Research Advanced in Personalized Federated Learning

Zizhuo Wang[a]
School of Internet of Things Engineering, Jiangnan University, Wuxi, China

Keywords: Personalized Federated Learning, Data Heterogeneity, Machine Learning.

Abstract: Federated learning eliminates the need for local device data sharing by enabling diverse users to team up to build shared global models, and has gradually become a research hotspot in machine learning communities in recent years. Despite being widely used in many application fields, the convergence of disparate data types and the absence of tailored solutions continue to pose challenges for federated learning models. In this context, pFL(personalized federated learning) has rapidly developed, aimed at reducing heterogeneity and creating personalized models for each device through personalized processing at the device, data, and model levels. The latest research progress in personalized federated learning is systematically reviewed in this article. Focusing on the aspects of global model personalization and learning personalized models, this article first introduces representative personalized federated learning algorithms, including their design ideas and key steps. This article also discusses the challenges in the field of personalized federated learning and looks forward to its future development direction, which aims to bring more new insight to this field.

1 INTRODUCTION

With the ongoing accumulation of multimedia data and the swift advancement of artificial intelligence technologies, how to acquire valuable information from large-scale data has emerged as a prominent research focus within the computing community. Based on data-driven a large amount of effort has been invested, greatly promoting the development of numerous scene understanding tasks (Xu, 2021; Sun, 2023) and practical applications (Xu, 2020; Wang, 2023). However, existing model construction often requires training scattered data sets before completing the training. The above centralized training paradigm not only damages data privacy and security, but also increases additional communication and storage costs. To address the aforementioned issues, federated learning has gradually attracted numerous research interests.

Federated learning facilitates collaboration among multiple users to develop a global model that is shared without necessitating the exchange of data derived from local devices. The central server receives the updated model after every client trains it according to their local data. After gathering all the updates sent back by the clients, the central server revises the global model on a single occasion. Utilizing the aforementioned multi-round learning and communication techniques, FL eliminates the necessity of collecting all data onto a solitary device, allows machine learning models to analyze the data that is kept by different users (or clients), while simultaneously addressing privacy and communication hurdles encountered in the process of machine learning tasks. However, federated learning's success is determined by the assumption that the information across all data centers is autonomous and uniformly dispersed, which is not realistic in practical complex applications. To deal with the above problems, personalized federated learning was born as a technique that combines federated and personalized learning approaches to train a model that takes into account each user's unique data characteristics and needs while protecting user privacy. This method has the ability to handle problems associated with non-IID(not independent and identically distributed) data in an effective manner, enhance the efficacy of the model in performing particular tasks, and diminish the requirement for data transfer, and save bandwidth and computing resources. Personalized Federated learning could provide a more accurate and personalized user experience, adjust to various scenarios, and meet the needs of different industries

[a] https://orcid.org/0009-0006-1754-8031

Wang, Z.
Research Advanced in Personalized Federated Learning.
DOI: 10.5220/0012958400004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 536-540
ISBN: 978-989-758-713-9

and users. Current studies can divide personalized federated learning into two stages: global model personalization and the learning personalized models.

(1) Global model personalization allows personalizing the data distribution for each client or user, which also maintain a global shared model at the same time. The advantage is that certain commonalities can be preserved, and the unique characteristics of each user are taken into consideration, so as to strengthen the applicability and accuracy of the model. Two stages can be assigned to the method: global model training and individual adaptation. First, all users train a shared federated learning model, serving as the basic model to capture the common features of all users' data. Secondly, after the model training is completed, each user who participates in the training carries on further training on the local client, to make the model better meet the personalized needs of users.

(2) The solution of Learning personalized models emphasizes the differences between individual users. The method gives all users a separate model to learn, which is suitable for the case of large differences in data distribution among users, and can finely capture the characteristics of each user's data. The steps used in this model are relatively complex, including: goal setting, data acquisition, model initialization, local training and personalized adaptation, local optimization, model evaluation, model aggregation and update, continuous iteration, and deployment.

