Federated Learning Meets Urban Opportunistic Crowdsensing in 6G Networks: Opportunities, Challenges, and Optimization Potentials

Wenjun Zhang, Xiaoli Liu, Chao Zhu, Samu Varjonen, Fangxin Wang, and Sasu Tarkoma

Abstract

6G networks are envisioned to enable the Internet of Things (IoTs) and foster ubiquitous sensing. Urban opportunistic crowdsensing, which leverages participants carrying mobile sensing units (MSUs) in their daily activities to collect data, enables low-cost and large-scale urban sensing for applications such as air quality monitoring, pothole detection, and noise classification. However, urban opportunistic crowdsensing poses challenges to conventional cloud-based centralized learning due to the unpredictable nature of crowdsensed data collection and privacy concerns from uploading personal information to the centralized cloud. Federated learning (FL), where MSUs act as data sources and computing nodes, offers a promising alternative to mitigating these issues. Despite FL's potential, urban crowdsensing contexts' spatial-temporal diversity, mobility, constrained resources, and emerging privacy concerns present new challenges. In this paper, we explore the opportunities and challenges of FL in urban opportunistic sensing in 6G networks and suggest potential optimization strategies. Furthermore, we conduct field experiments in Helsinki, Finland, and design an FL-based air quality calibration method for opportunistic crowdsensing to demonstrate the feasibility of our vision.

Introduction

As urbanization progresses, populations are increasingly concentrated in urban areas. It reveals that, as of now, 55% of the world's population lives in urban areas, and a proportion is expected to increase to 68% by the year 2050 [1]. The growing urban population and the behaviors of city residents continuously affect both the urban environment and its ecology, impacting the well-being of inhabitants. Consequently, the continuous monitoring of various urban characteristics and dynamics, such as air quality, construction noise, and traffic congestion, has become a critical issue.

With the development of 6G networks, ubiquitous sensing becomes possible. Among the various sensing methods, opportunistic crowdsensing is an efficient, cost-effective, and scalable approach for urban sensing monitoring [2], [3]. This method leverages mobile devices (e.g., phones, smartwatches, and vehicles) owned by citizens as mobile sensing units (MSUs), taking advantage of the opportunistic encounters of these individuals to perceive, collect, and share sensed common interests (e.g., temperature, moisture, and road condition) via collaborative wireless communication.

Compared to established fixed sensor networks, opportunistic crowdsensing offers greater flexibility and broader coverage, allowing for seamless integration into daily life for real-time dynamic monitoring of the urban environment. Opportunistic crowdsensing therefore not only reduces the cost of monitoring but also guarantees comprehensive coverage across spatial and temporal dimensions of data collection, offering robust support for urban management and decision-making processes.

Previous opportunistic crowdsensing scenarios have adopted methods that involve collecting and processing sensing data in the cloud. However, in urban environments, the end networks for opportunistic crowdsensing are typically provided by cellular networks. Therefore, transmitting massive amounts of sensing data can occupy precious cellular bandwidth and may even overload urban networks. In addition, sensing data often contains not just information about the common interest but also citizens' private information (e.g., location and activity patterns). Thus, uploading raw sensing data to the cloud could likely lead to breaches of user privacy. To tackle these network congestion and privacy concerns, federated learning (FL), introduced by Google in 2017, shifts towards decentralizing computational models by transitioning from reliance on remote data centers to processing directly on user devices [4]. FL leverages the onboard computation resources and performs the model training in a decentralized manner as illustrated in Fig. 1. Instead of uploading private data to a central server, only model weights or updates are communicated. Compared to transmitting raw data, updating models in FL requires considerably less bandwidth and occurs less frequently, ensuring that private data remains locally and

Digital Object Identifier: 10.1109/MNET.2024.3520552
Date of Current Version: 18 March 2025
Date of Publication: 19 December 2024

Wenjun Zhang, Xiaoli Liu (corresponding author), and Sasu Tarkoma are with the Department of Computer Science, University of Helsinki, 00014 Helsinki, Finland; Chao Zhu (corresponding author) is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China; Samu Varjonen is with the Department of Computer Science, University of Helsinki, 00014 Helsinki, Finland, and also with the Finnish Meteorological Institute (FMI), 00101 Helsinki, Finland; Fangxin Wang is with the School of Science and Engineering, The Chinese University of Hong Kong at Shenzhen, Shenzhen 518172, China.

substantially conserves communication resources. Moreover, the inherently distributed manner of FL naturally complements the local awareness of urban sensing, making it a suitable selection for opportunistic crowdsensing efforts.

However, deploying FL for urban opportunistic crowdsensing poses its own practical challenges. One significant challenge stems from the vast spatial and temporal spans of cities, where data collected by distributed MSUs exhibits heterogeneity, posing a challenge to FL in maintaining model coherence across diverse data types and sources. Additionally, the variability in computational resources across MSUs leads to inconsistent training times for high-load data tasks (e.g., videos), making the synchronization of model updates more complex. Finally, the MSUs engaged in opportunistic crowdsensing may be moving at high speeds (e.g., vehicles), which increases the risk of disconnection between MSUs and the regional server before model aggregation is completed. Therefore, minimizing processing and transmission time while ensuring model accuracy also remains a critical challenge.

