

Making crowdsourcing easier

August 24 2012, by Larry Hardesty, MIT News Office



Graphic: Christine Daniloff

Crowdsourcing is a technique for farming out labor-intensive tasks over the Internet by splitting them into small chunks that dozens, hundreds or even thousands of people complete at their desks for a few cents each.

Researchers at MIT's Computer Science and Artificial Intelligence Laboratory are developing a new [database system](#), called Qurk, that will automatically crowdsource tasks that are difficult or impossible to perform computationally. Images stored in a standard database system,

for example, could be sorted according to date of creation or some other data tag, whether applied automatically or by hand. Images in a Qurk database, however, could be sorted according to the approximate age of the people depicted, or the appeal of the depicted locations as travel destinations, or any other attribute whose assessment would require human judgment.

In a pair of conference papers last year, the researchers described and demonstrated Qurk's general [computational framework](#). In a new paper they're presenting this month at the 38th International Conference on Very Large Databases, they get into the nitty-gritty, describing a series of experiments on how best to crowdsource the common database operations "sort" and "join." The researchers found that, using the most obvious implementation of the join operation, it cost \$67 to combine two sets of images through Amazon's Mechanical Turk crowdsourcing service. With an improved implementation that they arrived at experimentally, they could get the cost down to \$3.

"When you write database queries in a declarative style using a language like SQL," says Rob Miller, an associate professor of computer science and one of the authors of the Qurk papers, "the database system can optimize them: It can find the fastest way, or the way that's cheapest in resources—whatever those resources are—to do it. You didn't specify all the details about how it should be done. The system figures all that out."

In the same way, says Adam Marcus—who leads Qurk's development with fellow graduate student Eugene Wu—Qurk is intended to spare the user from specifying the details of how to crowdsource database operations. "You can just say, 'I have this collection of images, and I want to sort them by how cute they are,' and the system will actually figure out how to implement a sort over your data set," Marcus says.

[Crowdsourcing](#) is particularly useful for tasks that are trivial for humans

but difficult, if not impossible, for computers. The paradigmatic such task is image recognition: Even the most complex, time-consuming image-recognition algorithms can't identify objects in images nearly as consistently as people can. So in their experiments, the researchers concentrated on databases of images. On all three Qurk papers, Marcus, Wu and Miller have been joined by professors Sam Madden and David Karger, both of the Department of Electrical Engineering and Computer Science.

If you were going to use Mechanical Turk to sort images according to, say, how cute they were, the most obvious approach would be to ask recruits—"Turkers" as they're generally known—to compare two images at a time and rank them; algorithms could then stitch the pairwise rankings into a master list. One of the things the MIT researchers were investigating was how many images a Turker could be expected to rank at once and still provide useful data, and their conclusion was somewhere between five and 10, depending on the task.

They were also comparing ranking schemes with rating schemes, in which Turkers would assign each image a rating of one to five stars. Ratings systems, which are favored by websites such as Amazon and Netflix, have notorious drawbacks, but the MIT researchers developed an interface that depicted, in addition to a large version of the image to be rated, a row of 10 smaller images randomly drawn from the database. A Turker rating his or her first image would thus have some sense of the average cuteness of the images in the database and could calibrate the ratings scale accordingly.

The researchers found that, while ranking provided more accurate sorting, the calibrated ratings system fared surprisingly well and was much cheaper. Depending on the sorting task, the importance of perfect accuracy and the user's budget, Qurk could thus use ranking or rating or some hybrid of the two, in which rating provided an initial ordering that

the more expensive ranking then refined.

Besides the sort operation, the researchers also tested crowdsourced implementations of the join operation, which merges data sets containing complementary information. One data set, for instance, might contain entries that refer to companies by their names, while the other refers to them by stock ticker symbols, and the idea is to combine the records that refer to the same entities. In the new experiments, the Turkers were joining data sets that contained some images of the same people or objects.

Again, the researchers found that Turkers could accurately process images in small batches, usually three or four pairs of images at a time. The interface was designed so that a few images from each data set were displayed in columns; the Turker then had to draw a line connecting any two images of the same person or object.

To further improve the efficiency of the join algorithm, the researchers also recruited Turkers to "pre-filter" the data. Some Turkers, that is, would be shown a few images at a time and asked to determine, say, the hair color of the subjects, or whether they had beards or wore glasses. The system could then use that information to select the images that other Turkers would compare.

Comparing images in small batches rather than in pairs made the crowdsourced join operation seven times as efficient; pre-filtering added another threefold increase, getting the cost of a merging two databases down from \$67 to \$3.

Of course, in order for Qurk to pre-filter data, it would need to know what attributes to ask Turkers to look for; identifying subjects with beards would be a useless request if the task were to join databases containing pictures of cars. But, Marcus says, Qurk will allow the user to

specify a group of attributes that might be useful for pre-filtering. The system will then evaluate those attributes on the fly, determining which, if any, actually increase the efficiency of the join operation. Similarly, with batch sorting tasks, Qurk will try out different-sized batches—within the five- to 10-image range—and see which yields the best results.

Michael Franklin, a professor of [computer science](#) and the University of California, Berkeley, is also researching ways to crowdsource database operations. "It's a really exciting area," he says. "There's lots of cases where those hybrid systems already are and are going to be increasingly important."

Franklin's group is developing a system that crowdsources operations written in a slightly extended version of SQL; by contrast, the MIT researchers "came up with a pluggable template language," Franklin says. Since Franklin's system, CrowdDB, uses the familiar SQL language, "I think ours is easier to use," he says. "But theirs probably covers more cases. It probably gives the programmer more options and more control."

Ultimately, however, both groups are addressing the same basic research question. "If you view people as the processing units, how do you let them know what they're doing?" Franklin says. "The programming interface for people is just fundamentally different than it is for computers."

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: Making crowdsourcing easier (2012, August 24) retrieved 24 April 2024 from <https://phys.org/news/2012-08-crowdsourcing-easier.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.