Solve recommendation problem by stochastic block modeling and Message-passing

Recommendation problem is one kind of information filtering problem that asks to predict the ‘rating’ or ‘preference’ that user would give to an item. For an example, given database of users rating movies, how to predict the non-existing rating between a user and a movie. Doing prediction based on from user’s past behavior is also called ‘collaborative filtering’.

Recommendation systems which give prediction has been widely used in big companies like Amazon, Facebook, Netflix, Ebay... And different recommendation algorithms have been studied widely in both academia and industry.

My proposal is to use stochastic block model which has been widely used in community detection problem to do recommendation, and solve the stochastic block model by scalable message passing algorithms. There is already one study [PLOS One, 7, 9(2012)] uses the similar idea that shows the power of stochastic block modeling in recommendation systems. However, their method of inferring the model is based on Monte-Carlo simulation, which is slow and not scalable. Thus their method works only on small systems. I propose to use Belief Propagation algorithm to do the model inference and Expectation-Maximization algorithm to learn parameters of the model. Belief Propagation is a scalable message passing algorithm that has computational complexity almost linear to number of edges of the system.

I already have all the necessary theoretical studies e.g. BP equations for the problem. It would be nice if one REU can implement the algorithm by C++ and do some experimental studies on large dataset and benchmarks.

I expect that from this project, REUs can learn something in the field of machine learning, recommendation system and message passing algorithms, and practice scientific programming using C++.

Pan Zhang
03/31/2014