Efficient models to identify groups in group recommender systems

Thesis submitted in partial fulfillment
for the award of degree

Doctor of Philosophy

by

Chintoo Kumar

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
Indian Institute of Technology
(Banaras Hindu University)
Varanasi

Roll No: 17071017 2022

To

The

Almighty God, My Family, Friends, Guruji..

CERTIFICATE

It is certified that the work contained in the thesis titled **Efficient models to identify groups in group recommender systems** by **Chintoo Kumar** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree. It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

Supervisor

Dr. Ravindranath Chowdary C

Dept. of Computer Science & Engineering

IIT(BHU)

Varanasi - 221005

DECLARATION

I, **Chintoo Kumar**, certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of **Dr Ravindranath Chowdary C** from July-2017 to November-2022, at the Departement of Computer Science & Engineering, Indian Institute of Technology (BHU), Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports dissertations, theses, etc., or available at websites and have not included them in this thesis and have not cited as my own work.

Date: 26-04-2023

Place: Bengaluru

Chintoo Kumar

(Chintoo Kumar)

CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my/our knowledge.

Supervisor

Dr. Ravindranath Chowdary C

IIT(BHU), Varanasi

Department of Computer Sc. & Engg

Indian Institute of Technology

(काशी हिन्दू विश्वविद्यालय)

(Banaras Hindu University)

वाराणसी/Varanasi-221005

VK Singh

Signature of Head of Department/Coordinator of School

Professor & Head

Department of Computer Sc. & Engg

Indian Institute of Technology

(Banaras Hindu University)

वाराणसी–221005 / Varanasi-221005

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: **Efficient models to identify groups in group recommender systems**
Name of Student: **Chintoo Kumar**

Copyright Transfer

The undersigned hereby assigns to the Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the Doctor of Philosophy.

Date: 26-04-2023

Place: Bengaluru

Chintoo Kumar
(Chintoo Kumar)

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

Acknowledgments

Though only my name appears on this dissertation's cover, many great people have contributed to accomplishing this job. I owe my gratitude to all who have made this thesis possible and because of whom my post-graduate experience has been one that I will cherish forever. It took a lot of passion, drive and perseverance to make it to the finish line.

I take this opportunity to express my profound gratitude and deep regards to my supervisor Dr. Ravindranath Chowdary C, Associate Professor, Computer Science and Engineering Department, IIT(BHU), Varanasi, for his moral guidance, encouragement and constant support throughout this dissertation. His excellent management skills, strong work ethic, and kind heart set a perfect example for me to follow, which indeed pushed me to nurture similar characters in my life. I am indebted and grateful for his invaluable guidance, advice, and mental and academic support throughout my candidature.

I am obliged and thankful to all other faculty members and staff members of the Computer Science and Engineering Department, IIT(BHU), Varanasi. I would like to express my gratitude to the Research Program Evaluation Committee members Dr. Ravi Shankar Singh and Dr. Lavanya Selvaganesh for their insightful comments, encouragement, and suggestions, which helped me in improving my research from various perspectives. I would like to convey my sincere gratitude to the other faculty members of the Department of Computer Science and Engineering, Prof. Sanjay Kumar Singh, Prof. Kaushal Kumar Shukla, Prof. Anil Kumar Tripathi, Prof. Rajeev Shrivastava, Dr. Anil Kumar Singh, Dr. Bhaskar Biswas, Dr. Sukomal Pal, Dr. Bidyut Kumar Patra, Dr. Pratik Chattopadhyay, Dr. Hari Prabhat Gupta, Dr. Amrita Chaturvedi, Dr. Tanima Dutta, Dr. Lakshamanan Keilasham, Dr. Ajay Pratap , Dr. Prasenjit Chanak, Dr. Mayank Swarnkar, Dr. Obbattu Sai Lakshmi Bhavana, Dr. Vignesh Sivaraman and Dr. Vinayak Shrivastava for their guidance and support throughout this tenure. When I look beyond the work environment, I would not be who I am now without the help of many senior people. I want to express my deepest gratitude to a few of my seniors, Dr. Shivang Agarwal, Dr. Shushant Kumar Pandey, for their kind acts and moral support. Besides, my colleagues Mrs. Deepika Shukla, Dr. Amit Biswas and Mr. Paras Tiwari, whose valuable comments and discussions are also important in accomplishing my thesis. I also want to offer my sincere thanks to my other colleagues Mr. Siba Shankar Sahu, Mr. Rahul Mishra, Ms. Dipty Tripathi, Ms. Shruti Bajpai, Mr. Vishal Srivastava, Mr. Randheer Bagi, Mr. Ankit Jaiswal, Mr. Amit Kumar, Mr. Agrya Halder and Social Projects' club's family members Mr. Yash Jain, Mr. Ravi Kumar, Mr. Shishir Gupta, Mr. Harshit Krishna Chowdhary, Ms. Shrishti Sharma, Mr. Piyush Singh, Mr. Prajwal M, Mr. Aditya Patodia, Mr. Shivam, Ms. Yashika Chaudhary, Mr. Ritwik Ghorui, Ms. Sai Prashritha, Ms. Yashvi Bansal, Ms. Bhavika Guiney, Mr. Kumar Nachiketa who provided me with technical, moral and emotional support during my research program. I feel humbled and grateful to everyone for all the help and support I have received along the way.

