

Table 2. The list of hyperparameters used in the evaluation.

Variable	Description	Value
Q	Training workload per sample	10
ε_c	Computation coefficient	1
f_c	CPU frequency of the PRC	0.5-2.5 GHz
f_i	CPU frequency of the i^{th} PID	3 GHz
$V_{i,k}$	Sample size of the i^{th} PID under FPMS	According to the evaluation setting
$V_{sum,k}$	Sample size of the PRC under CPMS	According to the evaluation setting
S_m	Model size	0.01 kb
S_s	Data size per sample	0.07 kb
B	Bandwidth	1 MHz
ε_t	Communication coefficient	50
r_i	Uploading rate of the i^{th} PID under FPMS	$100 - 1000$ kb/s

Evaluation Results. First, the training time consisting of computation and communication time in the first 150 rounds is analyzed for CPMS and FPMS, respectively. Due to the fact that FPMS can utilize the idle resources of each PID for local model training, and requires only the transmission of model parameters through the network, FPMS can maintain a much lower time in computation and communication as shown in Figure 6 (a) and (b), respectively. Moreover, even though the communication time of CPMS overcomes FPMS when the training gets stabilized, FPMS can still significantly reduce the overall training time as illustrated by Figure 6 (c), as the computation time required for the training is much larger than the one for communication.

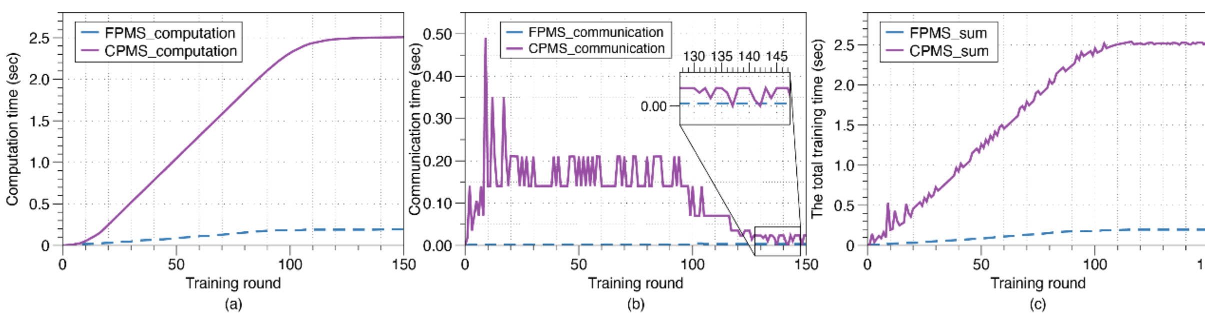

Figure 6. Evaluation results of training time for CPMS and FPMS.

Second, the resource consumption is also analyzed. Similar to the training time, FPMS consumes less computation and communication resources than CPMS as shown in Figure 7 (a) and (b), since CPMS requires the data to be gathered and processed in a centralized manner. Moreover, even though the overall consumption of CPMS decreases as less data are required to be transmitted, FPMS can still remain an obvious advantage in optimizing the utilization of related resources in the service cluster as shown in Figure 7 (c).

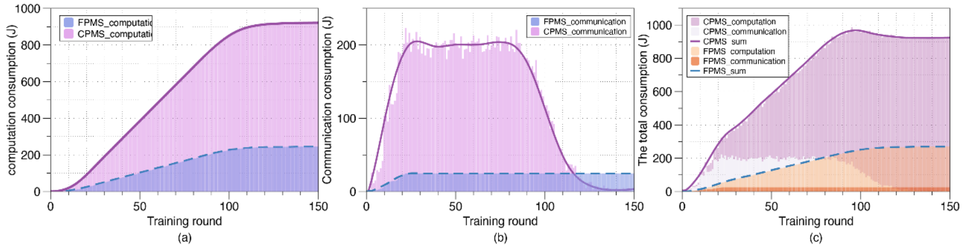

Figure 7. Evaluation results of resource consumption for CPMS and FPMS.

In summary, FPMS outperforms CPMS with significantly reduced training time and resource consumption, which shows its merits in utilizing service resources efficiently and effectively to train a global model with data protected.