In this article, we present a comprehensive perspective for utilizing FL in urban opportunistic crowdsensing from the following questions:

- What are the typical scenarios of urban opportunistic crowdsensing? What opportunities would be observed when utilizing FL to them?
- What are the challenges of deploying FL in urban opportunistic crowdsensing, particularly in terms of spatial-temporal, mobility, resource-constrained, and privacy-sensitive characteristics?
- What are the potentials of optimization for FL to ensure accuracy, reliability, and privacy preservation?

The remainder of the article is organized as follows. First, we investigate urban sensing scenarios and FL's opportunities for their implementation. Then, we introduce the optimization potentials of FL to tackle practical challenges in urban opportunistic crowdsensing. Following that, we present a case study of FL based calibration method on air quality monitoring through field experiments in Helsinki, Finland. Finally, we summarize the article.

Scenarios and Opportunities

Opportunistic crowdsensing has emerged as a prominent method for urban sensing and attracted considerable interest. This section begins by presenting the detailed profiles of several urban opportunistic crowdsensing applications, encompassing key aspects such as MSU mobility, update frequency, computing sensitivity, privacy sensitivity, and participant scale, as illustrated in Table 1. Following that, we explore the opportunities offered by FL's novel capabilities in those applications' implementations in detail.

Environmental Sensing

Maintaining a healthy urban environment is pivotal for the well-being and sustainability of urban populations. For instance, it has been reported that air pollution is responsible for killing more than 7 million people prematurely every year [5].

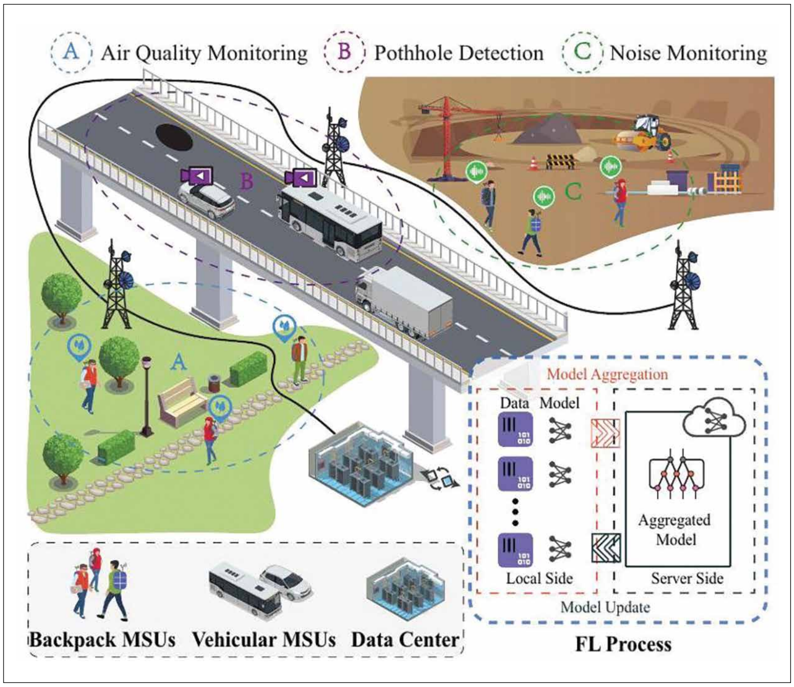

FIGURE 1. Urban opportunistic crowdsensing with FL.

Compared to transmitting raw data, updating models in FL requires considerably less bandwidth and occurs less frequently, ensuring that private data remains locally and substantially conserves communication resources.

Urban environmental sensing, covering air quality monitoring, noise classification, and disaster awareness, entails deploying city-wide sensors to capture the dynamic nature of urban ecosystems effectively. It plays a vital role in sustainable urban development and enables critical decisions in urban resource allocation (e.g., green space planning).

To monitor the urban environment, one method involves the establishment of reference stations (RSs) with infrastructures and precision sensors. Nonetheless, the deployment of RSs requires considerable construction and maintenance expenses. Alternatively, a more cost-effective method is employing crowdsensing with portable and low-cost MSUs [6]. These MSUs typically comprised of batteries and affordable sensing elements, are small and light enough to fit in a backpack. By distributing these portable sensing units to volunteers who carry them during their daily urban commutes, such as working or shopping, a more dynamic urban environmental sensing network can be created.

However, low-cost MSUs are usually energy-constrained and face accuracy challenges due to less powerful sensors and processors, lower sampling rates, and energy-saving measures. Meanwhile, data drift further affects reliability. To improve the accuracy of widespread sensing, it is crucial to employ calibration algorithms [7] to adjust sensor measurements. This method requires target measurements from high-accuracy RSs to build the model and calibrate the measurements from MSUs during their opportunistic encounters to adjust the bias and drift.

	MSU mobility	Update frequency	Computing sensitivity	Privacy sensitivity	Participant scale
Air quality monitoring	Low	Med	Low	Low	Large
Noise classification	Low	High	Med	High	Large
Disaster awareness	Low	High	High	Med	Low
Tourism recommendation	Low	Low	Med	High	Low
Traffic monitoring	High	High	High	Med	Large
Infrastructure monitoring	Med	Low	High	Med	Med

TABLE 1. Urban opportunistic crowdsensing applications profiles.

As previously mentioned, cloud computing faces network congestion and privacy concerns. Additionally, it is well understood that a dataset from a single MSU is insufficient to train a robust model. In light of these issues, FL enables MSUs to calibrate locally without uploading raw data, thereby significantly reducing network resource demands and effectively addressing privacy, network congestion, and measurement accuracy concerns.

