"Breakthrough" algorithm exponentially faster than any previous one

Smarter, faster algorithm cuts number of steps to solve problems

New algorithm uses adaptive sampling to exponentially speed up computation

What if a large class of algorithms used today — from the algorithms that help us avoid traffic to the algorithms that identify new drug molecules — worked exponentially faster?

Computer scientists at the [Harvard John A. Paulson School of Engineering and Applied Sciences](#) (SEAS) have developed a completely new kind of algorithm, one that exponentially speeds up computation by dramatically reducing the number of parallel steps required to reach a solution.

The theoretical research was presented at the [ACM Symposium on Theory of Computing](#) (STOC), June 25-29 and the experimental research will be presented at the [International Conference on Machine Learning](#) (ICML), July 10 -15.

A lot of so-called optimization problems, problems that find the best solution from all possible solutions, such as mapping the fastest route from point A to point B, rely on sequential algorithms that haven't changed since they were first described in the 1970s. These algorithms solve a problem by following a sequential step-by-step process. The number of steps is proportional to the size of the data. But this has led to a computational bottleneck, resulting in lines of questions and areas of research that are just too computationally expensive to explore.

"These optimization problems have a diminishing returns property," said [Yaron Singer](#), Assistant Professor of Computer Science at SEAS and senior author of the research. "As an algorithm progresses, its relative gain from each step becomes smaller and smaller."

Yaron Singer, Assistant Professor of Computer Science at SEAS, designed a new type of algorithm. (Photo courtesy of Eliza Grinnell/Harvard SEAS)

Singer and his colleague asked: what if, instead of taking hundreds or thousands of small steps to reach a solution, an algorithm could take just a few leaps?

"This algorithm and general approach allows us to dramatically speed up computation for an enormously large class of problems across many different fields, including computer vision, information retrieval, network analysis, computational biology, auction design, and many others," said Singer. "We can now perform computations in just a few seconds that would have previously taken weeks or months."

"This new algorithmic work, and the corresponding analysis, opens the doors to new large-scale parallelization strategies that have much larger speedups than what has ever been possible before," said Jeff Bilmes, Professor in the Department of Electrical Engineering at the University of Washington, who was not involved in the research. "These abilities will, for example, enable real-world summarization processes to be developed at unprecedented scale."

Traditionally, algorithms for optimization problems narrow down the search space for the best solution one step at a time. In contrast, this new algorithm samples a variety of directions in parallel. Based on that sample, the algorithm discards low-value directions from its search space and chooses the most valuable directions to progress towards a solution.

Take this toy example:

You're in the mood to watch a movie similar to The Avengers. A traditional recommendation algorithm would sequentially add a single movie in every step which has similar attributes to those of The Avengers. In contrast, the new algorithm samples a group of movies at random, discarding those that are too dissimilar to The Avengers. What's left is a batch of movies that are diverse (after all, you don't want ten Batman movies) but similar to The Avengers. The algorithm continues to add batches in every step until it has enough movies to recommend.

This process of adaptive sampling is key to the algorithm's ability to make the right decision at each step.

"Traditional algorithms for this class of problem greedily add data to the solution while considering the entire dataset at every step," said Eric Balkanski, a graduate student at SEAS and co-author of the research. "The strength of our algorithm is that in addition to adding data, it also selectively prunes data that will be ignored in future steps."
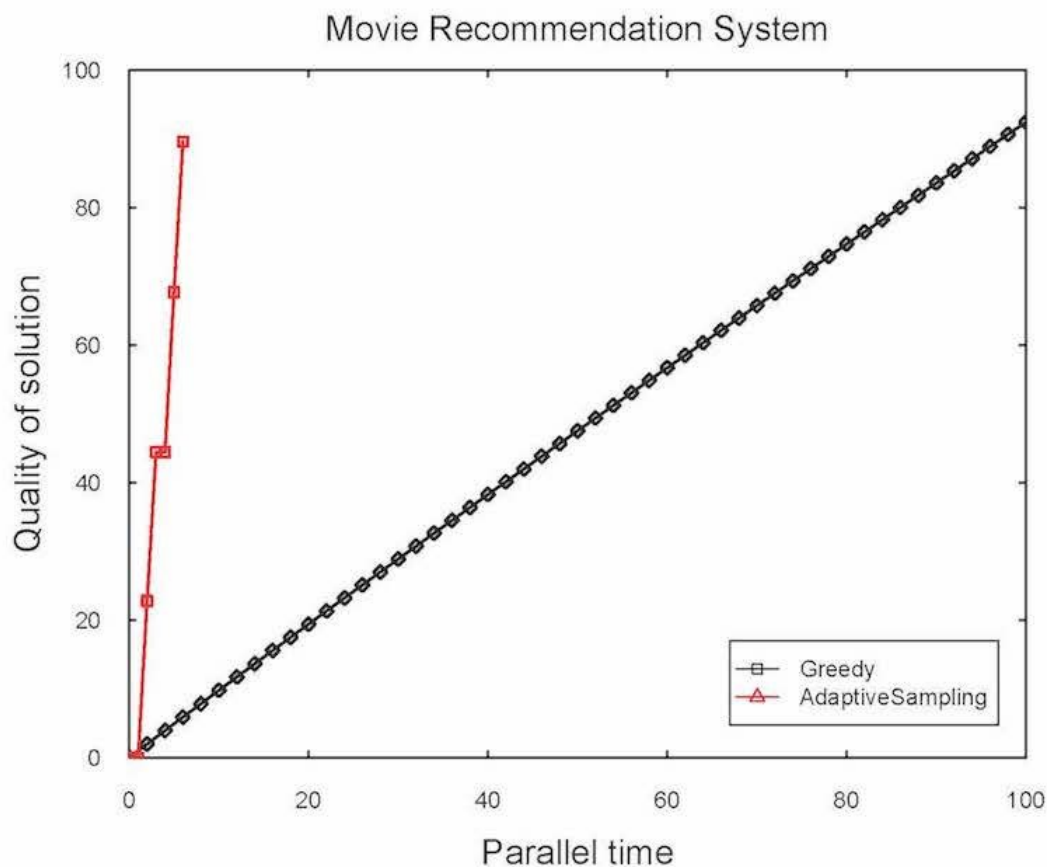
The black line shows the number of steps a traditional algorithm takes to solve a problem while the red line demonstrates the number of steps the new algorithm needs.

