SPECTRUM
Journal of Social Sciences

ASPUR PUBLISHING

Vol. 02, No. 03 (2025) 151-156, doi: 10.61552/SJSS.2025.03.003 - http://spectrum.aspur.rs

PRIVACY-PRESERVING COLLABORATIVE FILTERING FOR PERSONALIZED RECOMMENDATIONS

Yogita Sawant [1]
Pallavi Thigale
Harsha Patil
Sakshi Thigale
Vikas Mahandule

Received 03.08.2024.
Revised 02.09.2024.
Accepted 05.10.2024.

Keywords:

Homomorphic encryption, Federated learning, Differential privacy, Secure multiparty computation (SMC), Privacy-aware user preference modeling , Anonymized data sharing, Decentralized privacy-preserving recommendation models, Privacy-enhanced collaborative filtering, Encrypted user data, Secure aggregation.

A B S T R A C T

Personalized recommendations have become very important for many online platforms. They aim to make the user experience better by offering content that matches each user's individual interests. This paper discusses the importance of personalized recommendations in different areas and their impact on keeping users engaged and satisfied. Through looking at existing research and case studies, we identify key challenges and opportunities in developing effective personalized recommendation systems. We also propose strategies to overcome these challenges and make personalized recommendations more effective. The other parts introduce the proposed system design, the technologies we will use, and how we will evaluate the methods. The paper concludes by presenting our current research progress.

Personalized recommendations have become a cornerstone of many online platforms, aiming to enhance user experience by offering content tailored to individual preferences. This paper explores the significance of personalized recommendations in various domains and discusses their impact on user engagement and satisfaction. Through an analysis of existing research and case studies, we identify key challenges and opportunities in developing effective personalized recommendation systems. Furthermore, we propose strategies to overcome these challenges and enhance the effectiveness of personalized recommendations. The other parts introduce the proposed system architecture, the technologies that we will use, and the evaluation methods. The paper is concluded by presenting the current research progress.

1. INTRODUCTION

In today's digital age, individuals are inundated with an abundance of information and content across various online platforms (Barnett et al., 2020). Personalized recommendation systems address this challenge by leveraging user data and behavioral patterns to deliver tailored content recommendations (Deldjoo et al., 2020). These systems have gained immense popularity across diverse domains, including e-commerce, streaming services, social media, and news platforms. etc (Shankar et al., 2024). Some of the benefits of our scheme could be obtained by pseudo-identities. Users could use persistent online identities that do not link to their actual

[1] Corresponding author: Yogita Sawant
Email: sawanty295@gmail.com

identity, and make anonymous purchases (Oyserman & Schwarz, 2020). This approach addresses several of the difficulties with traditional collaborative filtering systems. While it does prevent vendors from gathering customer data. Today’s server-based collaborative filtering systems have a number of disadvantages (Khan, et al., 2020; Rafique et al., 2023). First of all, they are a serious threat to individual privacy. In today's digital age, individuals are inundated with an abundance of information and content across various online platforms (Barnett, et al., 2020). Personalized recommendation systems address this challenge by leveraging user data and behavioral patterns to deliver tailored content recommendations (Wu et al., 2023; Chen et al., 2024). These systems have gained immense popularity across diverse domain. today's digital world, people are overwhelmed with too much information and content online. Users often struggle to find relevant content that matches their interests and preferences (Stray et al., 2021). Personalized recommendation systems solve this problem by using user data and behavior patterns to deliver customized content recommendations (Javed et al., 202; Ko et al., 2022). These systems are very popular across different areas, including e-commerce, streaming services, social media, and news platforms.

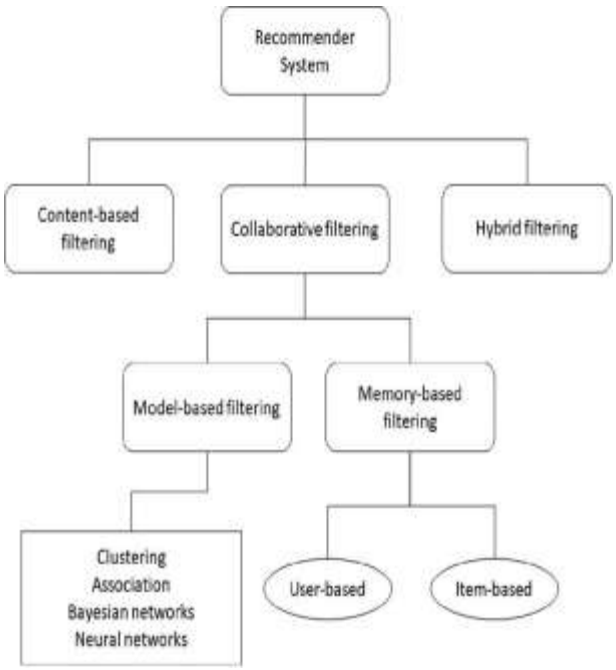

Figure 1. Recommendation systems

Recommendation systems are gaining popularity for online services, providing personalized recommendations for various products (Figure 1). These services have become effective revenue drivers for online businesses, with 35% of Amazon purchases and 75% of Netflix viewings attributed to personalized recommendations. Existing systems typically involve collaborative efforts to support personalized recommendations.

