Zhou, Xiaokang; Liang, Wei; Ma, Jianhua; Yan, Zheng; Wang, Kevin I-Kai

2D Federated Learning for Personalized Human Activity Recognition in Cyber-Physical-Social Systems

Published in:
IEEE Transactions on Network Science and Engineering

DOI:
10.1109/TNSE.2022.3144699

E-pub ahead of print: 01/01/2022

Document Version
Peer reviewed version

Please cite the original version:
2D Federated Learning for Personalized Human Activity Recognition in Cyber-Physical-Social Systems

Xiaokang Zhou, Member, IEEE, Wei Liang, Member, IEEE, Jianhua Ma, Member, IEEE, Zheng Yan, Senior Member, IEEE, and Kevin I-Kai Wang, Member, IEEE

Abstract—The proliferation of the Internet of Things (IoT), wearable computing, and social media technologies bring forward the realization of the so-called Cyber-Physical-Social Systems (CPSS), which is capable of offering intelligent services in many different aspects of our day-to-day life. While CPSS offer a wide variety of data from wearable devices, challenges such as data silos and secure data sharing still remain. In this study, a 2-Dimensional Federated Learning (2DFL) framework, including the vertical and horizontal federated learning phases, is designed to cope with the insufficient training data and insecure data sharing issues in CPSS during a secure distributed learning process. Considering a specific application of Human Activity Recognition (HAR) across a variety of different devices from multiple individual users, the vertical federated learning scheme is developed to integrate shareable features from heterogeneous data across different devices into a full feature space, while the horizontal federated learning scheme is developed to effectively aggregate the encrypted local models among multiple individual users to achieve a high-quality global HAR model. A computationally efficient somewhat homomorphic encryption (SWHE) scheme is then improved and applied to support the parameter aggregation without giving access to it, which enables heterogeneous data sharing with privacy protection across different personal devices and multiple users in building a more precise personalized HAR model. Experiments are conducted based on two public datasets. Comparing with three conventional machine learning methods, evaluation results demonstrate the usefulness and effectiveness of our proposed model in achieving faster and smoother convergence, with better precision, recall, and F1 scores for HAR applications in CPSS.

Index Terms—Federated Learning, Convolutional Neural Networks, Human Activity Recognition, Privacy Protection, Cyber-Physical-Social Systems

1 INTRODUCTION

One of the key requirements for intelligent systems is to be able to offer services according to human context. The ability to identify human activities correctly is of critical importance in offering context-aware and intelligent services. While technologies such as mobile computing, wearable computing, and ambient intelligence are available in our surrounding environments, the ability of recognizing human activities and offering intelligent services accordingly is still very limited.

To date, Human Activity Recognition (HAR) is still facing many challenges, such as lack of sufficient training data, computationally intensive learning, risk of sensitive information leakage, and lack of personalized models [1-3]. The concept of Cyber-Physical-Social Systems (CPSS), which integrates cyber-physical systems, Internet of Things (IoT) technologies, and social media technologies, offers an opportunity for addressing some of these challenges. For example, through collecting heterogeneous CPSS data, a more complete and accurate human activity model can be trained. However, many challenges still remain: First, with the proliferation of CPSS, it also comes with data heterogeneity [4]. How to break the data silos and make use of heterogeneous data to realize a more complete human activity model needs to be addressed. Second, different types of data are very likely to be collected from different devices of different vendors. These data may be highly sensitive to an individual user and should not be shared with another user, nor across his/her own personalized devices for security reasons [5]. For example, a person may not necessarily want to share his/her vital sign monitoring data (e.g., heart rate) with a social media application on the smartphone.

Therefore, it is important to investigate a new approach that can manage heterogeneous data across multiple personal devices to break the data silos, while supporting knowledge sharing and distributed learning among multiple users to achieve better modelling of human activities. The solution proposed in this paper is referred to as the 2-Dimensional Federated Learning (2DFL) that consists of the...
concepts of both horizontal and vertical federated learning, enabling the use of the widespread smartphone, wearable devices, ambient devices, and social media to acquire heterogeneous personal activity data. A vertical federated learning scheme is designed to ensure that sensitive data collected by different devices of different vendors are kept safe and uncompromised by sharing with a trusted third party. This enables heterogeneous data to be used to train a more complete HAR model. The local HAR model is then enhanced by the horizontal federated learning which supports aggregation of personal human activity models to achieve a more accurate global model. The global model will be shared with individual users iteratively to improve their corresponding local models. Specifically, contributions of the proposed 2DFL model can be summarized as follows.

i) A 2DFL framework, which takes the advantage of both the vertical and horizontal federated learning, is designed to overcome the limitation of insufficient training data and insecure data sharing in CPSS.

ii) Two federated learning schemes, including a vertical federated learning scheme to integrate sharable features from heterogeneous data into a full feature space across multiple personal devices, and a horizontal federated learning scheme to effectively aggregate the encrypted local models among multiple individual users, are developed to facilitate the training of a high-quality global model for HAR during a secure distributed learning process.

iii) A computationally efficient somewhat homomorphic encryption (SWHE) scheme is improved and applied in our 2DFL to support the parameter aggregation without giving access to it, which enables heterogeneous data sharing with privacy protection across multiple personal devices and multiple users in building a more precise personalized HAR model.