Focusing on the aforementioned two aspects, this article aims to report the research advanced in the personalized federated learning topic and is structured as following steps. In Section 2, the key and the main challenges of federated learning and the background of personalized federated learning are discussed. Furthermore, Section 2 also analyzes the progress of the research of global model personalization and learning personalized models, including their representative algorithms. Section 3 introduces some commonly used datasets for personalized federated learning and reports the performance of some representative methods. Section 4 discusses the future development of personalized federated learning and Section 5 finally concludes the whole work.

2 METHOD

2.1 Revisiting Federated Learning

The fundamental concept of federated learning involves the process of downloading models to each data center for training and upload them to the cloud for aggregation. The success of it is highly dependent under the presumption that the data across all data centers is independent and identical. However, the actual data distribution is not ideal, and federated learning confronts two primary challenges: inadequate convergence due to substantial data heterogeneity and a scarcity of tailored solutions. For the former, the accuracy of federated learning methods will significantly decrease when identically distributed local data distributions that are non-IIDare used for local training and synchronization. For the latter, given that traditional federated learning involves the training of an individual global shared model and its subsequent adaptation for use across various clients, the global model will likely struggle to generalize effectively in situations where the local distributions differ markedly from each other and from the overall global distribution. To counteract the sluggish convergence issue of federated learning approaches when dealing with non-IID data and the absence of customized models tailored for specific local tasks or data sets, personalized federated learning proposes two targeted strategies: global model personalization and learning personalized models.

The objective of personalizing global models is to resolve the performance challenges encountered during the training of globally shared federated learning models on heterogeneous data, whose main goal is to maintain a shared global model while adapting and optimizing it to suit the specific needs of individual clients well. The setting of personalized federated learning closely follows the general federated learning training process, which first trains a individual global federated learning model. Then, by performing additional local adaptation steps on each local dataset, an individualized global federated learning model is fine-tuned for each individual federated learning client. This type of personalized technology can be disconnected from data-based methods and model-based methods. While the model-based methods are intended to study a robust global model for the specific processing of personalized clients or to enhance the adaptive performance of local models in the future, the data-driven methods address client drift by reducing statistical fluctuations among client datasets.

Learning personalized models aims to address the challenges of personalized solutions. Contrary to the personalized strategy of training One worldwide model, methods belonging to this category train the single personalized federated learning model and then modify the model aggregation process to construct the global personalized models. Existing methods of learning personalized models are mainly based on architecture and similarity. The architecture-based approach aims to provide a custom-fitted model framework designed individually for each customer, while the similarity-based approach aims to enhance the efficiency of personalized models by utilizing customer relationships, where similar individualized models are constructed for the respective customers.

2.2 Global Model Personalization

2.2.1 Data-Based Methods

Data-based strategies strive to minimize the variability in the statistical characteristics of client data distributions, mainly including data augmentation and client selection. Client selection focuses on devising mechanisms to select clients that allow for sampling from a data distribution that is more uniform. For the data augmentation methods, their basic idea is to generate identically distributed data to reduce data imbalance. Common methods include oversampling techniques (such as SMOTE and ADASYN) and undersampling techniques (such as the Tomek link). However, due to the private nature of the client's data, it is not practical to apply directly. the aforementioned data augmentation methods in the context of federated learning. A key hurdle in data augmentation within the realm of federated learning is the necessity for data sharing or the presence of surrogate datasets that adequately mirror the overall data distribution.

2.2.2 Model-Based Methods

The model-based methods are used to learn a powerful global federated learning model so that each client can be personalized to enhance the adaptive capabilities of local models. Meta-learning is a solution that consists of training multiple learning tasks and producing highly adaptive models (Nichol, 2018). For example, Finn et al. (Finn, 2017) proposed a model-independent meta-learning (MAML) algorithm that constructs global models on multiple tasks and adjusts global models on individual tasks. In addition, federated education has employed transfer learning. Schneider and his team succeeded in personalizing models in non-federated environments (Schneider, 2019), while Wang et al. proposed to re-learn the parameters of pre-trained global models on local data (Wang, 2019).