Tourism Recommendation

Recommendation systems have emerged as a key technology in navigating the vast choices in urban tourism, such as pinpointing the best sights, eateries, and shopping areas. By capturing and predicting user preferences, these systems can effectively mitigate information overload, ensuring travelers can easily find options that match their tastes and interests.

To enhance the performance of recommendation system, extensive data collection is essential. Crowdsensing significantly contributes to this effort by leveraging widespread users' engagements. Utilizing social media platforms on MSUs, it efficiently compiles a comprehensive dataset from public contributions. By analyzing this collective information, recommendation systems can refine their algorithms to offer more precise, contextually relevant suggestions tailored to user preferences and current trends, thereby enhancing the overall effectiveness of urban tourism recommendations.

However, data from social media platforms often contain a significant amount of private information about users, such as user attributes, behaviors, social relations, and context information. Uploading these data to the cloud can pose serious risks, such as illegally selling it to third parties or being targeted by malicious attacks. As regulations become increasingly stringent, collecting user data through application backends for cloud-based training has become untenable. On the other hand, the FL-based crowdsensing scheme allows participants to collaboratively train a machine learning (ML) model by uploading their weights or updates rather than raw data to the edge server during their opportunistic encounters via wireless communication and then sending them to the server via bandwidth, emerges as a promising approach for developing privacy-preserving recommendation systems.

Despite avoiding transmitting raw user data, FL-based recommendation systems still encounter several challenges. Privacy concerns persist since adversaries might infer users' preferences from the shared parameters. Additionally, variations in MSUs' capabilities and data distribution significantly affect recommendation accuracy and efficiency. Given FL's critical importance for privacy-compliant recommendation systems, tackling these challenges is key to unlocking its potential and ensuring its future viability.

Visual Crowdsensing

Images or videos encompass a richer array of content compared to time series sensor data. By employing powerful artificial intelligence algorithms, a vast amount of information can be extracted from images or videos.

Compared to numerical data, visual data encompasses a richer array of content. By employing powerful ML models, a vast amount of information can be extracted from visual data. Consequently, visual crowdsensing, which harnesses the power of smart devices' built-in cameras coupled with advanced image processing techniques, has emerged as the leading approach to capturing detailed and extensive information about specific subjects of interest.

The integration of advanced image processing techniques with visual crowdsensing facilitates the real-time monitoring and analysis of phenomena such as intelligent transportation, infrastructure maintenance, and public safety. In intelligent transportation, traffic congestion monitoring can be identified by processing visual data crowdsourced from dash cameras [8]. Within infrastructure maintenance monitoring, damages to urban infrastructure, such as the absence of road signs, the toppling of street lights, and potholes, can be swiftly spotted, thereby accelerating the repair process. For public safety, emergencies or criminal activities can be quickly detected in video feeds, speeding up response efforts.

A common approach to enhance the precision of image processing models is to transmit all the raw visual data generated from the MSUs to the cloud server for model training. However, video data is larger in data item size, and uploading this data raises pertinent concerns regarding communication costs and the potential to overwhelm urban networks. An alternative approach is to leverage FL, enabling data to remain stored on MSUs while models are trained locally without transferring raw data. Additionally, each MSU's model is updated by integrating parameters uploaded by each unit to the server.

However, visual data faces several unique issues, such as multi-objectivity, data redundancy, and high data processing costs. Efficiently processing visual data under the FL mechanism with limited computational resources on MSUs remains a challenge to be addressed.

Challenges and Optimization Potentials

The upcoming 6G network's massive connectivity and high-speed transmission capabilities are bringing profound changes to urban opportunistic crowdsensing. With the support of 6G's massive connectivity protocols, any urban entity (such as cars, smartwatches, streetlights, etc.) can become part of the opportunistic crowdsensing ecosystem, contributing to the sharing of data and greatly enriching data sources. On the other hand, the low-latency, high-speed transmission of the 6 G network enhances the data dimensions, making it possible for crowdsensing high-dimensional data such as video, audio, and radar data, paving the way for future city metaverse applications based on AR, VR, and other immersive technologies.

However, the transformative nature of 6 G for urban opportunistic crowdsensing also brings challenges to the application of FL. For example, data collected by MSUs across different times and locations leads to data heterogeneity, resource-constrained MSUs may produce models with lower accuracy, and different MSUs have varying requirements for privacy protection. Additionally, there are challenges related to multimodal data fusion and multi-task training. In this section, we explore potential optimization strategies to address the practical challenges mentioned above, making the implementation of FL in urban opportunistic crowdsensing more efficient and reliable.

Clustering

In urban opportunistic crowdsensing in a 6 G environment, MSUs (e.g., vehicles) exhibit a diverse spatiotemporal movement throughout the city. This dispersion results in highly heterogeneous collected data. For instance, vehicles at intersections often capture more pedestrians than those on highways.

Particularly, in FL, where end users' devices act as both data sources and computing nodes, data heterogeneity becomes an inherent challenge. This heterogeneity notably impacts FL by causing a drift in each client's updates, which leads to slow and unstable model convergence. Remarkably, this client drift persists even when full batch gradients are utilized and all clients participate consistently throughout the training process [9].

