

# Optimization-based User Group Management: Discovery, Analysis, Recommendation

— PhD Thesis by Behrooz Omidvar-Tehrani

User data is becoming increasingly available in multiple domains ranging from phone usage traces to data on the social Web. User data is a special type of data that is described by user demographics (e.g., age, gender, occupation, etc.) and user activities (e.g., rating, voting, watching a movie, etc.) The analysis of user data is appealing to scientists who work on population studies, online marketing, recommendations, and large-scale data analytics. However, analysis tools for user data is still lacking.

In this thesis [1], we believe there exists a unique opportunity to analyze user data in the form of **user groups**. This is in contrast with individual user analysis and also statistical analysis on the whole population. A group is defined as set of users whose members have either common demographics or common activities. Group-level analysis reduces the amount of sparsity and noise in data and leads to new insights. In this thesis, we propose a user group management framework consisting of following components: user group **discovery, analysis** and **recommendation**.

— The very first step in our framework is group discovery, i.e., given raw user data, obtain user groups by optimizing one or more quality dimensions. For more details of our first component, refer to [2].

— The second component (i.e., analysis) is necessary to tackle the problem of information overload: the output of a user group discovery step often contains millions of user groups. It is a tedious task for an analyst to skim over all produced groups. Thus we need analysis tools to provide valuable insights in this huge space of user groups. For more details about our second component, refer to [3,4].

— The final question in the framework is how to use the found groups. In this thesis, we investigate one of these applications, i.e., user group recommendation, by considering affinities between group members. For more details on this component, refer to [5].

All our contributions of the proposed framework are evaluated using an extensive set of experiments both for quality and performance. The datasets we have employed are MovieLens, BookCrossing, Nokia, DBLP and Facebook.

[1] Behrooz Omidvar-Tehrani: *Optimization-based User Group Management: Discovery, Analysis, Recommendation* (PhD Thesis), University of Grenoble, 2015

[2] Behrooz Omidvar-Tehrani, Sihem Amer-Yahia, Pierre-Francois Dutot and Denis Trystram: *Multi-Objective User Group Discovery on the Social Web*, PKDD 2016

[3] Behrooz Omidvar-Tehrani, Sihem Amer-Yahia, Alexandre Termier: *Interactive User Group Analysis*, CIKM 2015

[4] Behrooz Omidvar-Tehrani, Sihem Amer-Yahia, Alexandre Termier, Aurelie Bertaux, Eric Gaussier and Marie-Christine Rousset: *Towards a Framework for Semantic Exploration of Frequent Patterns*, IMMOA 2013

[5] Sihem Amer-Yahia, Behrooz Omidvar-Tehrani, Senjuti Basu Roy and Nafiseh Shabib: *Group Recommendation with Temporal Affinities*, EDBT 2015