The rest of the paper is structured as follows. Section 2 summarizes the existing related works of HAR, federated learning and CPSS. The proposed methodology is described in Section 3, followed by the experimental evaluation in Section 4. Section 5 concludes the paper with promising future works.

2 RELATED WORK

Several significant issues relating to the topics discussed in this paper, including existing researches on HAR, studies on federated learning techniques, and intelligent applications based on CPSS, are summarized respectively.

2.1 Researches on HAR with Smart Devices

HAR is a key technique in monitoring human-related behaviors among smart sensor enhanced applications, including well-being management, online medical diagnosis, and smart surveillance, especially when facing heterogeneous data integrated with hybrid contexts across CPSS.

Specifically, smartphone, wearable devices, and IoT sensors, have been widely employed for HAR related analysis in smart computing environments. For instance, Cruciani et al. [6] introduced a semi-population-based method to build a personalized model for smartphone-based HAR. They proposed an online training mechanism based on a semi-population neural network classifier, which could be updated only using a small amount of labeled data for user adaptation. Chen and Shen [7] analyzed motion sensor behaviors based on sensory data sequences collected from smartphones. They developed a cycle detection algorithm to extract cross-domain features from the obtained activity unit for activity recognitions. Zhang et al. [8] built a multi-layer architecture for smart home design and implementation based on the ontology and multi-agent technologies. They introduced an inference algorithm to learn the knowledge from unordered actions and temporal property of activities, in order to provide personalized services based on the inferred composite activities. Munoz-Organero [9] developed a deep Recurrent Neural Network (RNN) based algorithm to improve the outlier detection in HAR using wearable sensor data, which could extract data segments of a particular sub-activity from the main activity to optimize sub-activity recognitions. Huang et al. [10] focused on wearable HAR, and constructed a so-called two-stage end-to-end Convolutional Neural Network (CNN) model based on a single accelerometer. They then proposed a data augmentation method which could improve the accuracy in identifying the specific ascending stairs and descending stairs activities, and reduce the power consumption as well.

In particular, online and real-time recognition, has become an emerging issue, in terms of the dynamics in HAR tasks. Sun et al. [11] designed a multi-task learning model to deal with the personalized activity recognition task, in which a structured classification scheme was proposed to automatically discover task relationships, and further improve the online task learning, based on task-similarity approximations and probabilistic sampling. Li et al. [12] proposed a two-level clustering based identification method, and designed a dynamically updated activity plan, in which the probability of user adherence was analyzed according to their past activity patterns and current activity target.

Currently, researchers have tried to leverage a variety of machine learning techniques to realize HAR in a data-driven way. Hong et al. [13] focused on a semi-population based mechanism which only required a small number of labeled data from a new user to fit his/her features in a trained activity learning model. They further constructed a hybrid model using Bayesian networks and SVM, to improve the accuracy of HAR. Aiming to provide data-driven activity planning and typical activity pattern, Magherini et al. [14] implemented a system called automated recognizer of activities of daily living, in which the temporal logic and model checking could be utilized to generate the realistic patterns, and improve the automated real-time recognition in smart computing environments. Bianchi et al. [15] employed deep learning techniques in a HAR system, aiming to recognize single user’s daily activities in ambient assisted living environments. Based on the data collected from smart and wearable devices, they coupled sensors to a CNN on cloud, in order to facilitate the long-term personalized monitoring in smart home environments.
2.2 Studies on Federated Learning Technique

As an emerging learning paradigm, federated learning has shown its potential in globally training a large amount of user-generated data from diversified personal devices while preserving privacy issues in a decentralized way.

Current studies mainly focus on improving the global model or aggregation mechanism, in order to enhance the learning accuracy with a more efficient communication architecture. Wu et al. [16] presented a personalized federated learning framework, in which a cloud-edge architecture was employed to enable the personalized model deployment for intelligent IoT application development. Khan et al. [17] built a dispersed federated learning framework, to formulate an integer linear optimization, including the association and resource allocation problem. They developed a rounding technique based algorithm, which might enhance the learning robustness in an iterative way. Oh et al. [18] presented a communication-efficient framework based on federated learning for privacy-preserving, in which two algorithms, called federated learning after distillation and mixup data augmentation, were used to handle the asymmetric uplink-downlink channel capacities, and improve the accuracy and convergence time in MNIST classification tasks. Sattler et al. [19] constructed a framework for sparse ternary compression in the federated learning environment. They extended the top-k gradient sparsification, and improved the compression for both upstream and downstream communications with a newly designed communication protocol. Wang et al. [20] proposed a federated transfer learning that integrates federated learning and transfer learning to address the challenge of limited dataset for new applications and new devices in modern smart manufacturing applications. Chen et al. [21] discussed a federated learning scheme enhanced by an asynchronous learning strategy with a temporally weighted aggregation mechanism. The asynchronous learning strategy was designed to update parameters in the shallow and deep layers according to their different frequencies, while the temporally weighted aggregation mechanism was developed to improve the learning accuracy based on a better integration of local models.