2.3 Learning Personalized Model

2.3.1 Architecture -based Methods

The architecture-based methods achieve personalization by designing individual model structures for each data center, with representative approaches being parameter decoupling and knowledge distillation. Each client's personalized layer is implemented utilizing the parameter decoupling approach, which is bolstered by a personalized model architecture framework. FedMD (Li, 2019), integrating knowledge distillation and transfer learning methodologies, allows customers to independently construct their neural networks by merging their own private data with the universal public dataset. Yu et al. (Yu, 2020) set a model that is both global and personalized as teacher and student networks respectively, directing pupils to transfer insights from the instructor's framework into their own, whilst emulating the results of the teacher's model.

2.3.2 Similarity-Based Methods

By modelling customer relationships, the similarity-based approach aims to achieve personalization. Learn a personalized model for each customer, and relevant customers learn similar models. This is often conceptualized as a multi-task learning process. In multi-task learning, each FL client is treated as a task, the relationships between clients can be learned and captured. MOCHA (Smith, 2017) learns personalized models for each federated learning client. Since all clients need to participate In every iteration of federated learning model training, it's improper for cross-device applications.

3 EXPERIMENT

3.1 Common Dataset

Electricity consumption dataset (Jiao, 2024) are the real electricity consumption data of 10 districts and counties in a city in northern China from November 25, 2016, to November 25, 2017. The time granularity of data collection is 1 hour, with a total of 24 load values throughout the day. In addition to load data, the dataset includes calendar factors, including year, month, day, hour, and meteorological factors such as maximum and minimum temperatures.

Fuel consumption dataset (Han, 2024) from 20 different ships, including container ships, bulk carriers, tankers, and multi-purpose vessels, were collected for training and analysis. This data, along with sea state information from the ECMWF and CMEMS, covered a year with 6-hour sampling intervals, totaling roughly 24,000 samples. Key variables included ship speed, draft, tilt, RPM, and 18 sea state metrics like wind, wave, and current conditions. After the data is fused, cleaned, and filtered, a process of personalized federated learning takes place.

3.2 Performance Analysis

To examine how various representative personalized federated learning algorithms perform, this section compared the performance of five models: FedAvg, FedPer, pFedMe, local learning (LL), and central learning (CL) on Electricity consumption dataset. LL is a method where every client independently uses its local data for model training and testing. In contrast, CL involves uploading all clients' data to a primary server is responsible for conducting consolidated model training and subsequently redistributing the refined model to each individual client for evaluation. The FedAvg algorithm is a federated learning approach that trains a global model by aggregating data from various clients, which is then used for testing on each client's data. The FedPer method selects 4 out of 6 layers in a neural network as the globally shared base layers, which are trained using FedAvg, while the remaining 2 layers serve as personalized layers, trained independently by each client. In pFedMe, the clients utilize the Moreau envelope as their regularization loss function, an individualized federated learning method, optimizing both the customized model and the global model simultaneously. In short, pFedD achieved the best prediction effect. Compared with LL, CL, FedAvg, FedPer and pFedMe, the average MAPE of pFedD was reduced by 1.82%, 1.26%, 1.67%, 0.53% and 0.38%, respectively. The average RMSE was reduced by 8.64, 6.66, 8.46, 3.19 and 2.47, respectively.

4 DISCUSSION

As an important branch of federated learning, although significant progress has been made in the research of personalized federated learning, there are still many challenges for its future development:

(1) Technology maturity and commercial landing. Federal learning has emerged in China since 2018, and after several years of development, technology and engineering have gradually matured. With the emergence of numerous platforms and products, federated learning has begun to move towards large-scale commercial implementation. This suggests that personalized federated learning is also benefiting from this trend and is gradually moving from the research phase to practical applications.