Leveraging clustering techniques [10] within FL significantly improves the management of data heterogeneity. As illustrated in Fig. 2, training distinct models for groups of MSUs with similar data distribution makes the entire process more customized and efficient for each group. For example, clustering could organize vehicles into clusters based on their location environment. Vehicles at intersections might form one cluster, where the focus could be on training models to recognize dangerous pedestrian behaviors, such as jaywalking or crossing against the traffic signal. Conversely, highway vehicles could constitute another cluster identifying risky vehicular actions like unsafe lane changes or speeding. This clustering strategy enables the deployment of specialized models meticulously calibrated to meet specific requirements, thereby enhancing precision and fostering effective decision-making across diverse scenarios.

Knowledge Transfer

In the 6G network, transmission units such as antennas and basebands will become miniaturized, and many resource-constrained devices (e.g., smartwatches) will be involved in urban crowdsensing. On the other hand, in most ML models, beginning training with randomly initialized parameters and progressing through hundreds to thousands of iterations to achieve parameter stability is a widely adopted method. Therefore, it is almost unrealistic to train a model with high robustness from randomly initialized parameters in a short duration.

Transfer learning [11] is an ML method that enables the application of a model, already trained on a certain task or dataset, to a related but different task or new dataset. During the transition to a new task, fine-tuning is often employed, which involves adjusting some of the model's parameters to better fit the new task's data. This method is particularly suited for scenarios with limited labeled data or computational resources, as it leverages existing knowledge from models pre-trained on large datasets to avoid training a model from scratch, accelerating the model training process and enhancing the model's performance on specific tasks, especially when the new task's data volume is relatively small or differs from the original training data. Through transfer learning, we can adapt to new data and tasks more quickly and effectively without sacrificing model performance.

Consequently, integrating knowledge transfer into FL can significantly enhance the performance of urban opportunistic crowdsensing. For example, with their numerous sensors and powerful GPUs, smart vehicles like Teslas can leverage large datasets to train robust pre-trained models. Subsequently, standard vehicles, equipped only with dash cameras and basic CPUs, can fine-tune these models through FL among themselves instead of developing models from scratch.

This heterogeneity notably impacts FL by causing a drift in each client's updates, which leads to slow and unstable model convergence.

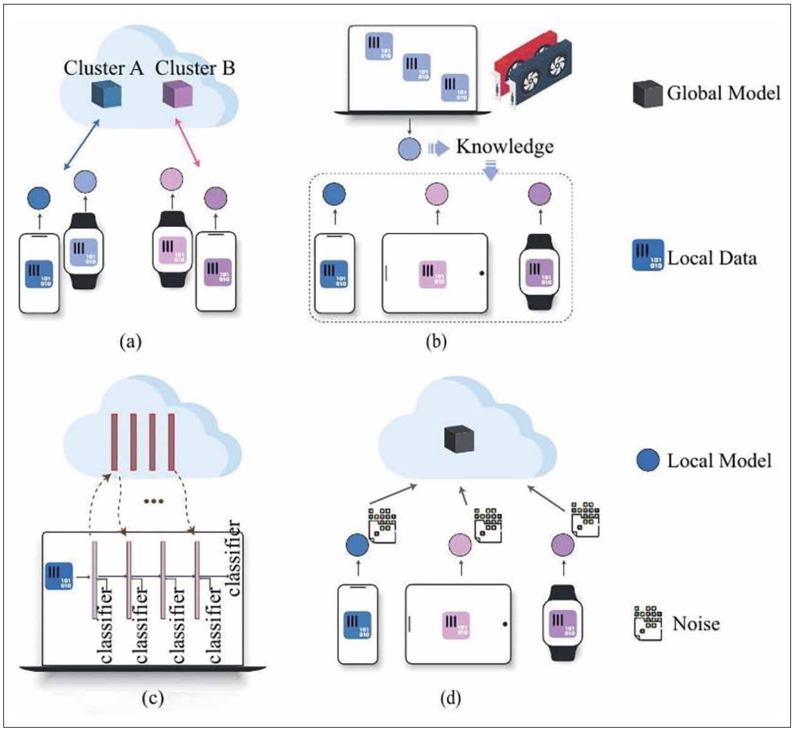

FIGURE 2. Optimization potentials of FL. a) Clustering. b) Knowledge Transfer. c) Layer Decoupling. d) Local Differential Privacy.

Layer Decoupling

In 6G networks, device-to-device (D2D) technology is expected to advance rapidly. For instance, cellular-assisted D2D communication enables two or more devices to communicate directly under the coordination of the cellular network without relying on the base station for data routing. In urban opportunistic crowdsensing, MSUs may move quickly relative to each other, potentially breaking their D2D communication links as distances increase. Although FL avoids the need to send large amounts of raw data, transmitting entire large neural network models (e.g., the size of the YOLO v5 model can reach up to a few dozen megabytes) also carries the risk of communication disruptions before the transmission is complete.

Layer-by-layer decoupling was initially developed to help understand and optimize the complex interactions between layers in deep architectures [12]. Its fundamental principle involves training one layer at a time, allowing each layer to stabilize and capture relevant features before moving on to the next. In urban opportunistic crowdsensing, this layer-by-layer decoupling can also reduce the size transmitted to ensure timely transmission. Although the training outcomes of layer-by-layer models might not match the performance of training the whole model at once, trade-offs between latency and effectiveness in different mobility scenarios can be achieved by carefully adjusting the number of decoupled layers. Besides, layer-wise decoupling also segments large training tasks into smaller ones. Within a multitasking environment, these smaller tasks can be executed in parallel across MSUs via multithreading, significantly enhancing the system's throughput.