In addition, different learning schemes have been integrated into the federated learning framework, to pursue an efficient distributed training process. Zhou et al. [22] introduced a hierarchical federated learning structure, in which a two-layer model aggregation scheme was designed, targeting modern vehicular network infrastructure in 6G environments. Zhan et al. [23] focused on the improvement of model training during the federated learning processes, based on an incentive mechanism. They proposed a deep reinforcement learning based mechanism to incent a pricing strategy for parameter optimization in edge. To minimize the communication cost and reduce the test error in the global model, Zhu and Jin [24] developed a multi-objective evolutionary algorithm with an optimized neural network model. They further introduced a scalable method to improve the encoding network connectivity in deep neural networks, aiming to enhance the federated learning efficiency. Chiu et al. [25] built an edge learning system, in which a single model considering both the semi-supervised learning and federated learning technologies was designed to periodically train and upload the data shared from edge devices. In addition, they proposed a so-called federated swapping scheme to alleviate the adverse impact of weight divergence during federated training process. Kwon et al. [26] focused on a specific case in Internet-of-Underwater-Things, thus developed a multi-agent deep reinforcement learning based algorithm to facilitate federated learning computation in ocean environments. They proposed an optimized resource allocation method to deal with the unexpected environment changes and channel unreliability based on deep deterministic policy gradient. Considering the combination of federated learning and reinforcement learning, Mowla et al. [27] introduced a security framework of flying ad-hoc network, in which the federated learning mechanism was used to realize the on-device attack detection, while the reinforcement learning was employed to optimize defense paths based on an adaptive epsilon-greedy policy.

2.3 Intelligent Applications on CPSS

CPSS have been widely talked due to the rapid development of IoT and mobile computing techniques, which inevitably incorporates human factors into intelligent systems and smart applications [5]. Zhang et al. [28] summarized the state of the art and perspectives of CPSS, and surveyed a variety of applications ranging from manufacturing, control systems, transportation, energy, etc.

Some studies considered graph models to analyze the network structure and dynamic connections in CPSS for smart design. Amin and Choi [29] built a graph model to detect the high communication traffic hotspots based on social network analysis. They measured the importance of hotspots using network metrics from the telecom data. Li et al. [30] introduced a mathematical structure to deal with the optimal scheduling issue for electric vehicle charging and discharging. They further discussed a case study in which a hierarchical optimal algorithm was integrated with a moving horizon method based on CPSS framework. Li et al. [31] constructed a hybrid graph model to handle the associated risks in CPSS, in which a hidden Markov model was employed to analyze the dynamic user activities, and a Bayesian risks graph model was used to evaluate the potential risks based on the classification of three different levels of activities. Xia et al. [32] developed a reputation-based mechanism for trust management in CPSS. They proposed a fuzzy logic model to analyze the trustworthiness, and utilized the directed acyclic graph with an expurgation mechanism to optimize service discovery based on the quantified indirect trustworthiness in a lightweight paradigm.

On the other hand, multiple factors were considered to improve the performance in CPSS. Wang et al. [33] et al. focused on the multi-objective optimization problem for high performance computing in CPSS. They developed an evolutionary algorithm and discussed a floorplanning case study based on a technical framework for CPSS. Su et al. [34] introduced an incentive scheme, in which the reputation of social users was considered as an important factor in improving the crowd sourcing service. They further
constructed an auction model, in order to facilitate the selection of social users with more needed data based on the analysis of interactions between CPSS and users. Aiming at formulating the general optimization problem in secure CPSS, Feng et al. [35] presented a tensor-based strategy with four different sub-models, in which multiple factors, such as execution time, energy consumption, and experience quality, were taken into account to optimize big data computing issues in CPSS. Zheng et al. [36] designed a privacy preservation framework in CPSS, and considered both the privacy concerns and user expectations in a data publishing mechanism with a heuristic algorithm using the formulated real-world knowledge.

3 2-DIMENSIONAL FEDERATED LEARNING FRAMEWORK

In this section, an overview of the proposed 2DFL framework is introduced, using personalized HAR as the motivating scenario. The vertical and horizontal federated learning schemes are explained to achieve personalized HAR models. The complete 2DFL algorithm is presented finally.

3.1 Framework Overview

Fig. 1 illustrates a typical scenario of a modern CPSS infrastructure, where multiple users may participate in a system while individual users have their own collection of wearable devices, home automation devices, and social media channels. In this study, it is assumed that each individual user will have a smartphone to host applications and communicate with other personal devices to collect heterogeneous data, such as location, motion, vital signs, and social media data. The vertical federated learning ensures that sensitive data collected by different devices of different vendors are kept safe and uncompromised, while allowing model enhancement using all the available data. On the other hand, the horizontal federated learning enables knowledge sharing by supporting secure model aggregation across multiple users. In this study, a personal smartphone is considered as a semi-honest aggregator and other personal devices are considered as honest-but-curious devices where no data sharing should happen in between those devices and their associated apps.

As shown in Fig. 1, the first part of the 2DFL aims to aggregate heterogeneous data collected from personal devices, such as smart watches, smart glasses, smart shoes, smart thermostat, and smart tracker. While CPSS envisage these devices to work together and offer intelligent services, in reality there are security concerns for these personal devices to exchange data with one another. In the proposed 2DFL framework, vertical federated learning is employed to address the limitation of heterogeneous data sharing across different devices to enhance a personalized HAR model while preserving privacy between devices of different vendors. At individual user level, personal devices do not exchange their data, and encrypted data are collected from personal devices and aggregated by a trusted third party (i.e. a smartphone). These data will be used to train a local HAR model with a more complete feature space.

