POI 에서 딥러닝을 이용한 개인정보 보호 추천 시스템

펭소니 [1], 박두순 [1*], 김대영 [1], 양예선 [1], 이혜정 [2], 싯소포호트 [1]
[1] 순천향대학교 소프트웨어융합학과
[2] 순천향대학교 AI · SW 교육원

peng.sony61@gmail.com, parkds@sch.ac.kr, dyoung.kim@sch.ac.kr

Personal Information Protection Recommendation System using Deep Learning in POI

Sony Peng[1], Doo-Soon Park[1*], Daeyoung Kim[1*], Yixuan Yang[1], HyeJung Lee[2], Sophort Siet[1]
[1]Dept. of Software Convergence, Soonchunhyang University, South Korea
[2]Institute for Artificial Intelligence and Software, Soonchunhyang University, South Korea

요 약

POI refers to the point of Interest in Location-Based Social Networks (LBSNs). With the rapid development of mobile devices, GPS, and the Web (web2.0 and 3.0), LBSNs have attracted many users to share their information, physical location (real-time location), and interesting places. The tremendous demand of the user in LBSNs leads the recommendation systems (RSs) to become more widespread attention. Recommendation systems assist users in discovering interesting local attractions or facilities and help social network service (SNS) providers based on user locations. Therefore, it plays a vital role in LBSNs, namely POI recommendation system. In the machine learning model, most of the training data are stored in the centralized data storage, so information that belongs to the user will store in the centralized storage, and users may face privacy issues. Moreover, sharing the information may have safety concerns because of uploading or sharing their real-time location with others through social network media. According to the privacy concern issue, the paper proposes a recommendation model to prevent user privacy and eliminate traditional RS problems such as cold-start and data sparsity.

1. Introduction

Recommendation engines have emerged as one of the most crucial systems in the big data era. The massive data generator from mobile devices and location acquisition technologies have made real-time location information more approachable for users. The tremendous development in location services enables a wide range of research on Location-Based Social Networks (LBSNs), including Yelp, Gowalla, Facebook, Foursquare, and more [1]. People may connect through LBSNs, pin or share current locations, and provide their experiences related to a place called Point-of-Interest (POI). However, with the significantly increased amount of information (user and place interaction), users may be inconvenient to filter their favourite locations. Thus, in this situation, recommendations emerge on the LBSNs in the shape of POI recommendation systems [2].

There are two main types of RSs, including Collaborative filtering (CF), Content-based filtering (CBF). These two mechanisms are the most successful techniques in the flood of information (Big data) era. CBF create the user profile according to the user's past behaviours from explicit or implicit action; it utilizes item characteristics to suggest other items similar to what users previously preferred. Unlike CBF, CF recommends items to the active user through the similarity of the nearest neighbour who shares the same interest. In addition, the CF approach collects and analyzes user data, continuously contributing their interaction data (clicks, likes, dislikes, logs, etc.) to the centralized (cloud) server and trains a machine learning model on gathered interactions to create relevant recommendations. In addition, the interaction between users (similar users) and items (similar items) provides a more successful technique to suggest suitable items to the target users. However, CF fails to recommend items to new users; suppose the users do not have any interaction with the system. In short, this problem is called the cold-start problem [3]. Many researchers are paying much attention to solving the problem and enhancing the system's performance of recommendation system.

Our paper's main task is to propose a Federated Learning model with a Recommendation model. The contribution aims to solve the existing problem of the Traditional RS approach, which happened to Collaborative filtering and the Content-

based approach. Moreover, we would like to protect the user's privacy issue from conventional techniques, and another goal is trying not to push data to centralized storage.

2. Related works

2.1. Problem statement

In this sub-section, we address POI recommendations in LBSNs. The recommendation is based on Spatial and Temporal data [4], as well as the set of user-poi interactions, location-location interaction, user-user interactions, and user context information to form a predictable recommendation model. Figure 1 is just an example of user interaction through real-time check-in to the point of the interest location. The traditional machine learning model requires data (the user action or user-sensitive) to upload to a centralized server to make a prediction; moreover, the data must be stored in centralized storage, which highly impacts the users' privacy issues. According to the privacy problem, google introduced a model which can protect users' privacy. Here, we present a well-known approach called Federated Learning (FL) [5]. The framework enables smart devices (mobile phones, computers) to learn a shared model collaboratively while keeping all the training data on the local device.