Differential Private Techniques

As illustrated in Table 1, many urban sensing applications involve acquiring data related to the users themselves, such as voice, portrait, and location details. Although the FL eliminates the need for direct transmission of raw data, it still poses a risk of privacy breaches through model update leakage. Model update leakage jeopardizes FL security by exploiting flaws in the transfer of model gradients or weights. For example, in a noise monitoring application, an attacker can reveal phrases that have been used in other participants' training sets by inverting the real-valued vector representations.

In response to the privacy concerns, researchers have suggested enhancing the security of FL with complementary techniques such as secure multi-party computation (SMC) and homomorphic encryption (HE), which are cryptography-based approaches that typically necessitate advanced encryption computing protocols. However, the iterative communication and transmission of encrypted data between MSUs and the server results in significant communication and computation costs. In comparison, the differential privacy (DP) [13] technique is a much more cost-effective privacy protection method by adding random noise to the data or model weights, ensuring individual privacy in crowdsensing scenarios.

Among existing works, Local Differential Privacy (LDP) [14] possesses inherent local awareness that each data contributor adds noise to their local data before sharing it, preserving privacy at the source. In urban opportunistic crowdsensing, LDP can effectively mask the MSUs' specific data details. For example, in a noise monitoring scenario, LDP would allow each participant to add a certain amount of noise to the original audio data, ensuring that the exact contribution of each device remains obscured, protecting personal privacy. Nevertheless, in 6G networks, the massive number of MSUs may have varying privacy requirements, and in FL, privacy protection and model performance often affect each other. Therefore, how to provide personalized privacy protection for MSUs while balancing privacy and model performance is still a key research question in LDP-enabled FL.

A Case Study: Federated Learning Based Calibration on Air Quality Monitoring

In this section, we introduce a strategy that leverages opportunistic crowdsensing for air quality monitoring by using affordable, portable sensors as MSUs. Additionally, we design a calibration algorithm based on FL to enhance the sensing accuracy of these portable sensors. Through field experiments, we attempt to illustrate the effectiveness of FL in urban opportunistic crowdsensing.

Sensor Hardware and Deployment

In the air quality field experiments, we designed and used our own MSU built around a Nordic Semiconductors nrf52840 System on a Chip. Communications are facilitated by a cellular phone using an LTE connectivity. For uplink transmission, the cellular phone uses SC-FDMA modulation, achieving upload speeds of up to 250 kbit/s. Due to its single-carrier characteristic, SC-FDMA has a lower Peakto-Average Power Ratio (PAPR), making it highly advantageous for future 6 G scenarios that involve massive connectivity of resource-constrained IoT devices. Each MSU in our experiment comprises multiple sensors shown in Fig. 3(a). Furthermore, to eliminate any potential data variation due to equipment discrepancies, we deployed all MSUs from the same batch with identical components. The Sensirion SPS30 sensor collects PM_1, $PM_{2.5}$, PM_4, and PM_{10}, the MiCS-4514 sensor collects CO and NO_2, the MQ-131 sensor gathers O_3, and the BME-288 sensor is responsible for collecting temperature (T), relative humidity (RH), and pressure (P). The Grimm 180/FH 62 I-R sensor installed on RS SMEAR III, managed by the Institute for Atmospheric and Earth System Research (INAR) and situated on the Kumpula Campus of the University of Helsinki, Finland, is used to collect $PM_{2.5}$ as reference measurement.

We conduct our opportunistic crowdsensing scenario using three MSUs from 19^{th} to 31^{st} July 2022. Fig. 3(c) and (d) show their encounters with SMEAR III. The MSUs, labeled with A, B, and C, collect data approximately every three minutes. Note that the three MSUs began data collection at different starting timestamps, resulting in different samples across each unit. However, these sensors collecting O3 data malfunction, leading to data loss. Ultimately, the remaining measurements

