SCIENCE CHINA
Information Sciences

• EDITORIAL •

April 2025, Vol. 68, Iss. 4, 140100:1–140100:2
https://doi.org/10.1007/s11432-025-4334-2

Special topic on cloud-edge collaboration for on-device recommendation*

In numerous e-commerce and online service platforms, cloud-based recommender systems (RSs) have long been regarded as a powerful tool to mitigate information overload and provide personalized experiences for end-users. However, traditional cloud-based RSs face inevitable challenges such as resource-intensive computation, reliance on network access, and privacy breaches. With the growing public demand for algorithmic trustworthiness and emerging privacy legislation, a new paradigm called on-device recommender systems (ODRSs) has become prevalent in recent years. By advocating for a cloud-edge collaboration paradigm that additionally leverages users' devices for learning and deploying RSs, ODRSs retain users' sensitive data on their personal devices, where dedicated resource-efficient recommendation models are deployed to perform on-device inference. To promote further research in this area, we have put forward this special topic section "Cloud-edge collaboration for on-device recommendation" in *SCIENCE CHINA Information Sciences*.

As a typical paradigm of ODRSs, federated learning has been widely adopted to facilitate cloud-edge collaborations, where cross-user federated recommendation is one of the most representative variants. In cross-user federated RSs, the training process of distributed user models is coordinated by the central server, where the cloud-edge communications are facilitated by exchanging information (e.g., model gradients or logits) that does not contain the raw, private user data. In the paper "A survey on cross-user federated recommendation", Yang et al. provide a comprehensive review on the four key aspects concerning cross-user federated RSs, namely their privacy protection, security against adversarial attacks, recommendation accuracy, and resource efficiency. This survey also explores high-potential directions in this line of research, shedding light on the future of federated RSs.

Among different recommendation services, the diagnostic recommendation in e-health is becoming increasingly appealing given its strong convenience and accessibility. However, in terms of privacy, a centralized diagnostic recommendation engine is a suboptimal solution due to the need for collecting and hosting a large number of users' medical records via a cloud server. Thus, Nguyen et al. propose a novel on-device diagnostic recommendation framework in their paper "On-device diagnostic recommendation with heterogeneous federated BlockNets". Built upon the federated learning architecture, the proposed framework utilizes a modular design that allows users to flexibly adjust their local model configurations to fit their on-device computational capacity, and a data-free knowledge distillation mechanism is in place to further optimize each local model's performance.

In ODRSs, to mitigate the scarcity of local training data, a common practice is to use additional self-supervised learning objectives such as contrastive learning. In the paper "Robust federated contrastive recommender system against targeted model poisoning attack", Yuan et al. have uncovered the fact that when contrastive learning is used in federated RSs, the federated RSs will become more vulnerable to adversarial attacks due to the increased embedding uniformity. To respond to this observation, Yuan et al. enhance the vanilla contrastive learning loss function in federated RSs by designing a distance-aware regularizer. As such, items with a similar level of popularity can maintain their mutual distances in the learned embedding space, where experiments well support the improved robustness of federated RSs once trained with the regularizer.

On top of the item interaction history, some recommendation services like food delivery and point-of-interest recommendation will require additional spatiotemporal information from the user. For users who are concerned with privacy breach risks, they are likely to replace fine-grained spatiotemporal information with coarse-grained one, thus leading to deteriorated service quality. In the work "Privacy-preserving recommendation with coarse-grained spatiotemporal contexts", Chen et al. address this challenge by constructing a coarse-grained spatiotemporal knowledge graph with the knowledge from spatiotemporal

* **Citation** Yin H Z, Cui B, Zhou X F, et al. Special topic on cloud-edge collaboration for on-device recommendation. Sci China Inf Sci, 2025, 68(4): 140100, https://doi.org/10.1007/s11432-025-4334-2

co-occurrence patterns and common sense. Owing to the structured knowledge carried by the knowledge graph, the learned user interests are better aligned with the coarse spatiotemporal context, enhancing the recommendation quality for location- and time-sensitive services.

Considering the resource-constrained nature of user devices, a key challenge in ODRSs is to deploy a recommendation model under tight memory budgets without significantly sacrificing its accuracy. For ID-based recommendation models that need to represent each user/item with a unique embedding vector, the heavy parameterization of all embeddings is the main memory bottleneck. Though there are hashing-based solutions that define meta-embeddings and compose unique user/item embeddings by combining different meta-embeddings, in the paper “Coarse-to-fine lightweight meta-embedding for ID-based recommendations”, Wang et al. point out those methods are prone to semantic information loss from the compression process. Thus, a coarse-to-fine meta-embedding learning scheme based on graph neural networks is designed, so as to effectively incorporate semantic information at different granularities and optimize the balance between memory efficiency and accuracy.

The other article that discusses an innovative means of utilizing knowledge graphs in ODRSs is “PerFedKG: two-stage information-loop federated knowledge graph for personalized privacy-preserving recommendation systems”. In this article, Wang et al. identify two major challenges when straightforwardly adopting knowledge graphs in federated RSs, which are the lack of model personalization and semantic incompleteness. As a countermeasure, this article presents a novel federated recommendation framework. This framework seamlessly fuses both global patterns and personalized user preferences into each user’s local model, facilitating personalized model learning. Furthermore, by proposing an information exchange module that aligns each item’s semantic information within both interaction and knowledge graphs, the new federated recommendation framework offers immense promises on its performance.

To conclude, the six articles to be included in this special topic section have demonstrated substantial research advances in ODRSs. We believe the contributions, suggestions, and insights brought by these papers will further stimulate technical innovations in this active research area.

Guest Editors:
Hongzhi YIN
The University of Queensland, Australia
Bin CUI
Peking University, China
Xiaofang ZHOU
The Hong Kong University of Science and Technology, China
Tong CHEN
The University of Queensland, Australia
Quoc Viet Hung NGUYEN
Griffith University, Australia
Xiangliang ZHANG
University of Notre Dame, USA