2. METHODOLOGY

However, the model requires a lot of domain knowledge due to the hand-engineered feature representation of items, and can only make recommendations based on existing user interests, articles, case studies, and industry reports to gain insights into the challenges, trends, and best recommends items relevant to the user, using a similarity metric like dot product.

Additionally, we interview experts in machine learning, data science, and user experience to gather diverse perspectives.

Personalized Recommendation Techniques: Personalized recommendations are vital in various industries like e-commerce, streaming, social media, and more. They aim to provide users with content, products, or services tailored to their preferences, behaviors, and interests.

Recommendation Systems Benefits and Use Cases: There are several key benefits of recommendation systems, including. Personalization: Recommendation systems provide personalized recommendations to users, based on their past behavior, preferences, and interactions with the system. Personalized recommendation techniques are essential in various industries, such as e-commerce, streaming services, and social media, to provide users with content, products, or services tailored to their preferences, behaviors, and interests. Common techniques used for personalized recommendations include content-based filtering, collaborative filtering, and hybrid recommendation. Netflix, for example, uses these techniques.

Recommendation systems offer several benefits, including personalization based on past behavior, preferences, and interactions with the system. This allows users to receive tailored recommendations based on their individual needs and interests. This increased engagement can lead to increased sales or conversions. Additionally, recommendation systems improve the user experience by making it easier for users to find the content they are looking for, reducing frustration and increasing user satisfaction.

- Personalization: Recommendations tailored to each user's needs and interests.
- Increased engagement: Providing interesting recommendations keeps users engaged.
- Improved user experience: Makes it easier for users to find what they want.
- Increased revenue: Personalized recommendations can lead to more sales.

2.1 Content –Based Filtering

Content-based filtering is a technique that predicts a user's preferences based on past preferences, such as genre, director, or actor. This approach can be applied to any item with explicit or implicit attributes, including books, movies, music, and products. The algorithm uses user's past behavior, such as ratings or purchase history, to build a profile of the user's preferences, which is then used to recommend items with. To build a content-based recommendation system, basic concepts and techniques, such as extraction, similarity measures, and recommendation algorithms, are used. The system uses agenre column and vectorizers like binary feature matrix, bags of words, and tf-idf, and adds tags for movies using genome-tags and genome-score files from the movie lens data. The model limiting its ability to expand on users' existing interests. It focuses on item attributes like genre, director, actor, etc.

2.2 Collaborative Filtering

Collaborative filtering is a method used in recommender systems to produce new recommendations based on past interactions between users and items. It aims to find similar users and recommend items based on their preferences. Memory-based approaches work with recorded data and are based on nearest neighbors search. Model-based collaborative filtering uses machine models to predict customer ratings and predict unrated products. These algorithms are divided into subsets, such as matrix factorization-based algorithms, deep learning methods

3. COLLABORATIVE FILTERING PROCESSES OVERVIEW

3.1 Model-based Collaborative Filtering

- Uses machine learning algorithms to make rating predictions.
- Develops a compact model of ratings data by uncovering latent factors.
- Operates over entire user-item data directly, but can suffer from scalability issues.
- Can uncover latent factors explaining observed ratings and detect signals in the data.
- Higher accuracy for dense data, preferred for sparse datasets.
- Both techniques have strengths, with many practical recommender systems combining both approaches.Examples: of collaborative filtering include suggesting famous and interesti collaborative filtering. Both methods require machine learning and can be learned automatically without relying on hand-engineering of features information on Reddit, which are voted positively by a group of people. As the group becomes more diverse, the publicized stories show better community interest. Wikipedia is another application of Collaborative filtering is based solely on past interactions between users and items to produce new recommendations. It finds what similar users would like and recommends accordingly.
- There are two types: memory-based (calculates user/item similarity from data) and model-based (uses machine learning to predict ratings).