(2) Solving the problem of data heterogeneity. If data is non-IID, traditional methods of federated learning may experience performance degradation, such as highly heterogeneous data. Personalized federated learning aims to improve the convergence and accuracy of models on this type of data by combining strategies for global models and personalized models.

(3) Algorithm innovation. To tackle the difficulties presented by non-IID data, research on personalized federated learning is exploring new algorithms. For example, one approach uses two-phase training, training a shared global FL model, followed by additional training on local data for personalization.

(4) Balancing safety and efficiency. Trends in federal learning suggest that future research will seek to balance safety and efficiency. This means that personalized federated learning also needs to be optimized in order to satisfy the requirements of practical applications in both respects.

5 CONCLUSION

This paper concentrates on the topic of personalized federated learning and introduces its latest research progress. More specifically, from the perspectives of personalizing a global model and learning an individualized model, this paper provides a detailed introduction to the data-based, model-based, architecture-based, and similarity-based approaches, including their design ideas and representative algorithms. This paper also discusses the main challenges of personalized federated learning. While personalized federated learning confronts challenges, including data heterogeneity, its ability to provide personalized services while protecting privacy makes it promising in multiple fields. As technology continues to advance, personalized federated learning shows a promising future in driving the industry forward.

REFERENCES

Finn, C., Abbeel, P., Levine, S., 2017. Model-agnostic meta-learning for fast adaptation of deep networks. *in*

Proceedings of the 34th International Conference on Machine Learning, Volume 70. JMLR. org, 2017, pp. 1126‑1135.

Han, P., Sun, Z., Liu, Z., Yan, C., 2024. Heterogeneous ship fuel oil consumption prediction at sea based on personalized federated learning. *Computer Integrated Manufacturing Systems*.

Jiao, R., Chu, J., Li, J., Zhang, W., 2024. Personalized Federated Multi-region Load Forecasting Method Based on Data Decomposition. *Proceedings of the CSEE.*

Li, D., Wang, J., 2019. Fedmd: Heterogenous federated learning via model distillation. *arXiv preprint arXiv:1910.03581.*

Nichol, A., Achiam, J., Schulman, J., 2018. On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999*.

Sun, H., Lang, W., Xu, C., Liu, N., Zhou, H., 2023. Graph-based discriminative features learning for fine-grained image retrieval. *Signal Processing: Image Communication*, 110:116885.

Schneider, J., Vlachos, M., 2019. Mass personalization of deep learning. *arXiv preprint arXiv:1909.02803*.

Smith, V., Chiang, C.-K., Sanjabi, M., Talwalkar, A., S., 2017. Federated multi-task learning. *in Advances in Neural Information Processing Systems*, pp. 4424-4434.

Wang, J., Wang, Y., Liu, L., Yin, H., Ye, N., Xu, C. , 2023. Weakly Supervised Forest Fire Segmentation in UAV Imagery Based on Foreground-Aware Pooling and Context-Aware Loss. *Remote Sensing* 15, no. 14: 3606.

Wang, K., Mathews, R., Kiddon, C., Eichner, H., Beaufays, F., Ramage, D., 2019. Federated evaluation of on-device personalization. *arXiv preprint arXiv:1910.10252*.

Xu, C., Jiang, H., Peter, Y., Khan, Z., A., Chen, Y., 2020. MHW-PD: A robust rice panicles counting algorithm based on deep learning and multi-scale hybrid window. *Computers and Electronics in Agriculture*, 173:105375.

Xu, C., Lang, W., Xin, R., Mao, K., Jiang, H., 2021. Generative detect for occlusion object based on occlusion generation and feature completing. J*ournal of Visual Communication and Image Representation*, 78:103189.

Yu, T., Bagdasaryan, E., Shmatikov, V., 2020. Salvaging federated learning by local adaptation. a*rXiv preprint arXiv:2002.04758*.