様式第１号その１（修士）

学 位 授 与 申 請 書

2024年　2月　9日

大阪公立大学長　様

大阪公立大学大学院 　　情報学研究科 　基幹情報学専攻

氏　名　　毛利　憲竜

大阪公立大学学位規程第５条第１項の規定により

修士（ 情報学 ）の学位の授与を申請します。

（注意）

１．　この申請書は、1通提出すること。

　　（A4）

注：　右肩の日付は西暦年表示。

粒状性を考慮したラフ集合ベースの混合多項分布型

共クラスタリングに基づく協調フィルタリング

Collaborative Filtering Based on Rough Set-Based Co-clustering Induced by Multinomial Mixture Models Considering Granularity

|  |  |  |  |
| --- | --- | --- | --- |
| 専攻名  （分野） | 基幹情報学専攻 | 氏 名  （氏名） | 毛利　憲竜 |
| Department | Department of Core Informatics | Name | Kenryu Mori |

In clustering-based collaborative filtering (CF), clusters of users with similar preference patterns are extracted, and items with high preferences within the cluster are recommended. Since data in CF tasks contain uncertainties arising from human sensibilities, represented as co-occurrence relationships between users and items, approaches such as rough clustering and co-clustering can be effective. Thus, rough co-clustering induced by multinomial mixture models (RCCMM) and its application to CF (RCCMM-CF) have been proposed.

However, RCCMM has a problem in that it does not consider the granularity, an important viewpoint in rough set theory. Therefore rough set CCMM (RSCCMM) was proposed as a rough co-clustering method that considers the granularity of the target space.

In this study, we propose a CF method based on RSCCMM (RSCCMM-CF) that considers the granulatity of the user set in CF tasks. Furthermore, we verified the recommendation performances of the proposed methods by applying it to real-world datasets, namely, NEEDS-SCAN/PANEL dataset and MovieLens-100k dataset. In addition, we discuss the effect of granulation in rough set theory on CF tasks by comparing the proposed method with conventional CF based on HCCMM (HCCMM-CF) and RCCMM-CF. We used ROC-AUC as an evaluation indicator of the recommendation performance.

From the result of numerical experiments, in both datasets, we confirmed that RSCCMM-CF obtained better recommendation performance than HCCMM-CF and RCCMM-CF by adjusting the roughness parameter and it maintains high performance regardless of the initial number of clusters . Also, in the MovieLens-100k dataset, we confirmed that the recommendation performance was better when the data was binarized and the missing values were imputed with 0.5 than when the missing values were imputed with the average evaluation value of each user.

In summary, the results show that consideration of granularity based on rough set theory is effective for the CF tasks. We expect that the preprocessing of missing values and the analysis of clustering results in this research will advance and lead to even more effective collaborative filtering methods. グラフ

自動的に生成された説明

Figure 1: Changes in AUC by the initial number of clusters, C, in various methods in NEEDS-SCAN/PANEL dataset