The other part of the 2DFL framework involves horizontal federated learning to aggregate local models with a common feature space across multiple users to further enhance the model accuracy. Individual model parameters are also encrypted and shared with the cloud server where model aggregation happens across all participating users that have the same feature space. Iteratively, a high quality global HAR model will be built based on the contributions of multiple CPSS users. This part of the framework addresses the limitation of insufficient training data of each individual user and insecure data sharing across multiple
users.

The proposed framework demonstrates a natural complement of horizontal and vertical federated learning, in the application of HAR within CPSS. Horizontal federated learning offers a potential path to achieve a better global model based on the collective effort of all participating users. However, it relies on the quality of all local models which are limited by the data silos at individual user level. Therefore, vertical federated learning is adopted in this study to facilitate secure data sharing across multiple personal devices via a trusted third party (i.e., a smartphone) and to establish a local model with better quality. The proposed 2DFL framework will exploit the full potential of CPSS by making use of all the available heterogeneous data and achieve a more robust personalized HAR model.

### 3.2 Local Vertical Federated Learning

One of the key benefits of CPSS is the ability to collect heterogeneous data. However, devices developed by different vendors form data silos where direct data sharing is neither possible nor secure. This issue significantly limits the potential of CPSS in building a complete and accurate HAR model.

In the proposed 2DFL framework (illustrated in Fig. 1), vertical federated learning is adopted to aggregate heterogeneous data (i.e., from different devices) with different feature space, while preserving data privacy across multiple personal devices. In a typical vertical federated learning, the first step is to align common user groups from different parties, and the second step is to conduct encrypted model training across multiple parties [37]. In the designed vertical federated learning scheme, the first step is omitted. This is due to the fact that there is only one user in the targeted HAR application across multiple devices, and the purpose of the vertical federated learning is to aggregate different feature spaces offered by heterogeneous data. The encrypted model training process is illustrated in the vertical federated learning block in Fig. 1.

Referring to Fig. 1, the smartphone is considered as a trusted third party that is responsible for initiating the learning process by first distributing public keys to all participating devices (e.g., smart watch and smart glass in Fig. 1). Upon receiving the public keys, the participating devices encrypt and exchange their encrypted model parameters, Øsw and Øsg (i.e., parameters of the smart watch and smart glasses), with the trusted third party (i.e., the smartphone) to perform the training process. This is slightly different to the vertical federated learning proposed in [37] because embedded personal devices do not directly communicate with each other in the targeted HAR application. The vertical federated learning combines the different feature spaces of heterogeneous device data, and therefore the cost function is defined as follows.

$$L_{VFL} = \sum_i (u_i^{sw} + u_i^{sg} - y_i)^2 + \frac{\lambda}{2} (||\theta_{sg}||^2 + ||\theta_{sw}||^2).$$

(1)

where $u_i^{sw} = \Theta_{sw}x_i^{sw}$ and $u_i^{sg} = \Theta_{sg}x_i^{sg}$.

Based on the exchanged encrypted parameters, these parameters and the training model will be computed and updated according to the following equations.

$$\frac{\partial L_{VFL}}{\partial \theta_{sw}} = \sum_i (u_i^{sw} + (u_i^{sg} - y_i))x_i^{sw} + \lambda \theta_{sw}. \quad (2)$$

$$\frac{\partial L_{VFL}}{\partial \theta_{sg}} = \sum_i (u_i^{sg} + (u_i^{sw} - y_i))x_i^{sg} + \lambda \theta_{sg}. \quad (3)$$

### 3.3 Global Horizontal Federated Learning

The aforementioned vertical federated learning builds a local HAR model with a complete feature space (i.e., combining feature spaces from heterogeneous device data). However, the local HAR model may not perform very well due to the limited training data. The second phase of the proposed 2DFL framework is to perform horizontal federated learning to aggregate local HAR models that share the same feature space from multiple users to enhance the quality of the trained model. The aggregated global model is built based on the collective knowledge of all the participating users, which addresses the issue of limited training data in modern HAR application. The global model will be redistributed to each individual user, where the model will be enhanced and personalized again by local heterogeneous data. Iteratively, high quality personalized HAR models will be achieved.

As illustrated in Fig. 1, considering a typical CPSS setting, there can be $N$ participating users with their personal smartphones collecting common CPSS data. These users (or their smartphones) form a set of data owners $\{d_i\}_1^N$, which collects personalized data $D = \{d_i\}_1^N$ where each data set contains a collection of CPSS data $x_i$ (e.g., location, motion, and social media), and the corresponding label $y_i$ as $d_i = (x_i, y_i)$ of a particular user. A local HAR model, $h_{ij}$, is trained by each individual user using the corresponding local data set, $d_i$. A global HAR model, $H$, is trained by aggregating a set of model parameters extracted from each individual local HAR model. To ensure privacy preservation, local parameters are encrypted before transmitting to the cloud server, and no sensitive data about individual data owners will be exchanged.

Specifically, in the proposed 2DFL model, a horizontal federated learning with CNN is implemented. Since the focus of this study is on the proposed 2DFL, CNN is adopted for its direct applicability in federated learning. However, the proposed model has the potential to work with other machine learning algorithms, which will be further studied in the future. Initially, a local CNN-based HAR model will be trained by the individual data owner to identify a set of common human activities represented by the hypothesis $a(x_j, \omega)$.

