ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Enhanced Deep Reinforcement Learning based Group Recommendation System with Multi-head Attention for Varied Group Sizes

Saba Izadkhah and Banafsheh Rekabdar

Portland State University - Department of Computer Science - Oregon - USA

Abstract. This paper introduces EnGRMA, an Enhanced deep reinforcement learning-based Group Recommendation system with Multi-head Attention for varied group sizes. EnGRMA adapts its recommendation strategy according to group sizes, using individual member preferences in smaller groups through a weighted average method, and leveraging multi-head attention to aggregate diverse opinions effectively in larger groups. This method helps model dynamic member-item interactions, enhancing the system's ability to deliver personalized recommendations. Our evaluation of the MovieLens-Rand dataset shows that EnGRMA not only outperforms GRMA and DRGR in Recall, NDCG, Precision, and F1 scores but also demonstrates superior performance in NDCG against AGREE.

1 Introduction

In response to the evolving landscape of online platforms and the increasing prevalence of group decision-making scenarios, there's a growing emphasis on enhancing group recommender systems [1]. Recent advancements, including the integration of attention networks [2], aim to address the complexities of group dynamics, offering more personalized and effective recommendations. However, existing approaches often rely on fixed strategies that may not fully capture the intricacies of group interactions.

To address these limitations, this study introduces EnGRMA, which is designed for groups of different sizes. For smaller groups, EnGRMA prioritizes individual member preferences and interactions using a weighted average of individual member preferences, where each member's preference has a higher weight. This can lead to more personalized recommendations within the group. For larger groups, the model shifts its focus to collective preferences and uses a multi-head attention mechanism to capture a wider range of opinions and interactions within the group, which increases system consistency and provides more personalized and context-aware recommendations. Our approach builds on the strengths of previous research in the GRMA (Deep Reinforcement Learning based Group Recommendation System with Multi-head Attention Mechanism) [1], DRGR (Deep Reinforcement Learning based Group Recommender System) [3], and AGREE (Attentive Group Recommendation) [4], while addressing limitations and challenges identified in the existing literature. The paper is structured as follows: Section 2 describes the proposed methodology. Section 3 provides the experimental setup and results, and the final Section 4 provides the conclusion.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

2 The Proposed Method

The EnGRMA is introduced as a significant enhancement over the previous GRMA [1]. For clarity, The EnGRMA method tailors policies for different group sizes, utilizing a weighted aggregation strategy for smaller groups and a multi-head attention mechanism for larger groups, which allows for more accurate modeling of group dynamics. In this section, we first simulate an online environment. Next, we provide an overview of the proposed reinforcement learning framework based on the actor-critic architecture.

Framework Overview. The proposed framework highlights the utilization of Markov Decision Processes (MDPs) and Reinforcement Learning (RL) to enhance recommendation strategies, with training conducted via the Deep Deterministic Policy Gradient (DDPG) algorithm [5] in addition to the Ornstein-Uhlenbeck Noise (OUNoise) [6]. We formally define the tuple (S, A, P, R, γ) of the Markov Decision Process (MDP) where at each time step, an agent based on state $s_t \in S$ takes an action $a_t \in A$ and receives a reward $r(s_t, a_t)$. We define U, G, and I to represent the set of users, groups, and items, respectively [1].

Simulating an online environment. The environment simulator [3], determines the next state s_{t+1} based on h_{t+1} and the reward r_t. The state h_{t+1} is updated using the action a_t, which represents the group recommendation. If $r_t > 0$, h_{t+1} is updated to incorporate the recommended items; otherwise, h_{t+1} remains unchanged. The next state s_{t+1} consists of the combination of g (group information) and h_{t+1}. Using this approach, by utilizing h_{t+1}, s_{t+1}, and r_t allows the simulator to predict unknown ratings and simulate online rewards.