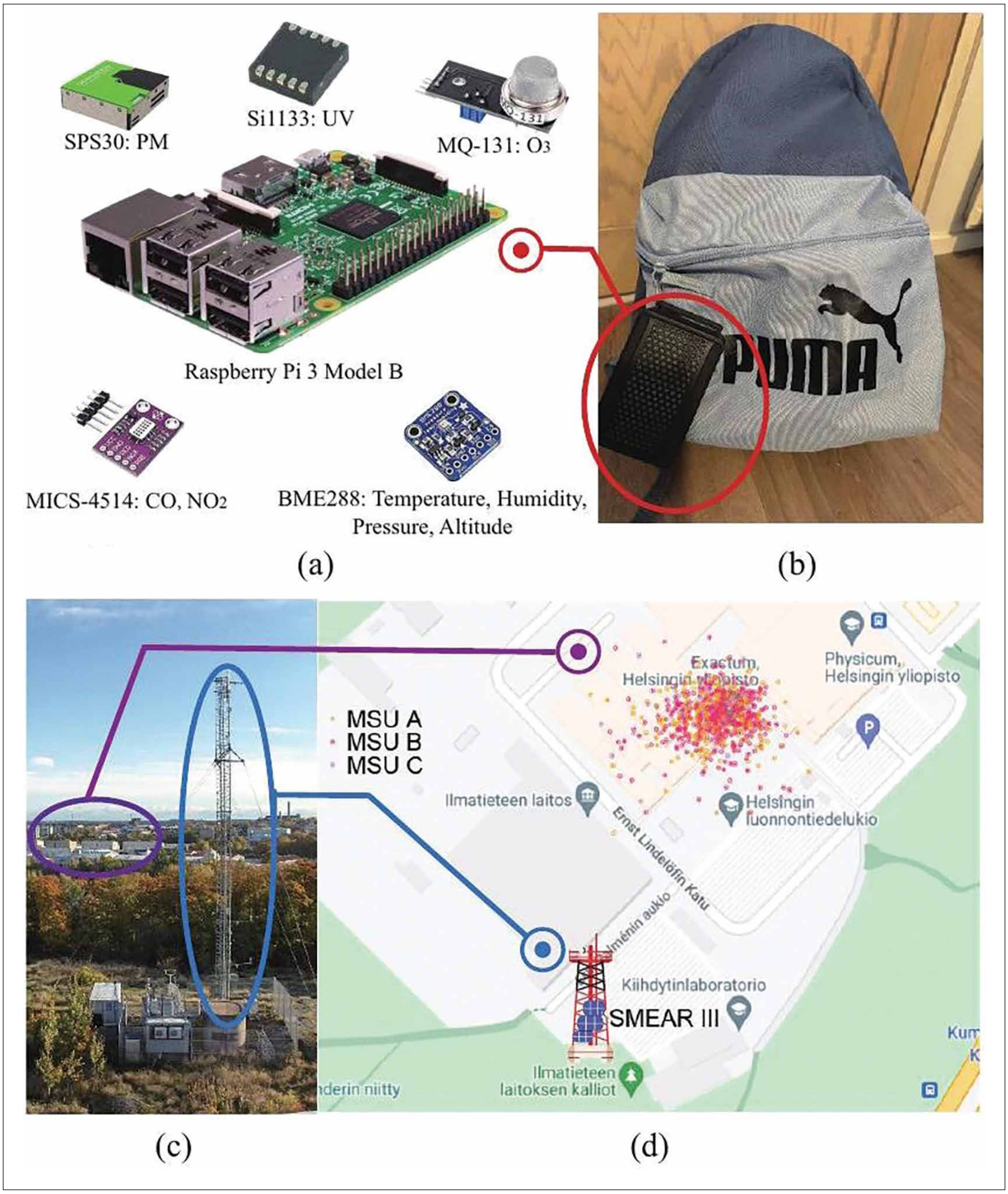

FIGURE 3. MSUs deployment for air quality monitoring. a) Low-Cost MSU Components. b) MSU on Backpack. c) Reference Station. d) Field Experiments.

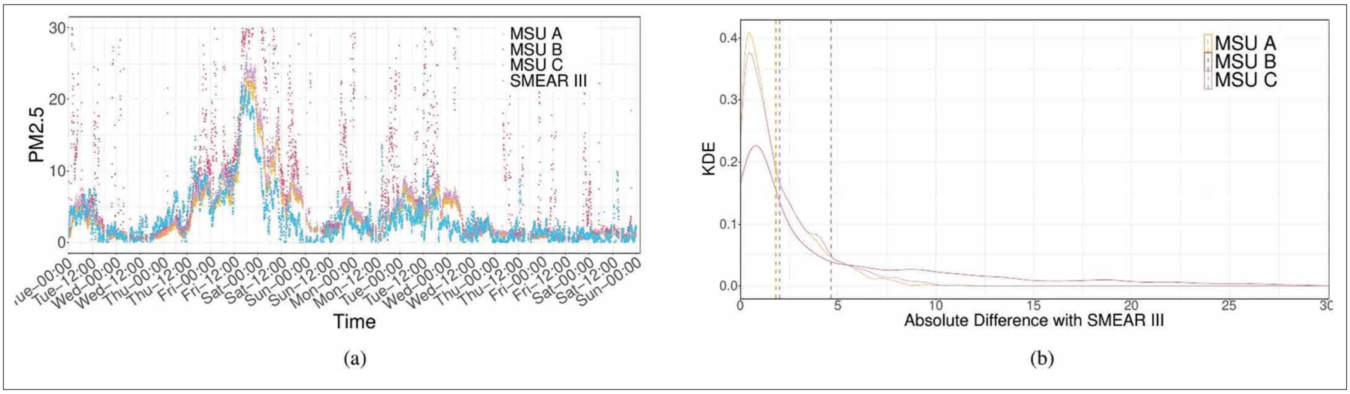

FIGURE 4. $PM_{2.5}$ detected by different MSUs: a) $PM_{2.5}$ measurements of MSUs and SMEAR III; and b) errors of different portable MSUs with SMEAR III before calibration.

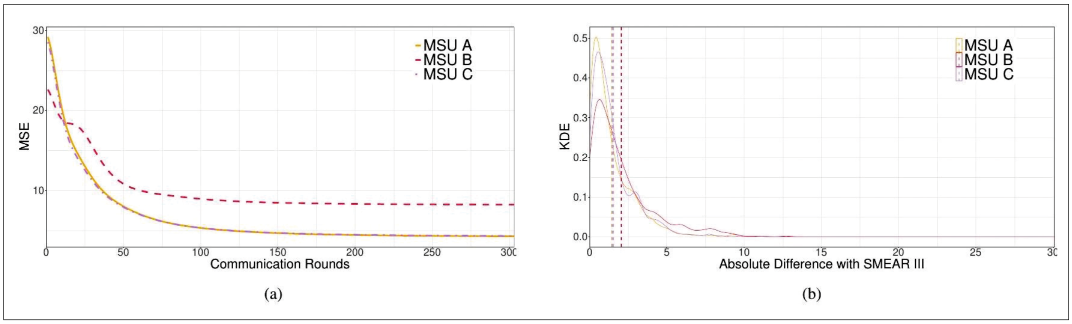

FIGURE 5. FL-based calibration results: a) the converging curves of MSE in iteration procedure; and b) errors of different MSUs with SMEAR III after calibration.