The prediction of the hypothesis is implemented by a SoftMax classifier defined in Eq. (4) as the activation function, targeting the HAR application which is a multi-class classification problem. The loss function of the classifier is defined in Eq. (5), and the gradient descent optimizer is applied to train the model with the loss function.

$$y_{j\text{predict}} = \text{Softmax}(a(x_j, \omega)) = \frac{e^{y_j}}{\sum_i e^{y_i}}. \quad (4)$$

$$L_{HFL} = -\frac{1}{k}[\sum_{q=1}^k \sum_{p=1}^l \{y_f = p\} \log(\frac{e^{a_{y_f}}}{\sum_{q=1}^k e^{a_{y_f}}})]. \quad (5)$$
where each data owner, \( p_j \), has \( k \) samples.

Iteratively, the local model parameters acquired by each individual user will be encrypted and transmitted to the cloud server for aggregating and forming a global HAR model. The global model is formed by combining the received parameters as follows.

\[
H = \frac{1}{N} \sum_{j=1}^{N} \omega_j.
\]

The horizontal federated learning exploits the potential of collective effort, where multiple users’ data can be used together in training a more accurate model. At the end of the horizontal federated learning, a global HAR model based on a collection of multiple user data will be established, which can then be distributed to individual users to perform further personalized training using their own heterogeneous device data.

### 3.4 Algorithm for 2-Dimensional Federated Learning

As the name suggested, the 2DFL is a combination of both the horizontal and vertical federated learning, conducted sequentially. Referring to the algorithm pseudo code shown in Algorithm 1, lines 2-12 correspond to the horizontal federated learning across multiple participating users using the data with common feature space. At the end of the horizontal federated learning, a global model will be produced, which will be used as the baseline model for further enhancement at individual user level.

The other phase of the 2DFL (lines 13-19) makes use of the designed vertical federated learning to utilize the heterogeneous data collected by multiple personal devices with different feature spaces. Model parameters from multiple personal devices are aggregated by a trusted third party (i.e., a smartphone) to iteratively compute new model parameters to minimize the loss function.

In conjunction, the horizontal federated learning addresses the issue of insufficient training data by allowing privacy-preserved parameter sharing to establish an accurate global model. The vertical federated learning enables utilization of heterogeneous data across multiple personal devices to achieve a more complete feature space, while ensuring no data leakage across different devices. The final outcome of the 2DFL is to achieve a more completed and accurate HAR model.

To protect the data privacy, the Brakerski-Gentry-Vaikuntanathan (BGV) scheme is introduced as a SWHE algorithm. The HElib is used to implement the BGV scheme. Given a set of original data, four key components, including Key Generation (KeyGen), Encryption (Enc), Evaluation (Eval), and Decryption (Dec) are applied with HElib (as shown in lines 5, 6, 9, 11). Noted that, the BGV scheme applied in this study is not a full homomorphic encryption scheme, which can only support limited operations in the 2DFL training process. The advantage of the BGV scheme is its low computational complexity in comparison with a full homomorphic encryption scheme. This strategy encrypts the raw parameters to cipherdata first with the public key generated from client side. The 2DFL consumes the cipherdata and performs limited operations (addition/subtraction/multiplication) specified by the BGV scheme. The process result will be sent back to clients and then decrypted with a private key, which delegates processing of the data without giving access to it.

#### Algorithm 1. 2DFL for HAR

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialize model parameter ( \omega ), maximum iterations ( T_1 )</td>
</tr>
<tr>
<td>2</td>
<td>for ( t = 1 ) to ( T_1 ) do</td>
</tr>
<tr>
<td>3</td>
<td>for each data owner ( p_i ) \in ( P ) do</td>
</tr>
<tr>
<td>4</td>
<td>Conduct local training for the model ( M ) on dataset ( d_i )</td>
</tr>
<tr>
<td>5</td>
<td>Generate public and private key ( (pk_i, sk_i) ) with secure parameter ( \lambda_i ); ( (pk_i, sk_i) ) \leftarrow KeyGen(( \lambda_i ))</td>
</tr>
<tr>
<td>6</td>
<td>Encrypt the local parameter ( \omega_i(t) ) to cipherdata: ( c_i(t) ) \leftarrow Enc(( \omega_i(t), pk_i ))</td>
</tr>
<tr>
<td>7</td>
<td>Upload cipherdata ( c_i(t) ) to the cloud server</td>
</tr>
<tr>
<td>8</td>
<td>end for</td>
</tr>
<tr>
<td>9</td>
<td>Aggregate local parameters using Eq. (4) with evaluation by ( \text{Eval}(c_i, c_j) ) for each pair ( c_i, c_j )</td>
</tr>
<tr>
<td>10</td>
<td>Distribute the global model parameters to each data owner</td>
</tr>
<tr>
<td>11</td>
<td>Decrypt the encrypted global model parameters ( c(t) ) for data owner ( p_i; \omega_i(t) ) \leftarrow Dec(( c(t), sk_i ))</td>
</tr>
<tr>
<td>12</td>
<td>end for</td>
</tr>
<tr>
<td>13</td>
<td>Initialize local models based on global model parameters ( \omega ), maximum iterations ( T_2 )</td>
</tr>
<tr>
<td>14</td>
<td>for ( t = 1 ) to ( T_2 ) do</td>
</tr>
<tr>
<td>15</td>
<td>Exchange intermediate model parameters</td>
</tr>
<tr>
<td>16</td>
<td>Compute new model parameters and loss function using Eq. (1-3)</td>
</tr>
<tr>
<td>17</td>
<td>Update model ( h_i ) by BGV scheme</td>
</tr>
<tr>
<td>18</td>
<td>end for</td>
</tr>
<tr>
<td>19</td>
<td>return ( har_i )</td>
</tr>
</tbody>
</table>