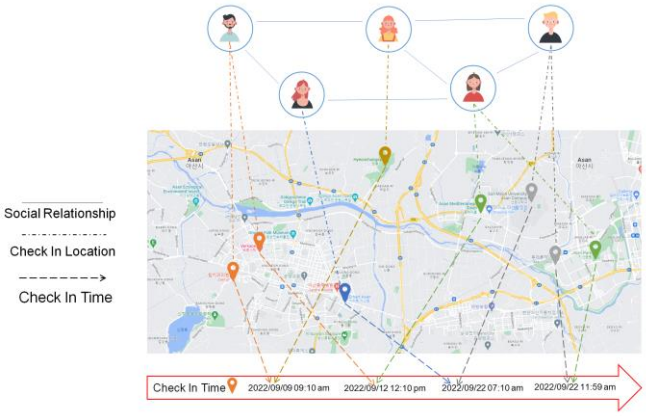

Fig.1. Addressing the location check-in and time of the Users

3. Material and Methods

3.1. Dataset

We select Foursquare to be used in our proposed architecture as it meets our requirements. The dataset is one of the most popular LBSN sites. They were taken from [6], which included user profile data (gender, friends, followers). It comprises 18,291 and 11,874 users who checked in New York City and Tokyo, respectively. According to [7], the corresponding user check-in can be founded in the global check-in dataset. Moreover, both datasets can be linked by the anonymized user ID (unique).

3.2. System Architecture

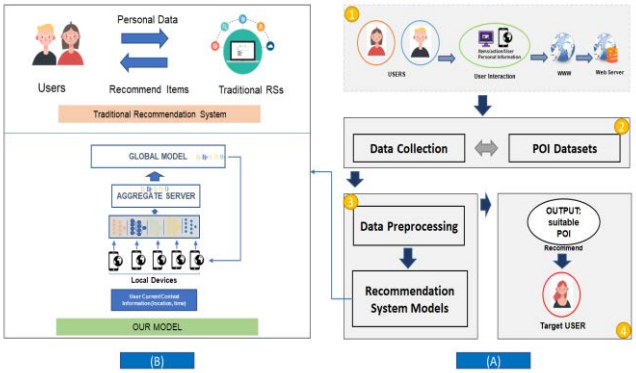

Fig.2. System Architecture of our recommendation system

To solve the mentioned challenges, we propose a model combining a Federated learning mechanism with a recommendation system in our paper. The main concept of our method is to protect user-sensitive data and enhance the recommendation's performance especially; we are giving the solution to alleviate the cold-start issues in the traditional recommendation approach. Figure 2 illustrates our main architecture, which contains two main sections. In the first section (A), we provide the overall flow of the system. For (B), we show the comparison of the traditional recommendation system model with our proposed model. Part B (light orange box) represents the classical recommendation model in that user data is uploaded to the centralized storage to allow machine learning to learn and train the model, which returns the prediction to the user. However, the method is high risk due to the user's information being stored in centralized storage (data or cloud storage). Again, in part B, the green box illustrates our proposed model, which aims to protect user privacy of the recommendation system. The steps of our model are described as follows: User Information: Users' information is not uploaded to the server to train the model; in short, it will be kept and not shared. More certainly, the FL mechanism is locally trained for each user (local devices), and no sensitive data are transferred to the server or other user devices. The deep learning model then transmits the model weight to aggregate, which returns the prediction to local devices through the global model.

4. Conclusion

Recommendation systems improve user living activity and how they look up helpful information from social media sites. Therefore, POI recommendation system plays a vital role in LBSNs. Indeed, the user may provide their identity to the server to get closely content (personalized content). Therefore, it must be a severe privacy issue that impacts users, and it is a problem that the recommender system cannot be ignored. Here, our approach focuses on the model to prevent privacy issues; moreover, we are trying to solve traditional problems such as cold-start and data sparsity.

참고문헌

[1] Frith, J. "Smartphones as locative media". John Wiley & Sons, 2018.

[2] Unger, M., Bar, A., Shapira, B., & Rokach, L. "Towards latent context-aware recommendation systems". Knowledge-Based Systems, 104, 165-178, 2016.

[3] Gope, J., & Jain, S. K. "A survey on solving cold start problems in recommender systems". In 2017 International Conference on Computing, Communication and Automation (ICCCA) (pp. 133-138). IEEE.

[4] Qi, L., Zhang, X., Li, S., Wan, S., Wen, Y., & Gong, W. "Spatial-temporal data-driven service recommendation with privacy preservation". Information Sciences, 515, 91-102, 2020.

[5] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Roselander, J. "Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems", 1, 374-388, 2019.

[6] Yang, D., Zhang, D., Qu, B., & Cudré-Mauroux, P. "PrivCheck: Privacy-preserving check-in data publishing for personalized location-based services". In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 545-556), 2016.

[7] Yang, D., Zhang, D., & Qu, B. "Participatory cultural mapping is based on collective behaviour data in location-based social networks". ACM Transactions on Intelligent Systems and Technology (TIST), 7(3), 1-23, 2016.