Through our analysis and experiments, we have validated that FL can effectively train a robust calibration model through multi-party collaboration from MSUs, enhancing the accuracy of all participants.

from MSUs which are strongly correlated with the PM2.5, namely *T, RH, P, CO*, NO_2, PM_1, $PM_{2.5}$, PM_4, PM_{10}, are designated as variables in our calibration model for all three MSUs. To train the model, we collect the $PM_{2.5}$ from SMEAR III for the same time period, with data being gathered at one-minute intervals, as model targets. Therefore, even though our MSUs collect data every three minutes and have different data samples, they can still extract their corresponding targets from the RS, which collects data every minute.

Federated Learning Based Calibration

To implement the calibration model leveraging the FL, we designate three MSUs as client nodes. Each client independently trains a local model using Multiple Linear Regression (MLR), a method previously applied by Badura et al. [15] for the calibration of the PMS7003 Particulate Matter ($PM_{2.5}$) sensors. Upon completion of local training, these clients transmit their model updates to the central server. For the aggregation of these local models, we employ the Federated Averaging (FedAvg) algorithm, as introduced by McMahan et al. [4]. This server-side aggregation process culminates in creating an enhanced, unified model, which is subsequently distributed back to clients for further iterations. Our experimental framework encompasses 300 communication rounds to ensure thorough convergence and model refinement. The optimal model w minimizes the empirical risk defined as:

$$\min_{w} \left\{ F(w) \triangleq \sum_{c=1}^{|C|} p_c \left(\frac{1}{|D_c|} \sum_{\eta \in D_c} f\left(w;(x,y)\right) \right) \right\} \quad (1)$$

Where $|C|$ is the number of clients; for each client c with a local training dataset D_c consisting of $|D_c|$ samples, the proportion p_c is calculated as $|D_c|$ divided by the total number of samples across all clients $\sum_{C=1}^{|C|} |D_C|$; and $f(w; (x,y))$ is the loss function of each client for quantifying the prediction error of model w for the sample $\eta = (x, y)$.

Experiment Results

To more directly compare the effectiveness of applying VFL on air quality calibration, we first conduct a data analysis on the original data. Fig. 4(a) presents the measurements over time of the raw $PM_{2.5}$ data collected by each MSU, compared with SMEAR III. Evidently, MSU B's raw data deviates from the values recorded by SMEAR III, whereas the distributions of MSU A and C remain relatively similar to SMEAR III. Fig. 4(b) presents a Kernel Density Estimation (KDE) plot comparing the absolute differences in $PM_{2.5}$ measurement between three MSUs and SMEAR III. Notably, the plot clearly shows that MSU B has a larger deviation from the SMEAR III measurement, with its mean absolute difference centered around 5 $\mu g/m^3$, whereas the other two MSUs maintain a mean difference closer to 2 $\mu g/m^3$. Additionally, the peak sharpness of the three KDE plots indicates that MSUs A and C have a more concentrated distribution of absolute differences with SMEAR III, implying lower variability. In contrast, MSU B exhibits larger disparities.

Fig. 5(a) displays the loss convergence for three MSUs under the VFL framework, with 300 communication rounds conducted using the Mean Squared Error (MSE) loss function. It is clearly observed that the loss of each MSU gradually decreases and stabilizes as communication rounds increase. Fig. 5(b) presents the KDE plot that vividly demonstrates the efficacy of the VFL framework in reducing the absolute difference in $PM_{2.5}$ between each MSU and the SMEAR III. Notably, the mean absolute differences for MSU A and C decrease from 2 to 1.5 $\mu g/m^3$, and more remarkably, for MSU B, from 5 to 2 $\mu g/m^3$, compared with Fig. 4(b). This substantial improvement, especially for MSU B , underscores the potential of leveraging higher-quality MSU data to enhance calibration for units with initially inferior data quality. Furthermore, the KDE plots in Fig. 5(b) exhibit increased peak sharpness and a closer convergence towards 0 for all three MSUs, starkly contrasting to those depicted in Fig. 4(b). This shift is particularly pronounced for MSU B, highlighting a significant reduction in absolute deviations and indicating an overall improvement in data accuracy.

Conclusion

In this article, we explore the applications, opportunities, and challenges associated with implementing FL in urban opportunistic crowdsensing in 6G networks. To demonstrate the feasibility of our concept, we use an air quality monitoring application as a case study and present the results of small-scale field experiments conducted with three portable, low-cost MSUs. Through our analysis and experiments, we have validated that FL can effectively train a robust calibration model through multi-party collaboration from MSUs, enhancing the accuracy of all participants. In addition, we train models on MSUs locally and transmit only weight updates to the server for aggregation. This approach safeguards against information leakage, as opposed to transmitting raw data, effectively mitigating two major challenges in opportunistic crowdsensing: substantial bandwidth usage and the risk of privacy violations.