3.2 Hybrid Filtering

Hybrid systems combine content-based and collaborative filtering approaches to leverage their advantages. Hybrid recommendation systems are a type of fuzzy recommendation system (FRS) that combine the capabilities of traditional centralized and distributed systems. They offer personalized recommendations while protecting user privacy by storing data in a central repository and executing the recommendation algorithm by a federated learning system. This ensures security and privacy by not sharing user data between the two systems. Hybrid systems can use different algorithms, such as collaborative filtering and contentbased filtering, and can adapt to different data sources like social media and web analytics. They also offer scalability, allowing algorithm updates without sharing additional user data. However, the primary privacy concern is the extensive collection of user data from various sources, including sensitive user behavior, profile, and contextual data. To mitigate data breaches and ensure user privacy, these systems must prioritize robust privacy-preserving techniques like data anonymization, homomorphic encryption, and differential privacy. This approach allows for more efficient and effective systems while accommodating more users. Hybrid recommendation systems use various sources of information, including demographic, contextual, and external data, to make more accurate recommendations. These systems predict user preferences based on objectives like clicks, revenue, and sales. Collaborative Filtering and Content-Base Filtering methods are commonly used, while Knowledge-Based recommender systems are based on user requirements. Hybrid systems combine these approaches for a robust performance across various applications. Hybrid Recommendation Systems Overview.

4. RECOMMENDER SYSTEMS

Recommender systems are used by communities to help individuals find information or products that are most likely to be interesting or relevant to their needs. The system consists of two entities: the user, who provides their opinion and receives recommendations, and the item, which is rated by users. The input data is usually arithmetic rating values, and the outputs can be predictions or recommendations.

The three main processes of recommender systems are representation, neighborhood formation, and recommendation generation. In the original representation, the input data is de-fined as a collection of numerical ratings of m users on n items, expressed by the $m*n$ user-item matrix R. Techniques have been proposed to reduce the sparsity of the initial user-item matrix.

Neighborhood formation involves determining the similarity between users in the user-item matrix, which is then used to estimate their possible preferences. The final step in the recommendation process is to produce either a prediction or a recommendation based on the neighborhood of users.

In a scenario where two e-commerce entities, like online bookstores, want to cooperate to strengthen their recommender systems and improve the precision of their recommendations for their customers, merging the databases without disclosing the actual commercial data is a challenge.

The paper introduces two approaches to secure multiparty computation: the homomorphic encryption based approach and the scalar product based approach. The homomorphic encryption framework uses ElGamal and Paillier homomorphic encryption, which preserve user preference privacy by encrypting data with a public key. The scalar product approach addresses the problem of Secure Multiparty Computation (SMC) by proposing a secure two-party product protocol. The paper compares the performance of both approaches using the Jester Joke Recommender System, a database with 4.1 million continuous ratings of jokes. The ElGamal algorithm is found to be about five times faster in a single operation than the Paillier algorithm. The paper concludes that these approaches can be used to maintain privacy in large-scale applications, but their prohibitive cost in protecting privacy makes them unsuitable for large-scale application.

Revised commodity approach: Revised commodity approach is a method for multi-party computation where all necessary data is prepared before executing tasks. This involves pre-producing and transporting random numbers to A and B before executing tasks, allowing for efficient communication between parties. The commodity server, a neutral third party, is not necessarily included in computing transportation, CPU, and total execution time. Both parties exchange information in one round, resulting in the lowest time cost and least computing resources among the three approaches. However, additional storage is needed for random numbers, and stored random numbers must not be disclosed to each party. The transportation time line for ElGamal's approach is not linear, possibly due to Ruby's fixed-size buffer for IO. The total time cost for a 100-dimensional scalar product is shown in the following table. The revised commodity approach has the best performance in preserving privacy and accuracy in collaborative filtering.

5. CONCLUSION

This overview presents different recommendation system methodologies.

Privacy-preserving collaborative filtering is a promising approach to address privacy concerns while still providing personalized recommendations. Despite challenges, it has potential for enhancing user privacy. The document suggests privacy-preserving collaborative filtering as a promising solution for personalized recommendation systems, using cryptographic protocols like homomorphic encryption. Despite challenges like computational complexity and potential trade-offs in recommendation accuracy, it suggests further research and development to enhance user privacy protections in recommendation systems.

The below is the node displacement summary in the table 1.