In experiments, Singer and Balkanski demonstrated that their algorithm could sift through a data set which contained 1 million ratings from 6,000 users on 4,000 movies and recommend a personalized and diverse collection of movies for an individual user 20 times faster than the state-of-the-art.

The researchers also tested the algorithm on a taxi dispatch problem, where there are a certain number of taxis and the goal is to pick the best locations to cover the maximum number of potential customers. Using a dataset of two million taxi trips from the New York City taxi and limousine commission, the adaptive-sampling algorithm found solutions 6 times faster.

"This gap would increase even more significantly on larger scale applications, such as clustering biological data, sponsored search auctions, or social media analytics," said Balkanski.

Of course, the algorithm's potential extends far beyond movie recommendations and taxi dispatch optimizations. It could be applied to:

- designing clinical trials for drugs to treat Alzheimer's, multiple sclerosis, obesity, diabetes, hepatitis C, HIV and more
- evolutionary biology to find good representative subsets of different collections of genes from large datasets of genes from different species
- designing sensor arrays for medical imaging
- identifying drug-drug interaction detection from online health forums

This process of active learning is key to the algorithm's ability to make the right decision at each step and solves the problem of diminishing returns.

Illustration: finding a **diverse summary** of 100 images
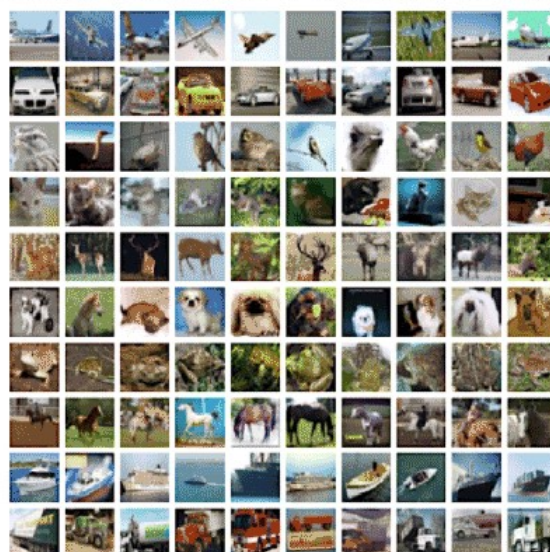
"This research is a real breakthrough for large-scale discrete optimization," said Andreas Krause, professor of Computer Science at ETH Zurich, who was not involved in the research. "One of the biggest challenges in machine learning is finding good, representative subsets of data from large collections of images or videos to train machine learning models. This research could identify those subsets quickly and have substantial practical impact on these large-scale data summarization problems."

Singer-Balkanski model and variants of the algorithm developed in the paper could also be used to more quickly assess the accuracy of a machine learning model, said Vahab Mirrokni, a principal scientist at Google Research, who was not involved in the research.

"In some cases, we have a black-box access to the model accuracy function which is time-consuming to compute," said Mirrokni. "At the same time, computing model accuracy for many feature settings can be done in parallel. This adaptive optimization framework is a great model for these important settings and the insights from the algorithmic techniques developed in this framework can have deep impact in this important area of machine learning research."

Singer and Balkanski are continuing to work with practitioners on implementing the algorithm.