### 4 Experiment and Analysis

To verify the effectiveness, the proposed 2DFL model is applied to execute collaborative learning tasks for data sharing and HAR in a simulated cloud-edging network. Evaluations of the proposed algorithm were conducted on CPSS data that consists of multiple sensor inputs collected using smartphones and on-body wearable devices in the real world. The execution environment of the experiment is configured with Intel i7-4790 @3.6GHz CPU, 32GB RAM, NVIDIA GeForce GTX 970 GPU, CentOS, Python 3.6, and PyTorch.

#### 4.1 Experiment Setup

To investigate the proposed 2DFL model on HAR applications in a typical CPSS environments, an accelerometer dataset collected using smartphones and on-body wearable devices (head, chest, upper arm, waist, thigh, and shin) [38], was applied here for the performance evaluation. This dataset contains 11,771 activities performed by 30 subjects of ages ranging from 18 to 60. Specifically, this dataset covers 9 types of activities in daily living and 8 types of falls. These acceleration data are utilized as high-level common sensor data to conduct training for the horizontal federated learning in CPSS. In this study, six types of activities (totally
10,510 samples) as shown in Table 1 were adopted to evaluate the performance of the proposed method. It can be observed in the table that there is an unbalanced distribution across the six types of activities. In addition, this dataset is partitioned into 30 groups to simulate the cloud-edge federated learning, targeting HAR applications in CPSS scenario.

Furthermore, we separate the acceleration dataset in Table 1 to build the personal gadget data, from fifteen subjects (seven females and eight males, age 31.9 ± 12.4, height 173.1 ± 6.9 cm, weight 74.1 ± 13.8 kg). This separated dataset is used as the personal gadget data for the vertical federated learning process. To simulate the different scale of distributed application scenario with various groups, the personal dataset is separated into a series of groups ranging from 5 to 30 in this study.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climbing down</td>
<td>2,632</td>
</tr>
<tr>
<td>Jumping</td>
<td>374</td>
</tr>
<tr>
<td>Standing</td>
<td>1,264</td>
</tr>
<tr>
<td>Lying</td>
<td>218</td>
</tr>
<tr>
<td>Walking</td>
<td>2,152</td>
</tr>
<tr>
<td>Jogging</td>
<td>3,870</td>
</tr>
</tbody>
</table>

In particular, each group represents a participant in the distributed CPSS scenario that collectively forms the dataset. Through this arrangement, each participant may hold a different category of data and therefore achieve a more random group composition. This will further increase the difficulty of HAR during the training process.

To evaluate the performance of the proposed method, three baseline approaches, DNN, Siamese Convolutional Autoencoder (SCAE) [39], and FSL-CNN [40] are applied in this experiment. The deep learning classifier DNN is implemented by a multi-layer perceptron structure with an input layer (12 nodes), three hidden layers (12 nodes, 24 nodes, 12 nodes), and an output layer with SoftMax activation function. The SCAE model, with a twin convolutional autoencoders for feature comprising, is involved for comparison evaluations. The FSL-CNN classifier is implemented by a structure with a convolution layer (filter size 7×7), a ReLU layer, a maxpool layer, a convolution layer (filter size 5×5), a maxpool layer, and two fully connected layers with SoftMax activation function. These three methods are chosen to compare with the 2DFL model targeting HAR applications with limited training data in CPSS environments.

Based on the experimental setup, we evaluate the performance of the proposed algorithm in both learning efficiency and accuracy. For learning efficiency evaluation, we examine the process of convergence according to all the six activities listed in Table 1 to evaluate how the model fits different situations (i.e., under six different activities). The convergence and its learning process are investigated by loss rate of the algorithms according to different types of activities in the evaluation. For accuracy evaluation, the widely used precision, recall, and F1 metrics are adopted to investigate the achieved accuracy of HAR in CPSS environment, comparing with the three baseline methods. It is important to note that the cross-validation approach is applied to ensure a generalizable model is achieved for the proposed method. Besides, the Receiver Operating Characteristic (ROC) curve is also used to illustrate and evaluate how the proposed 2DFL works on different distributed application scale. Specifically, the metrics of accuracy is quantified by the area under the ROC curve, which is expressed as Area Under the Curve (AUC).

4.2 Performance Evaluation and Discussion

First, we try to evaluate the performance of the proposed 2DFL model in a typical distributed CPSS environment, against the other three baseline methods. The loss rate curves for the 2DFL and the three baseline classifiers are illustrated in Fig. 2(a)-(f) respectively for the activities listed in Table 1. In this experiment, we choose Stochastic Gradient Descent (SGD) as the optimizer with learning rate of 0.5 to train the models.