CONCLUSION

Intelligent transportation system is evolving under the influence of emerging technologies and diversified demands towards a novel paradigm, called Autonomous Transportation Systems (ATS), which can control system operations and provide user services with less human interventions. This paper discusses the personal mobility service of ATS under the premises of data security and privacy, which incapacitates the centralized PMS (CPMS), and promotes the federated PMS (FPMS) in a decentralized manner. To illustrate how CPMS is transferred to FPMS, a reference architecture is proposed with three views to define the physical objects that participated, the business logic implemented, and the service functions involved in FPMS.

Additionally, through a performance evaluation between CPMS and FPMS, the merits of adopting federated learning in supporting personal mobility in a data-preserving way were analyzed. Compared to CPMS, FPMS can maintain a higher training speed with less resource consumption in terms of communication and computation.

As for the future, FPMS according to the proposed architecture will be implemented with 1) a multi-modal data integration mechanism to accommodate different kinds of data such as public and private data; and 2) an FL-based travel recommendation algorithm to orchestrate various heterogeneous devices in a asynchronous manner.

ACKNOWLEDGEMENTS

This research is supported by National Key Research and Development Program of China (2020YFB1600400), National Natural Science Fund of China (62002398), and National Natural Science Fund of China (41901188).

REFERENCES

Ahvar, E., Orgerie, A.-C., and Lebre, A. (2019). “Estimating energy consumption of cloud, fog and edge computing infrastructures.” *IEEE Transactions on Sustainable Computing, IEEE*, 1–12.

Azevedo, C. L., Seshadri, R., Gao, S., Atasoy, B., Akkinepally, A. P., Christofa, E., Zhao, F., Trancik, J., and Ben-Akiva, M. (2018). “Tripod: sustainable travel incentives with prediction, optimization, and personalization.” *Proceedings of the Transportation Research Record 97th Annual Meeting*.

Bajaj, G., Bouloukakis, G., Pathak, A., Singh, P., Georgantas, N., and Issarny, V. (2015). "Toward enabling convenient urban transit through mobile crowdsensing." *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, IEEE, Gran Canaria, Spain, 290–295.

Bierlaire, M., Axhausen, K., and Abay, G. (2001). "The acceptance of modal innovation: The case of Swissmetro." *Swiss Transport Research Conference*.

Cui, G., Luo, J., and Wang, X. (2018). "Personalized travel route recommendation using collaborative filtering based on GPS trajectories." *International Journal of Digital Earth*, Taylor & Francis, 11(3), 284–307.

Danaf, M., Becker, F., Song, X., Atasoy, B., and Ben-Akiva, M. (2019). "Online discrete choice models: Applications in personalized recommendations." *Decision Support Systems*, 119, 35–45.

Dinh, C. T., Tran, N. H., Nguyen, M. N., Hong, C. S., Bao, W., Zomaya, A. Y., and Gramoli, V. (2020). "Federated learning over wireless networks: Convergence analysis and resource allocation." *IEEE/ACM Transactions on Networking*, IEEE, 29(1), 398–409.

Liu, Y., Yu, J. J. Q., Kang, J., Niyato, D., and Zhang, S. (2020). "Privacy-preserving traffic flow prediction: A federated learning approach." *IEEE Internet of Things Journal*, 7(8), 7751–7763.

Logesh, R., Subramaniyaswamy, V., Vijayakumar, V., Gao, X.-Z., and Indragandhi, V. (2018). "A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city." *Future Generation Computer Systems*, Elsevier, 83, 653–673.

McMahan, B., Moore, E., Ramage, D., Hampson, S., and Aguera y Arcas, B. (2017). "Communication-efficient learning of deep networks from decentralized data." *Artificial Intelligence and Statistics*, 1273–1282.

Qin, Y., Sun, X., Ma, X., Wang, M., Jia, L., and Tang, T. (2019). "Research on the architecture of Intelligent Railway 2.0 and its applications." *Journal of Beijing Jiaotong University*, 43(1), 138–145.

Sun, X., Huang, Z., Peng, X., Chen, Y., and Liu, Y. (2019). "Building a model-based personalised recommendation approach for tourist attractions from geotagged social media data." *International Journal of Digital Earth*, Taylor & Francis, 12(6), 661–678.

Taivalsaari, A., and Mikkonen, T. (2018). "A taxonomy of IoT client architectures." *IEEE Software*, 35(3), 83–88.

Yang, Q., Liu, Y., Chen, T., and Tong, Y. (2019). "Federated machine lLearning: concept and applications." *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19.

Yu, R., and Li, P. (2021). "Toward resource-efficient federated learning in mobile edge computing." *IEEE Network*, 35(1), 148–155.