Evaluating Collaborative Filtering Over Time

Neal Lathia

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

Department of Computer Science
University College London

September 9, 2010

To my parents

and their fervent passion for education

Abstract

Recommender systems have become essential tools for users to navigate the plethora of content in the online world. Collaborative filtering—a broad term referring to the use of a variety, or combination, of machine learning algorithms operating on user ratings—lies at the heart of recommender systems' success. These algorithms have been traditionally studied from the point of view of how well they can *predict* users' ratings and how *precisely* they rank content; state of the art approaches are continuously improved in these respects. However, a rift has grown between how filtering algorithms are investigated and how they will operate when deployed in real systems. Deployed systems will continuously be queried for personalised recommendations; in practice, this implies that system administrators will iteratively retrain their algorithms in order to include the latest ratings. Collaborative filtering research does not take this into account: algorithms are improved and compared to each other from a *static* viewpoint, while they will be ultimately deployed in a *dynamic* setting. Given this scenario, two new problems emerge: current filtering algorithms are neither (a) designed nor (b) evaluated as algorithms that must account for time. This thesis addresses the divergence between research and practice by examining how collaborative filtering algorithms behave over time. Our contributions include:

1. A fine grained **analysis** of temporal changes in rating data and user/item similarity graphs that clearly demonstrates how recommender system data is dynamic and constantly changing.
2. A **novel methodology** and time-based **metrics** for evaluating collaborative filtering over time, both in terms of accuracy and the diversity of top-N recommendations.
3. A set of **hybrid algorithms** that improve collaborative filtering in a range of different scenarios. These include temporal-switching algorithms that aim to promote either accuracy or diversity; parameter update methods to improve temporal accuracy; and re-ranking a subset of users' recommendations in order to increase diversity.
4. A set of **temporal monitors** that secure collaborative filtering from a wide range of different temporal attacks by flagging anomalous rating patterns.

We have implemented and extensively evaluated the above using large-scale sets of user ratings; we further discuss how this novel methodology provides insight into dimensions of recommender systems that were previously unexplored. We conclude that investigating collaborative filtering from a temporal perspective is not only more suitable to the context in which recommender systems are deployed, but also opens a number of future research opportunities.

Acknowledgements

Over the past years, I have been very lucky: I have been surrounded by brilliant, intelligent and inspiring people. They contributed to this thesis with their questions, insights, encouragement, and support; I am much indebted to them all. I will never be able to thank my supervisors, Steve Hailes and Licia Capra, enough: being mentored by researchers of this calibre was often all the motivation I needed. Thanks to Cecilia Mascolo, who was the first to suggest that I apply for a PhD (would I be writing this had it not been for that suggestion?); Daniele Quercia, with his unrivalled and contagious passion for research (and blogging); and all of the members of the MobiSys group. Thanks to EPSRC Utiforo, for the financial support, and thanks to all the project partners for the colourful meetings. A special thanks to Torsten Ackemann: the experiments I ran over the past few years would still be running had it not been for his invaluable help with the department's Condor cluster.

A highlight of the recent years is the time I spent in Telefonica I+D's Multimedia Group in Barcelona. A big thanks to Xavier Amatriain, Josep M. Pujol and Jon Froehlich; I not only learned a lot during these summer months, but made some great friends and thoroughly enjoyed my time there. I hope to one day finally manage to go hiking with Jon.

While all those with whom I worked with deserve my utmost thanks, I am even more indebted to my family and friends, who were there to take my mind off of my PhD. Thanks to Paul, Usha, Fergal and Preeya; to Pavle and Justin (we await your return to London), and Viktor (who always turned up at my doorstep at the right time). Thanks to my sisters, Sheila and Anna (who has put up with living with me). A special thanks to Yasmin, who has always been there for me. Lastly, thanks to the bands I have been a part of over these years (The Hartes; Pavle, and The Jukebox Leans; Sean and Duncan), for allowing me to keep nurturing my love for music.

This thesis is dedicated to my parents.

Contents

List of Figures

List of Tables