

# 面向个性化推荐的两层混合图模型<sup>①</sup>

张少中<sup>1,2</sup> 陈德人<sup>1</sup> (1.浙江大学软件学院 浙江 宁波 315100;

2.江万里学院 浙江 宁波 315100)

## Hybrid Graph Model with Two Layers for Personalized Recommendation

Shaozhong Zhang<sup>1,2</sup>, Deren Chen<sup>1</sup> (1. Software College, Zhejiang University, Ningbo 315100;

2. Zhejiang WanLi University, Ningbo 315100)

**Abstract:** A hybrid graph model for personalized recommendation, which is based on small world network and Bayesian network, is presented. The hybrid graph model has two-layers. The bottom level means user's layer and the upper one means merchandise's layer. The user's layer is an undirected arcs graph, which describes the relation of the user's nodes by small world network. The undirected arcs inside the connected nodes of user's layer mean the similarity of the preference of users. These arcs are weighted by relational strength. The weight represents node's similarity or link's strength and intensity. Nodes in the same group are more similar to each other or more strongly connected. Users in a same group have the same or similar trendy of preferences. The merchandise's layer describes the relation of goods or produce to others. It is connected by directed links, which means an implicated definition among merchandises, a user that purchase certain merchandise also tends to purchase another. The properties and content of merchandise can be used to show the similarity of the merchandise. The relations between user's layer and merchandise's layer are connected by directed links. The start node of the directed links is a user node in user's layer belonging to some node group, which is gained by small world network. The end node of links is the node of some merchandise of the merchandise's layer. The directed links between the user's layer and the merchandise's layer are connected based on trade information of users. The strength of the relation between users and merchandises can be denoted by the probability parameter. The probability parameter shows a possibility of some users selecting for some merchandises.

Firstly, algorithms for users clustering and for analysis of new user interest are presented to construct a hybrid graph model. Two important characteristic parameters, which are in small-world network, are introduced. These are characteristic path length and clustering coefficient. New user interest analysis is to judge which clustering group is the best match by calculating the distance of the new user node to the others user nodes.

Secondly, Bayesian network for causality of merchandises and users is constructed. It can be divided two parts, structure learning and parameter learning. The paper adopts the maximal mutual information principle to restrict complexity based on degree of Bayesian network. A new maximal mutual information entropy score function with restriction is defined and a maximum likelihood estimate algorithm is used to calculated parameter.

Thirdly, recommending algorithm for new user is presented. In the algorithm, the initialized inputs can utilize some users information including the attributes and browsing process of a user. A proper user-clustering group will be gained by clustering matching with other users in small world network based on this information. Then all the other users nodes, which connect to this user, are selected based on a threshold of path length in the clustering. The recommended merchandise set of these users will be obtained by Bayesian network inference using these nodes as proofs. Finally, a set of recommendation of merchandise is presented for user according to their order of probability distribution.

The paper uses the mean absolute error to evaluate the model and MovieLens database is selected. The experimentation shows that the model be accomplished to represent the relationships from user to user, merchandise to merchandise, and user to merchandise. The result shows that the hybrid graph model has a good performance in personalized recommendation.

<sup>①</sup> Supported by the National Natural Science Foundation of China under Grant No.70671007; China Postdoctoral Science Foundation under Grant No.20060390391