Actor-Critic Network. The Actor network processes embedded states s_t to generate action weights w_t for each item, providing accurate and timely recommendations aligned with group preferences. Simultaneously, the Critic network evaluates the quality of these recommendations by estimating the state-action value $Q(s_t, a_t)$, enabling continuous refinement of recommendations over time based on received rewards. Both networks benefit from optimization through the Adam optimizer [7].

Embedding Network. In EnGRMA, for smaller groups, we use a weighted average aggregation strategy, while for larger groups, we use the multi-head attention mechanism. This dual approach ensures the embedding state, a combination of the current state of group s_t and the selected action a_t at a given time step t. In both scenarios, by merging the action weights w_t generated by the Actor with the state-action value $Q(s_t, a_t)$ estimated by the Critic, the embedding state offers a comprehensive group preferences to capture the local and global patterns in the group's decision-making process.

Multi-Head Attention Mechanism. For larger groups, EnGRMA uses a multi-head attention mechanism [8] that takes queries, keys, and values as input and subsequently calculates attention scores through the dot product between the projected queries and keys. Then, these attention scores are utilized to compute weighted sums of the corresponding values. Let $G = \{u_1, u_2, \ldots, u_n\}$ represent a group of users with embeddings and $I = \{i_1, i_2, \ldots, i_m\}$ represent a

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

set of candidate items for recommendation by their embeddings. The multi-head attention can be formulated as follows:

Projection of Queries, Keys, and Values: For each group member u_i, we project their embedding in three spaces using learned linear transformations:

$$\text{Query}(q_i) = W_q \cdot u_i, \text{Key}(k_i) = W_k \cdot u_i, \text{Value}(v_i) = W_v \cdot u_i \tag{1}$$

where W_q, W_k, and W_v are learnable weight matrices for each attention head. Attention scores, for each attention head h, between queries and keys can be calculated using the scaled dot-product attention mechanism [8]:

$$\text{Attention}(h) = \text{softmax}\left(\frac{\text{Query} \cdot \text{Key}^T}{\sqrt{d_k}}\right) \cdot \text{Value} \tag{2}$$

where d_k is the dimension of key vectors ($d_k = \dim(\text{Key})$).

Multi-head attention heads can capture diverse interactions and with that multi-head aggregation can be formulated as below by concatenating the outputs of all attention heads [8]:

$$\text{Aggregated Attention} = [\text{Attention}(1), \text{Attention}(2), \ldots, \text{Attention}(H)] \tag{3}$$

The final aggregated display of the candidate items is obtained as the weighted sum of the values based on the attention scores:

$$\text{Final Recommendation} = \text{Aggregated Attention} \cdot \text{Value}^T \tag{4}$$

The aggregated representation, which considers the preferences and interactions of all group members with the candidate items, serves as the final recommendation for the group.

Weighted Average Aggregation Strategy. For smaller groups, EnGRMA uses a weighted average strategy [9] where each member's preferences are given a specific weight that indicates their importance in the group's decision-making process. This strategy [9] implements as follows:

Member Embeddings: Each group member u_i is represented by an embedding vector $E(u_i)$.

Weight Assignment: Assign a weight w_i to each member, where w_i represents the relative importance or influence of member u_i in the group. These weights are normalized to ensure:

$\sum_{i=1}^{n} w_i = 1, \quad$ where n is the number of group members.

Weighted Average Calculation: Compute as:

$$E_{\text{group}} = \sum_{i=1}^{n} w_i \cdot E(u_i) \tag{5}$$

Here, E_{group} represents the aggregated embedding of the group that effectively captures the combined preferences of its members.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

