The power of 'random': 'Seemingly loopy' technique could dramatically improve communications networks

February 9 2010, by Larry Hardesty

Graphic: Christine Daniloff

A radical new approach to the design of communications networks, called "network coding," promises to make Internet file sharing faster, streaming video more reliable, and cell-phone reception better -- among other improvements.

MIT is in the thick of these new developments. Last year, MIT researchers shared in two awards from IEEE, formerly the Institute of Electrical and Electronics Engineers, for papers that made vital contributions to the field of network coding.

“Most networks right now are built roughly along the same principles as a transportation network, or any other network that’s trying to deliver tangible goods,” says Muriel Médard, a professor in the Research Laboratory of Electronics who was a coauthor on both papers. A packet of data traveling across the Internet, for instance, passes through a series of devices called routers before it reaches its destination. A router doesn’t tamper with the packet’s contents; it just sends it on to the next router.

This is the first article in a two-part series on MIT contributions to the fledgling field of network coding. (part two is available here).

With network coding, however, a router doesn’t just hand off the packets it receives; it mathematically combines them into new, hybrid packets. If the combination is done cleverly enough, this makes the whole network more efficient.

To see how this might work, suppose that we’re at a coffee shop with our laptops. I’m trying to send you a message over the coffee shop’s WiFi connection at the same time you’re trying to send me a message. Ordinarily, my message will travel to the coffee shop’s wireless router, and then the router will send it to you. Your message will travel to the router, and then the router will send it to me. That’s four total transmissions. But if, instead of forwarding our messages, the router combines them and broadcasts the combination, there are only three total transmissions. Since you have a copy of the message you sent me, you can subtract it from the combination, and I can do the same with the message I sent you. If our laptops and the router do a little extra processing, they reduce the system’s bandwidth consumption by 25 percent.

Of course, this example assumes that the receivers already have the data they need to decode the combination, which is rarely the case in the real world. And data traveling over a network usually pass through a number of routers: if each of those routers is recombining packets that are already combinations themselves, the decoding process becomes much more complicated. But in principle, there's a way to get it all to work.

Cracked code

Network coding was born around 1999 or 2000, in a couple of papers that suggested that combining data at routers could improve network efficiency. But how that combination should be done, and what kinds of efficiency gains were possible, were unclear.

Then, in 2003, Médard, her grad student Tracey Ho (who's now at Caltech), MIT professor of electrical engineering David Karger, and colleagues at the University of Illinois and Caltech proved a counterintuitive result: in many cases, the best way to combine data at a router is to do it randomly.

Today, cell phones and computers send messages digitally: every message is a sequence of 0s and 1s. But any sequence of 0s and 1s can be thought of as a single number. With random network coding, a router receives, say, three messages, multiplies each of them by a different, randomly selected number, and adds the results together. That final sum is the new, hybrid message. The router sends the hybrid on to the next router in the network — but it also includes information about the three random numbers it used to produce the hybrid.

Random coding yields the biggest gains in networks where connections are spotty, but where there are several possible routes between sender and receiver. Suppose, for instance, that you're in a densely populated city with good cell-phone coverage. You're within range of several different cell towers, but you're inside a building that's interfering with your transmissions. Your cell phone is sending out lots of packets of data, but there's not one nearby cell tower that's receiving all of them. If each tower simply "hybridizes" the packets it receives and sends them on, then as long as the recipient gets enough hybrids from enough different towers, it can reconstruct the original message.

Ho, Médard, and their colleagues proved mathematically that if the same group of messages was sent to several different receivers, random coding made the most efficient possible use of the network's bandwidth.

"The idea is seemingly loopy," Médard says. "I think it's fair to say that it was greeted with some amount of bemusement in some parts of the community." As a graduate student, Ho was charged with presenting the findings at a conference in Japan, and her audience was skeptical. "People said, 'You must be comparing it to bad routing,'" Médard says. Under cross-examination from a room full of seasoned researchers, Médard says, "Anyone else would have just curled into a fetal position." But Ho, she says, was "cool as a cucumber. She was also the collegiate pistol champion. So the girl can keep her cool."

The IEEE award last year was an indication that the bemusement had turned to recognition. "It's a theoretical result, but it has deep practical implications," says Chris Ramming, director of the Corporate External Research Office for chip manufacturer Intel. Before coming to Intel, Ramming worked for the U.S. Defense Department's Defense Advanced Research Projects Agency (DARPA). "When I was a program manager at DARPA, we had a big project that used that approach as the core of the implementation techniques," Ramming says. "It was definitely seminal, and people are trying to build on it. So it's hugely important."

The other paper honored by the IEEE last year, and MIT researchers' continuing work on network coding, will be the subject of part two of this series.

This is the first article in a two-part series on MIT contributions to the fledgling field of network coding. (part two is available here).

Provided by Massachusetts Institute of Technology