As shown in Fig. 2, the first 500 iterations for the loss curve are plotted to describe the training process for all the methods. All the methods demonstrate reasonable loss curves along with the training iterations. All four methods present different fluctuations in the loss curves and the gradient descent speed for the learning process in terms of different types of activities as shown in Fig. 2(a)-(f) respectively. However, it is obvious that the proposed 2DFL model achieve a faster convergence with a relatively lower loss rate in comparison with the other three methods when conducting the training for all activities. It is also worth noticing, the 2DFL achieves the lowest loss on all the activities, which is better than the other three methods. In addition, another merit of the 2DFL model observed from the learning process is that the loss curve of 2DFL is smoother (i.e. less fluctuation) than those baselines especially when performing the training for activities Jumping and Jogging, which indicates a more stable training process for the proposed 2DFL model. From the above results, it can be observed that integrating the heterogeneous personal data and features across multiple personal devices by the vertical and horizontal federated learning in the proposed model may improve the learning process effectively.

Second, to further evaluate the HAR accuracy for all the methods, the Precision, Recall and F1 Measures are used as the evaluation metrics in this experiment. The results are summarized in Table 2. A group of experiments are conducted according to different numbers of extracted features from the used sensors and personal data, which is different in each approach. There are totally 42 features extracted from different sensor devices, which are listed in Table 3. However, without proper protection embedded in the learning framework, other methods can not support embedded devices to safely exchange their data, and hence the derived HAR model can not aggregate all features based on data from all devices.
For example, the best results achieved by the SCAE and DNN are achieved by using the 24 features extracted from accelerometer and GPS that belong to one personal device, whereas the FSL-CNN achieves its best result by using only the 15 acceleration data features. In comparison, our 2DFL that is capable of accessing data from all wearable devices in a safe 2D federated learning scheme achieves the best performance.

Since the 2DFL model integrates data from multiple sensor devices and personal gadgets, it is able to access features extracted from all devices. Depending on the used devices, the best results achieved by the other three methods are listed in Table 2. The results clearly demonstrate

**Table 2**

<table>
<thead>
<tr>
<th>Methods</th>
<th>Number of Features</th>
<th>Accuracy</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCAE</td>
<td>24</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>DNN</td>
<td>24</td>
<td>0.77</td>
<td>0.67</td>
<td>0.71</td>
</tr>
<tr>
<td>FSL-CNN</td>
<td>15</td>
<td>0.87</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>2DFL</td>
<td>42</td>
<td>0.93</td>
<td>0.89</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Table 2 shows the best results that can be achieved by individual methods when they can only access data (and the associated features) from a specific wearable device.
that the proposed 2DFL model achieves a substantial improvement than the other three methods according to all three metrics. Depending on the locations of individual sensors and data offered by different personal devices, different features can be extracted to strengthen a specific activity recognition. Therefore, it is critical to make use of the entire personal dataset to achieve a more complete HAR model. This is obviously indicated in the results where the 2DFL shows a clear strength by the ability to access and use data (and hence features) from all sensor and personal devices, and clearly outperforms the other methods in terms of the speed of convergence and classification accuracy.

Table 3

<table>
<thead>
<tr>
<th>Features</th>
<th>42 Features from All Sensor Devices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (x, y, z)</td>
<td>Standard deviation (x, y, z)</td>
</tr>
<tr>
<td>Variance (x, y, z)</td>
<td>Median (x, y, z)</td>
</tr>
<tr>
<td>IQR (x, y, z)</td>
<td>MAD (x, y, z)</td>
</tr>
<tr>
<td>Kurtosis (x, y, z)</td>
<td>Energy (x, y, z)</td>
</tr>
<tr>
<td>MeanDC (x, y, z)</td>
<td>Orientation (x, y, z)</td>
</tr>
<tr>
<td>EntropyT (x, y, z)</td>
<td>EntropyF (x, y, z)</td>
</tr>
<tr>
<td>Color Context (x, y, z)</td>
<td>Gravity (x, y, z)</td>
</tr>
</tbody>
</table>

Finally, we examine and compare the performance variation of the proposed 2DFL and a typical horizontal federated learning with respect to different distribution scales. We evaluate the performance variation with different numbers of edge nodes, ranging from 1 to 30. The performance variation is depicted in Fig. 3. As shown in the figure, both two learning schemes show a general performance improvement with increasing number of edge nodes. Both the 2DFL and horizontal federated learning have a significant rise when increasing from 1 to 15 nodes. The ROC values for both methods are stable around 0.8 along with the increasing edge nodes from 15 to 30. This result shows that more edge nodes can improve the results in general as more collected data will be included in the learning process. However, by having access to all device data, it can also be observed that the 2DFL significantly outperforms the ordinary horizontal federated learning especially in the cases with small amount of edge nodes. It should also be noted that more edge nodes may also introduce additional uncertainties and noises, leading to a slightly worse performance in both methods as observed from 15 to 20 in the figure.

