

Improving recommendation system algorithms

July 8 2011, by Larry Hardesty



Graphic: Christine Daniloff

Recommendation algorithms are a vital part of today's Web, the basis of the targeted advertisements that account for most commercial sites' revenues and of services such as Pandora, the Internet radio site that tailors song selections to listeners' declared preferences. The DVD rental site Netflix deemed its recommendation algorithms important enough that it offered a million-dollar prize to anyone who could improve their predictions by 10 percent.

But Devavrat Shah, the Jamieson Career Development Associate

Professor of Electrical Engineering and Computer Science in MIT's Laboratory of Information and Decision Systems, thinks that the most common approach to recommendation systems is fundamentally flawed. Shah believes that, instead of asking users to rate products on, say, a five-star scale, as Netflix and Amazon do, recommendation systems should ask users to compare products in pairs. Stitching the pairwise rankings into a master list, Shah argues, will offer a more accurate representation of consumers' preferences.

In a series of papers ([paper 1](#) | [paper 2](#) | [paper 3](#)) published over the last few years, Shah, his students Ammar Ammar and Srikanth Jagabathula, and Vivek Farias, an associate professor at the MIT Sloan School of Management, have demonstrated algorithms that put that theory into practice. Besides showing how the algorithms can tailor product recommendations to customers, they've also built a [website](#) that uses the algorithms to help large groups make collective decisions. And at an Institute for Operations Research and Management Sciences conference in June, they presented a version of their algorithm that had been tested on detailed data about car sales collected over the span of a year by auto dealers around the country. Their algorithm predicted car buyers' preferences with 20 percent greater accuracy than existing algorithms.

Calibration conundrum

One of the problems with basing recommendations on ratings, Shah explains, is that an individual's rating scale will tend to fluctuate. "If my mood is bad today, I might give four stars, but tomorrow I'd give five stars," he says. "But if you ask me to compare two movies, most likely I will remain true to that for a while."

Similarly, ratings scales may vary between people. "Your three stars might be my five stars, or vice versa," Shah says. "For that reason, I strongly believe that comparison is the right way to capture this."

Moreover, Shah explains, anyone who walks into a store and selects one product from among the three displayed on a shelf is making an implicit comparison. So in many contexts, comparison data is actually easier to come by than ratings.

Shah believes that the advantages of using comparison as the basis for recommendation systems are obvious but that the computational complexity of the approach has prevented its wide adoption. The results of thousands — or millions — of pairwise comparisons could, of course, be contradictory: Some people may like *Citizen Kane* better than *The Godfather*, but others may like *The Godfather* better than *Citizen Kane*. The only sensible way to interpret conflicting comparisons is statistically. But there are more than three million ways to order a ranking of only 10 movies, and every one of them may have some probability, no matter how slight, of representing the ideal ordering of at least one ranker. Increase the number of movies to 20, and there are more ways to order the list than there are atoms in the universe.

Ordering out

So Shah and his colleagues make some assumptions that drastically reduce the number of possible orderings they have to consider. The first is simply to throw out the outliers. For example, Netflix's movie-rental data assigns the Robin Williams vehicle *Patch Adams* the worst reviews, on average, of any film with a statistically significant number of ratings. So the MIT algorithm would simply disregard all the possible orderings in which *Patch Adams* ranked highly.

Even with the outliers eliminated, however, a large number of plausible orderings might remain. From that group, the MIT algorithm selects a subset: the smallest group of orderings that fit the available data. This approach can winnow an astronomically large number of orderings down to one that's within the computational purview of a modern computer.

Finally, when the [algorithm](#) has arrived at a reduced number of orderings, it uses a movie's rank in each of the orderings, combined with the probability of that ordering, to assign the movie an overall score. Those scores determine the final ordering.

Paat Rusmevichientong, an associate professor of information and operations management at the University of Southern California, thinks that the most interesting aspect of Shah's work is the alternative it provides to so-called parametric models, which are more restrictive. These, he says, were "the state of the art up until 2008, when Professor Shah's paper first came out."

"They've really, substantially enlarged the class of choice models that you can work with," Rusmevichientong says. "Before, people never thought that it was possible to have rich, complex choice models like this."

The next step, Rusmevichientong says, is to test that type of model selection against real-world data. The analysis of car sales is an early example of that kind of testing, and the MIT researchers are currently working up a version of their conference paper for journal publication. "I've been waiting to see the paper," Rusmevichientong says. "That sounds really exciting."

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Improving recommendation system algorithms (2011, July 8) retrieved 27 April 2024

from <https://phys.org/news/2011-07-algorithms.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.