All Summary

			Horizontal	Vertical	Horizontal	Resultant	Rotational		
	Node	L/C	X mm	Y mm	Z mm	mm	rX rad	rY rad	rZ rad
Max X	257	122	109.838	-0.832	-1.437	109.850	-0.000	0.002	0.001
Min X	253	123	-109.848	-0.832	-1.438	109.861	-0.000	-0.002	-0.001
Max Y	78	103	-12.758	77.912	-103.845	130.448	-0.011	-0.000	0.000
Min Y	233	101	0.002	-643.827	72.669	647.915	0.009	0.000	0.000
Max Z	799	102	12.807	-0.107	271.860	272.125	-0.003	0.000	-0.000
Min Z	799	109	-0.576	-0.602	-284.610	284.611	0.006	-0.001	0.000
Max rX	22	124	-0.554	-1.779	112.075	112.091	0.035	0.000	-0.000
Min rX	840	109	0.000	0.000	0.000	0.000	-0.029	0.008	-0.000
Max rY	848	109	0.000	0.000	0.000	0.000	-0.025	0.022	0.000
Min rY	844	109	0.000	0.000	0.000	0.000	-0.025	-0.022	-0.000
Max rZ	877	101	-20.656	-39.977	31.982	55.206	-0.019	-0.003	0.043
Min rZ	886	101	20.738	-39.980	31.973	55.231	-0.019	0.003	-0.043
Max Rst	233	101	0.002	-643.827	72.669	647.915	0.009	0.000	0.000

Table 1. Node displacement summary

The algorithm uses a stochastic gradient descent-based approach and achieves robust and stable solutions. The federated model provides similar recommendation performance as the standard method while preserving user privacy. Future research aims to explore multiple directions, including simulator-based analysis, communication payloads, efficiency, and security in federated models. This work proposes a personalized privacy-preserving recommendation system using a Trusted-based Agent Network (TAN) and differential privacy in data communication. The system mainly relies on a small portion of the network, the local trusted network, which contains trusted people or experts. The global view is transmitted only to servers, avoiding direct interaction with real data. The system generates overall recommendation results based on local and global views. The privacy-preserving approach can predict user preferences better than traditional collaborative filtering systems while preserving privacy. The run time complexity is linear, slightly higher than offline but lower than online versions.

References:

Barnett, M. L., Henriques, I., & Husted, B. W. (2020). The rise and stall of stakeholder influence: How the digital age limits social control. *Academy of Management Perspectives*, *34*(1), 48-64.

Chen, J., Liu, Z., Huang, X., Wu, C., Liu, Q., Jiang, G., ... & Chen, E. (2024). When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web*, *27*(4), 42.

Deldjoo, Y., Schedl, M., Cremonesi, P., & Pasi, G. (2020). Recommender systems leveraging multimedia content. *ACM Computing Surveys (CSUR)*, *53*(5), 1-38.

Javed, U., Shaukat, K., Hameed, I. A., Iqbal, F., Alam, T. M., & Luo, S. (2021). A review of content-based and context-based recommendation systems. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(3), 274-306.

Khan, A., Ahmad, A., Rahman, A. U., & Alkhalil, A. (2020). A mobile cloud framework for context-aware and portable recommender system for smart markets. *Smart Infrastructure and Applications: Foundations for Smarter Cities and Societies*, 283-309.

Ko, H., Lee, S., Park, Y., & Choi, A. (2022). A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics*, *11*(1), 141.

Oyserman, D., & Schwarz, N. (2020). Identity-based motivation and the logic of conversations obfuscate loss of online privacy and what policy-makers can do about it. *Journal of Consumer Psychology*, *30*(4), 759-766.

Rafique, W., Hafid, A. S., & Qadir, J. (2023). Developing smart city services using intent-aware recommendation systems: A survey. *Transactions on Emerging Telecommunications Technologies*, *34*(4), e4728.

Shankar, A., Perumal, P., Subramanian, M., Ramu, N., Natesan, D., Kulkarni, V. R., & Stephan, T. (2024). An intelligent recommendation system in e-commerce using ensemble learning. *Multimedia Tools and Applications*, *83*(16), 48521-48537.

Stray, J., Vendrov, I., Nixon, J., Adler, S., & Hadfield-Menell, D. (2021). What are you optimizing for? aligning recommender systems with human values. *arXiv preprint arXiv:2107.10939*.

Wu, C., Wu, F., Huang, Y., & Xie, X. (2023). Personalized news recommendation: Methods and challenges. *ACM Transactions on Information Systems*, *41*(1), 1-50.

Yogita Sawant
MIT ACSC Alandi pune,
India
sawanty295@gmail.com

Sakshi Thigale
MIT ACSC Alandi pune,
India
sakshithigale2018@gmail.com

Pallavi Thigale
MIT ACSC Alandi pune,
India
pallavithigale3936@gmail.com

Vikas Mahandule
MIT ACSC Alandi pune,
India
vnmahandule@mitacsc.ac.in
ORCID: 0009-0007-5415-9227

Harsha Patil
MIT ACSC Alandi pune,
India
hvpatil@mitacsc.ac.in
ORCID: 0000-0001-6519-9987