As public awareness of privacy protection increases, FL is poised to play an even more significant role in urban crowdsensing. In the future, with the integration of intelligent agents like autonomous vehicles, the number of devices participating in urban crowdsensing will grow exponentially. To address this, future research could focus on optimizing groupings of participating devices (e.g., vehicles or sensors) based on factors such as geographical location, mobility patterns, and resource capacity. Additionally, Knowledge Distillation (KD) techniques can be employed to facilitate knowledge transfer between models, where smaller "student" models are trained using insights from "teacher" models within different clusters-without the need to share raw data. Furthermore, combining layer decoupling with differential privacy will help reduce the computational load on devices while protecting user data and ensuring privacy during model transmission.

References

[1] World Health Org. (2023). *Urban Health*. Accessed: Feb. 5, 2024. [Online]. Available: https://www.who.int/health-topics/urban-healthtab=tab_1

[2] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, Nov. 2011.

[3] N. H. Motlagh et al., "Quality of monitoring for cellular networks," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 381–391, Mar. 2022.

[4] B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *Proc. Artif. Intell. Statist.*, 2017, pp. 1273–1282.

[5] World Health Org. (2023). *Household Air Pollution and Health*. Accessed: Feb. 5, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health

[6] S. Tarkoma et al., "MegaSense: 5G and AI for air quality monitoring," in *Proc. 5th EAI Int. Summit, SmartCity360*, Braga, Portugal. Cham, Switzerland: Springer, Dec. 2019, pp. 13–19.

[7] X. Liu et al., "Estimating black carbon levels with proxy variables and low-cost sensors," *IEEE Internet Things J.*, vol. 11, no. 10, pp. 17577–17588, May 2024.

[8] C. Zhu et al., "Vehicular fog computing for video crowdsourcing: Applications, feasibility, and challenges," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 58–63, Oct. 2018.

[9] S. P. Karimireddy et al., "SCAFFOLD: Stochastic controlled averaging for federated learning," in *Proc. 37th Int. Conf. Mach. Learn.*, vol. 119, Jul. 2020, pp. 5132–5143.

[10] A. Ghosh et al., "An efficient framework for clustered federated learning," *IEEE Trans. Inf. Theory*, vol. 68, no. 12, pp. 8076–8091, Jul. 2022.

[11] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2009.

[12] T. Bengio et al., "Greedy layer-wise training of deep networks," in Advances in Neural Information Processing Systems, vol. 19, B. Schölkopf, J. Platt, and T. Hoffman, Eds., MIT Press, 2006. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2006/file/5da713a690c067105aeb2fae32403405-Paper.pdf

[13] C. Dwork, "Differential privacy," in *Proc. Int. Colloq. Autom. Lang. Program.* Berlin, Germany: Springer, 2006, pp. 1–12.

[14] Ú. Erlingsson, V. Pihur, and A. Korolova, "RAPPOR: Randomized aggregatable privacy-preserving ordinal response," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Nov. 2014, pp. 1054–1067.

[15] M. Badura et al., "Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements," *Social Netw. Appl. Sci.*, vol. 1, no. 6, pp. 1–11, Jun. 2019.

Biographies

Wenjun Zhang (wenjun.zhang@helsinki.fi) is currently pursuing the Ph.D. degree with the Department of Computer Science, University of Helsinki, Finland. Her research interests focus on distributed machine learning, artificial intelligence, and information privacy. She is also active in mobile and ubiquitous computing and environmental sensing and monitoring.

Xiaoli Liu (xiaoli.liu@helsinki.fi) is currently a Research Coordinator with the Department of Computer Science, University of Helsinki, Finland. Her research interests include federated learning, edge intelligence, and augmented reality.

Chao Zhu (chao@bit.edu.cn) received the Ph.D. degree from the Department of Computer Science, Aalto University, Finland, in 2021. He is currently an Assistant Professor with the School of Cyberspace Science and Technology, Beijing Institute of Technology, China. His research interests include mobile edge computing and vehicular networking.

Samu Varjonen (samu.varjonen@helsinki.fi) is currently a Senior Research Scientist with the Finnish Meteorological Institute (FMI) and a Docent contract with the Department of Computer Science, University of Helsinki, Finland. His research interests include the Internet of Things, overlay networks, and sensors.

Fangxin Wang (wangfangxin@cuhk.edu.cn) received the B.Eng. degree in computer science and technology from the Beijing University of Posts and Telecommunications, the M.Eng. degree in computer science and technology from Tsinghua University, and the Ph.D. degree in computer science and technology from Simon Fraser University. He is currently an Assistant Professor and a Presidential Young Fellow with The Chinese University of Hong Kong at Shenzhen (CUHKSZ). His research interests include multimedia systems and applications, cloud and edge computing, and distributed networking and systems. He was selected into the Stanford World's Top 2% Scientists List 2023. He served as the TPC Chair for IEEE Satellite 2023; a TPC Member for IWQoS, ICC, and BigCom; and a reviewer for many top conferences and journals. He is an Associate Editor of IEEE Transactions on Mobile Computing.

Sasu Tarkoma (Senior Member, IEEE) (sasu.tarkoma@helsinki.fi) received the Ph.D. degree computer science from the University of Helsinki in 2006. He is the Dean of the Faculty of Science and a Full Professor at the Department of Computer Science, University of Helsinki, Finland. He is also the Director of the Helsinki Center for Data Science (HiDATA). His research interests include distributed systems, data analytics, mobile and ubiquitous computing, artificial intelligence, and 6G.