Last but not least, I thank my parents and my younger brother and sisters for their constant support, prayer and encouragement, without which this thesis would not have been completed. I am grateful to my other family members who have supported me along the way.

Date: 26/04/2023 Chintoo Kumar

Abstract

The use of E-commmerce has increased due to the affordability and availability of network bandwidth. The user base has grown multi-folds for these online markets, that has led to the data explosion on the web. It is a challenging task to filter this large volumes of data. In this scenario, information filtering systems are pivotal in making content/item search more accessible and efficient to provide personalised recommendations. Recommender systems are information filtering systems that provide items/contents/services from a plethora of choices available on the web by mining users' past preferences and feedback. Recommendation helps to overcome the information overload problem and as there are many options (for a specific requirement) and more will be added at regular intervals. Recommender systems aim to help users by offering a choice from a large set of items.

Man is a social animal. Therefore, majority of his activities involve a group, like watching a movie with friends, planning a tour itinerary with family, etc. In these scenarios, there is a requirement to recommend items to a group based on preferences of users. Now, the research has gained visibility in discovering user groups with similar preferences and generating recommendations for each group. The prevalence of social group activities led to the development of a group recommender system. The task of a group recommender system is to generate a group recommendation vector by integrating individual user's diversified interest. In general, a group activity increases social bonding. In the group recommendation, individual member activities show his/her satisfaction or dissatisfaction score for a particular item, and all members are equally important in the recommendation process. The entire group must be satisfied with the recommended products in group recommendations.

The decision-making process of a group is more complex, as different group members have diverse interests and contribute unevenly to the final decision. The composition of

a group is a vital problem in group recommendation. A majority of the previous works have developed group recommender systems with prior knowledge about group members. However, in many realistic scenarios, this information is generally not available. The existing models auto-detect groups by computing correlations among all individual users and are thus computationally very inefficient. Traditional clustering approaches reduce time complexity but do not solve dimensionality reduction, sparsity problems, etc. So, this thesis aims to provide efficient models for auto-detecting groups in group recommendations to mitigate these issues. Exploring a particular group from the available information is a challenging task. In many cases, the group members who share similar characteristics are unfamiliar with each other. This motivated us to explore a group before generating the recommendation, as a group composition is essential in an automatically identified group in group recommendations.

As group formation is an unsupervised task, it is inevitable to auto-detect groups from the available information where each group member shares some common characteristics. This thesis focuses on presenting efficient models for identifying user groups automatically. Groups are formed using probabilistic hashing techniques, e.g., locality-sensitive hashing and one permutation hashing, dimensionality reduction technique, i.e, autoencoder, etc. Automatically identified groups are also formed by incorporating swarm intelligence in K-means technique; an optimization technique (Gaussian mixture model) to cluster similar users and a community detection-based algorithm (Louvain algorithm). This thesis also studies the performance of benchmark clustering approaches in group formation. We conducted extensive experiments on real-world datasets and found that the proposed models to identify groups in group recommendations maximize consensus among the users in a group.

A clustering approach on a utility matrix mainly considers users' ratings for items. It does not consider meta-data to form a group. The review texts are auxiliary information available in most of the datasets. This thesis also considers the textual similarities among the review texts to auto-detect a group. In a group recommendation scenario, group members will have different tastes and achieving consensus is a non-trivial task. A common set of interested items taken from users' preferences is the most vital information to provide recommendations. In reality, recommendations are generated by considering only the set of common interest items instead of considering uninterested items. We hypothesise that these uninterested items may influence the overall recommendations by interpreting users' preferences in a better way.

In this thesis, we first discuss some major issues and challenges through a detailed literature survey on existing group recommendations and auto-detecting groups in group recommender systems. Then, we propose effective solutions to those issues. We also include the importance of uninterested items of the group members while generating the recommendation list. We have performed experiments over automatically identified groups and random groups to check the efficacy of the proposed models. We conducted experiments on publicly available real-world datasets. We found that incorporating uninterested items in group recommendations play a positive role in group recommendations while considering order and flexibility in user preferences. We observed that the overall group satisfaction scores are better in an automatically identified group.

Contents

List of Tables

List of Figures