3 Experimental Setup and Results

Data Set and Group Generation. In EnGRMA, we utilize the public MovieLens 1M dataset[1], which includes over a million ratings for more than 6,000 users and almost 4,000 movies. To prepare this dataset, we created the MovieLens-Rand dataset [10] formed by randomly grouping users with different sizes of two, three, four, and five users, and then generates ratings for these groups based on the user-item ratings, and assuming a movie is favored by a group if all members rate it with four stars or above. Table 1 reports statistics of the MovieLens-Rand dataset. We focused on users with at least 20 interactions to ensure sufficient data to evaluate long-term recommendation performance. The dataset was randomly split into training, validation, and testing sets with proportions of 70%, 10%, and 20%, respectively by temporal order. The groups in our setup are ad-hoc, i.e., a group may appear only in test data, but not in training data. This approach highlights EnGRMA's ability to address cold-start challenges in group recommendations. Each test user was a cold-start user, meaning that no historical information was provided to the recommender system beforehand. However, DRL-based techniques gradually infer user preferences, which can be applied to cold-start users without prior interaction or other peripheral information.

Total Users	3,711	**Total Groups**	5,000	**Total Items**	2,623
Avg. Group Size	2.18	**Avg. Rating/User**	229.76	**Avg. Rating/Item**	54.40

Table 1: Dataset statistics of the MovieLens-Rand

Evaluation Setup. In EnGRMA model, we employ four key metrics [10] [11]: Recall (Recall@K), normalized discounted cumulative gain (NDCG@K), Precision (Precision@K), and F1 Score (F1 Score@K), with K at 20 and 50, represents the number of recommendations. These metrics are averaged across all testing groups and a higher value indicates better accuracy within top K recommendations. Key hyperparameters are shown in Table 2. All models are trained on the training set, tuned on the validation set, and evaluated on the testing set.

Hyperparameter	Value	Hyperparameter	Value	Hyperparameter	Value
Batch Size	64	Group Size	[2-5]	History Length	5
Learning Rate	0.001	Attention Heads	16	Epochs	1000
Embedding size	32	Discount Rate (γ)	0.9	Actor Hidden Size	(128,64,32)
Dropout Rate	0.2	Negative Sample Size	50	Critic Hidden Size	(32,16)

Table 2: Hyperparameter settings for EnGRMA

Overall Performance Comparison. In evaluating EnGRMA, we compare its performance with three baselines including GRMA [1], AGREE [4], and DRGR [3]. Figure 1 represents evaluated metrics such as Recall@K, NDCG@K, Precision@K, and F1 Score@K for $K = 20$ and 50. EnGRMA model demonstrates enhanced performance compared to the baseline GRMA and DRGR models when

[1] `https://grouplens.org/datasets/movielens/`

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2.
Available from http://www.i6doc.com/en/.

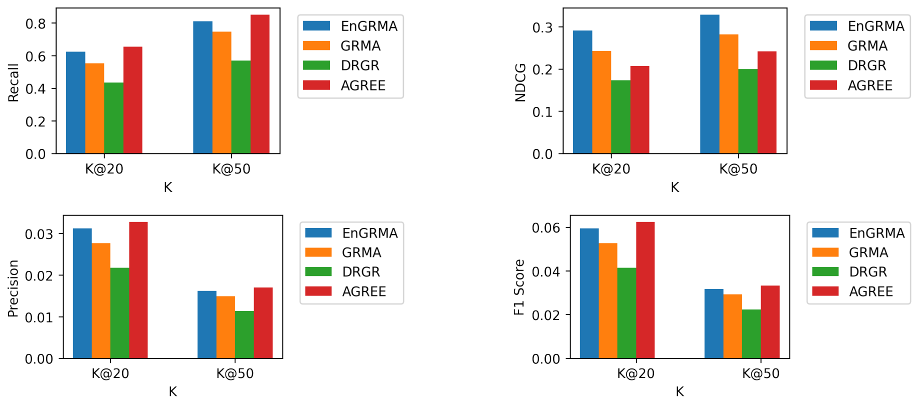

Fig. 1: Performance of group recommendation methods in terms of Recall, NDCG, Precision, and F1 Score at K = 20 and 50

evaluated by all metrics. EnGRMA also shows superior performance compared to AGREE in NDCG@K metric for both $K = 20$ and 50 indicating its effectiveness in ranking relevant items. The tailored approach of the EnGRMA model for different group sizes from small to large groups leads to a significant improvement in the quality of the recommendations. Table 3 compares the performance of EnGRMA with DRGR and AGREE and table 4 compares the performance of EnGRMA with GRMA.

