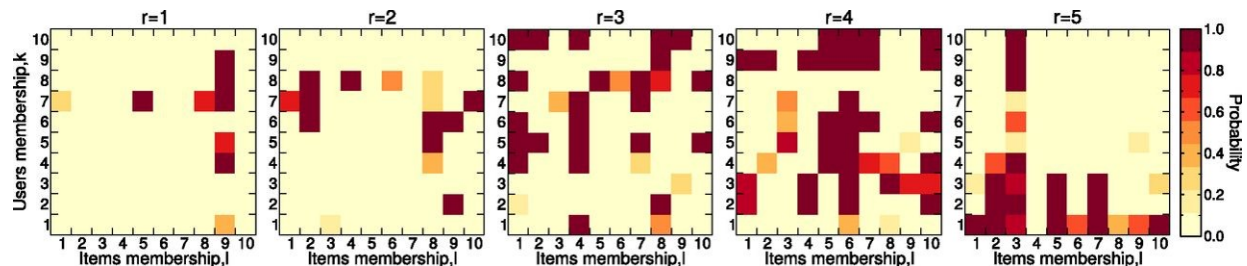# Suggestions for you: A better, faster recommendation algorithm

December 1 2016, by John German

Probability matrices in MMSBM. Credit: Antonia Godoy-Lorite et al. *PNAS* doi:10.1073/pnas.1606316113

The internet is rife with recommendation systems, suggesting movies you should watch or people you should date. These systems are tuned to match people with items, based on the assumption that similar people buy similar things and have similar preferences. In other words, an algorithm predicts which items you will like based only on your, and the item's, previous ratings.

But many existing approaches to making recommendations are simplistic, says physicist and computer scientist Cristopher Moore, a Santa Fe Institute professor. Mathematically, these methods often assume people belong to single groups, and that each one group of people prefers a single group of items. For example, an algorithm might suggest a science fiction movie to someone who had previously enjoyed

another different science fiction movie— - even if the movies have nothing else in common.

"It's not as if every movie belongs to a single genre, or each viewer is only interested in a single genre," says Moore. "In the real world, each person has a unique mix of interests, and each item appeals to a unique mix of people."

In a new paper in the *Proceedings of the National Academy of Sciences*, Moore and his collaborators introduce a new recommendation system that differs from existing models in two major ways. First, it allows individuals and items to belong to mixtures of multiple overlapping groups. Second, it doesn't assume that ratings are a simple function of similarity ;—instead, it predicts probability distributions of ratings based on the groups to which the person or item belongs.

This flexibility makes the new model more realistic than existing models that posit a linear relationship between users and items, says Moore. Not everyone enjoys rating things, and not everyone uses ratings in the same way—if a person rates a movie 5 instead of 1, that doesn't mean she likes it five times as much. The new model can learn nonlinear relationships between users and ratings over time.

Moore and his collaborators tested their model on five large datasets, including recommendations systems for songs, movies, and romantic partners. In each case, the new model's predicted ratings proved more accurate than those from existing systems—and their algorithm is faster than competing methods as well.

Moore is motivated by the opportunity to explore rich data sets and networks, where nodes and links have locations, content, and costs. "Our algorithm is powerful because it is mathematically clear," he says. "That makes it a valuable part of the portfolio of methods engineers can use."

"Now if we can just get people to read news they ought to, instead of what they like," Moore says. "But that's a much harder problem."

The paper, "Accurate and scalable social recommendation using mixed-membership stochastic block models," was published November 23 in *PNAS*, co-authored by Chris Moore of the Santa Fe Institute and Antonia Godoy-Lorite, Roger Guimerà, and Marta Sales-Pardo, all of the Universitat Rovira i Virgili, Spain.

Provided by Santa Fe Institute