5 Conclusions

The advancement of CPSS offers many benefits to real-world applications through abundant heterogeneous data and the corresponding data-driven intelligent services. However, several challenges remain in CPSS which limit the potential of CPSS in real-world applications. One of the critical challenges is to be able to share and use heterogeneous data across different users and devices without compromising data privacy. In this paper, a 2DFL model, which incorporates the advantages of both vertical and horizontal federated learning, was designed to allow secure distributed learning across different smart devices among multiple users, and further to make use of heterogeneous data to enhance personalized HAR in CPSS. The vertical federated learning scheme was developed to integrate sharable features into a more complete feature space of an individual user based on heterogeneous data generated across different devices. The horizontal federated learning scheme was developed to aggregate the encrypted local models among multiple individual users and achieve a high-quality global model for HAR. A computationally efficient SWHE scheme was improved and applied to support the parameter aggregation without giving access to it, which enabled heterogeneous data sharing with privacy protection across different personal devices and multiple users in building a more precise personalized HAR model. Experiments were conducted by comparing with three baseline methods. The results showed that the proposed 2DFL model outperformed the other three methods in both learning efficiency and accuracy, achieving faster converging loss rate and better precision, recall, and F1 scores.

In our future studies, we will go further to improve the efficiency of the algorithm for its use in real-world applications. More evaluations on different practical scenarios will be conducted to demonstrate the applicability of the proposed model and method.

Acknowledgment

The work was supported in part by the National Key R&D Program of China No. 2019YFE0190500, in part by the National Natural Science Foundation of China No. 62072171, No. 72091515.

References


Xiaokang Zhou (M’12) is currently an associate professor with the Faculty of Data Science, Shiga University, Japan. He received the Ph.D. degree in human sciences from Waseda University, Japan, in 2014. From 2012 to 2015, he was a research associate with the Faculty of Human Sciences, Waseda University, Japan. He also works as a visiting researcher in the RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, Japan, since 2017. Dr. Zhou has been engaged in interdisciplinary research works in the fields of computer science and engineering, information systems, and social and human informatics. His recent research interests include ubiquitous computing, big data, machine learning, behavior and cognitive informatics, cyber-physical-social systems, cyber intelligence and security. Dr. Zhou is a member of the IEEE CS, and ACM, USA, IPSJ, and JSAI, Japan, and CCF, China.

Wei Liang (M’19) received his M.S. and Ph.D. degrees in Computer Science from Central South University in 2005 and 2016. From 2005 to 2012, he worked in Microsoft (China) for soft engineering. From 2014 to 2015, he worked as an exchange researcher in the Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He is currently working at Key Laboratory of Human Province for New Retail Virtual Reality Technology, Hunan University of Technology and Business, China. His research interests include information retrieval, data mining, and artificial intelligence. He has published more than 20 papers at various conferences and journals, including FGCS, JCSS, and PUC. Dr. Liang is a member of the IEEE CS.

Jianhua Ma (M’98) is a professor with the Department of Digital Media, Faculty of Computer & Information Sciences, Hosei University, Japan. His research interests include multimedia, networking, pervasive computing, social computing, wearable technology, IoT, smart things, and cyber intelligence. Dr. Ma is one of pioneers in research on Hyper World and Cyber World (CW) since 1996. He first proposed Ubiquitous Intelligence (UI) towards Smart World (SW), which he envisioned in 2004, and was featured in the European ID People Magazine in 2005. He has conducted several unique CW-related projects including the Cyber Individual (Cyber-I), which was featured by and highlighted on the front page of IEEE Computing Now in 2011. He has published more than 300 papers, co-authored 5 books and edited over 30 journal special issues. He has founded three IEEE Congresses on ‘Smart World’, ‘Cybermatics’ and ‘Cyber Science and Technology’, respectively, as well as IEEE Conferences on Ubiquitous Intelligence and Computing (UIC), Pervasive Intelligence and Computing (PICom), Advanced and Trusted Computing (ATC), Dependable, Autonomic and Secure Computing (DASC), Cyber-Physical and Social Computing (CPSCom), Internet of Things (iThings), and Internet of People (IoP). Dr. Ma is a member of IEEE and ACM, the chair of IEEE SMC Technical Committee on Cybermatics, the founding chair of IEEE CIS Technical Committee on Smart World, and in the advisory board of IEEE CS Technical Committee on Scalable Computing.

Zheng Yan (M’96, SM’14) received the BEng degree in electrical engineering and the MEng degree in computer science and engineering from the Xi’an Jiaotong University, Xi’an, China in 1994 and 1997, respectively, the second MEng degree in information security from the National University of Singapore, Singapore in 2000, and the licentiate of science and the doctor of science in technology in electrical engineering from Helsinki University of Technology, Helsinki, Finland. She is currently a professor at the Xi’an University, Xi’an, China and a visiting professor at the Aalto University, Espoo, Finland. Her research interests are in trust, security, privacy, and security-related data analytics. Prof. Yan serves as a general or program chair for 30+ international conferences and workshops. She is a steering committee co-chair of IEEE Blockchain international conference. She is also an associate editor of many reputable journals, e.g., IEEE Internet of Things Journal, Information Sciences, Information Fusion, JNCA, IEEE Access, SCN, etc.

Kevin I-Kai Wang (M’04) received the Bachelor of Engineering (Hons.) degree in Computer Systems Engineering and PhD degree in Electrical and Electronics Engineering from the Department of Electrical and Computer Engineering, the University of Auckland, New Zealand, in 2004 and 2009 respectively. He is currently a Senior Lecturer in the Department of Electrical and Computer Engineering, the University of Auckland. He was also a research engineer designing commercial home automation systems and traffic sensing systems from 2009 to 2011. His current research interests include wireless sensor network based ambient intelligence, pervasive healthcare systems, human activity recognition, behavior data analytics and bio-cybernetic systems.