Model	**Recall**		**NDCG**		**Precision**		**F1 Score**	
	@20	@50	@20	@50	@20	@50	@20	@50
AGREE	0.65442	0.84952	0.20674	0.24158	0.03271	0.01698	0.06233	0.03332
DRGR	0.434134	0.569206	0.172883	0.200075	0.021701	0.01138	0.041346	0.022324
EnGRMA	0.623428	0.808839	**0.290788**	**0.328054**	0.031173	0.016177	0.059377	0.031715

Table 3: Group recommendation results of AGREE, DRGR, and EnGRMA, on Recall, NDCG, Precision, and F1 Score metrics at K = 20 and 50

Model	**Recall**		**NDCG**		**Precision**		**F1 Score**	
	@20	@50	@20	@50	@20	@50	@20	@50
GRMA	0.553445	0.747304	0.24287	0.28166	0.027675	0.014944	0.052708	0.029307
EnGRMA	**0.623428**	**0.808839**	**0.290788**	**0.328054**	**0.031173**	**0.016177**	**0.059377**	**0.031715**

Table 4: Group recommendation results of GRMA and EnGRMA on Recall, NDCG, Precision, and F1 Scoew metrics at K = 20 and 50

4 Conclusion

In conclusion, the proposed EnGRMA system significantly improves group recommender systems by effectively adapting to different group sizes. It effectively combines individual preferences in smaller groups with a weighted average aggregation and employs multi-head attention mechanisms for larger groups to handle diverse opinions. The system's adaptability in handling varying group dynamics and preferences leads to more personalized and context-aware recommendations. EnGRMA's robustness and effectiveness in diverse group scenarios position it as a compelling model for group recommendation tasks, outperforming existing models like GRMA and DRGR in all evaluation metrics and is also better than AGREE, especially in terms of NDCG@K, because the integration of the multi-head attention mechanism and weighted average aggregation strategy in EnGRMA significantly improves the ranking quality and ensures that more relevant items are placed at the top of the recommendation list, leading to higher NDCG scores.

References

[1] S. Izadkhah and B. Rekabdar, "Deep reinforcement learning based group recommendation system with multi-head attention mechanism," in *The Fifth International Conference on Transdisciplinary AI (TransAI)*, 2023.

[2] J. Zhu, Z. Li, C. Yue, and Y. Liu, "Trust-aware group recommendation with attention mechanism in social network," in *The 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*, pp. 271-276, 2019.

[3] Z. Liu, S. Wen, and Y. Quan, "Deep Reinforcement Learning based Group Recommender System," arXiv preprint arXiv:2106.06900, 2021.

[4] D. Cao, X. He, L. Miao, Y. An, C. Yang, and R. Hong, "Attentive Group Recommendation," in *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '18)*, pp. 645-654, New York, NY, USA, 2018.

[5] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N.Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.

[6] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, no. 3, pp. 823-841, 1930.

[7] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Computer Science*, 2014.

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, pp. 6000-6010, Long Beach, CA, USA, December 2017.

[9] T. Calvo, R. Mesiar, and R. R. Yager, "Quantitative Weights and Aggregation," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 1, February 2004.

[10] L. Baltrunas, T. Makcinskas, and F. Ricci, "Group recommendations with rank aggregation and collaborative filtering," in *The 2010 ACM Conference on Recommender Systems (RecSys 2010)*, pp. 119-126, Barcelona, Spain, September 2010.

[11] T. N. Pham, X. Li, G. Cong, and Z. Zhang, "A General Recommendation Model for Heterogeneous Networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 12, pp. 3140-3153, December 2016.