

## A Breakthrough in Algorithm Design

*Computer scientists at Carnegie Mellon University have devised an algorithm that might be able to solve a certain class of linear systems much more quickly than today's fastest solvers.*

**S**YSTEMS OF LINEAR equations are everywhere. They are used in telecommunications, transportation, manufacturing, and many other domains. The algorithms used to solve linear systems must be able to compute solutions to equations involving millions—or sometimes billions—of variables. Because calculating solutions for these systems is time-consuming on even the fastest computers, finding ways to accelerate these computations is an ongoing challenge for algorithm designers. Now, a group of computer scientists at Carnegie Mellon University (CMU) has devised an algorithm that the researchers say might be able to solve a certain class of linear systems much more quickly than today's fastest solvers.

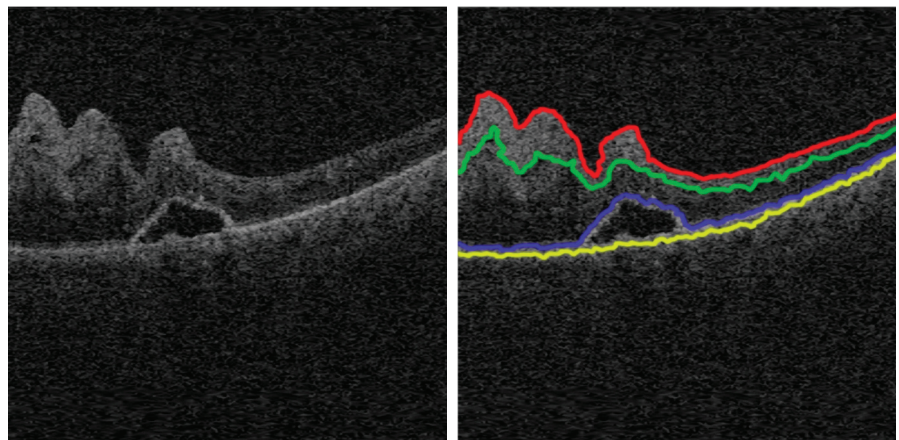
The researchers say the algorithm, which applies to a class of problems known as symmetric and diagonally dominant (SDD) systems, not only has practical potential, but also is so fast it might soon be possible for a desktop PC to solve systems with one billion variables in just seconds. “The main point of the new algorithm is that it is guaranteed to work and to work quickly,” says

Gary L. Miller, a professor of computer science at CMU and a member of the three-person team that developed the new algorithm.

SDD systems, characterized by system matrices in which each diagonal element is larger than the sum of the absolute values of all the other elements in the corresponding row, are used for a wide range of purposes, from online recommendation systems to industrial simulations, materials modeling, and

image processing. Algorithms used for this type of linear system fall into two broad classes: direct solvers, such as Gaussian elimination, and iterative solvers. In contrast to direct solvers, which compute exact solutions, iterative solvers produce a series of approximate solutions. Direct methods are usually memory-hungry, a limitation that makes iterative solvers, such as the kind developed by the CMU team, more effective for the large data sets generated by today's applications.

While iterative solvers eventually return satisfactory results, those results typically take a long time to produce because they require calculating many approximations. There have been hundreds of approaches to developing faster iterative solvers, but one method has proved to be the most effective and has become a guid-



**A linear system designed to improve the quality of retinal image segmentation through the use of an iterative solver technique called spectral rounding, developed at Carnegie Mellon University and the University of Pittsburgh Medical Center. Conventional segmentation algorithms tend to fail in the presence of retinal abnormalities. On the left is the input image. On the right is the segmented image.**

ing principle of sorts in this area of research. The idea is to solve computations on a massive linear system by quickly running computations on a sparser system that in some well-defined algebraic sense is similar to the larger one. The sparser system used to set up these computations for the larger system is called the preconditioner.

Producing the sparse matrix requires zeroing out some of the non-zero entries and increasing the weight of others in the larger matrix. “One key ingredient in the newer algorithms is the judicious use of randomization to determine which entries are zeroed out,” explains Miller, who likens the CMU algorithm’s process of sparsification to “flipping a biased coin” to determine the fate of an element in the system matrix. This sparsification process is designed to create a representation of the larger system matrix to generate the preconditioner that will guide later computations.

While finding a preconditioner might seem straightforward, finding a

good one is not. A reliable method for finding a good preconditioner to accelerate computations on large system matrices is an ongoing challenge in math and computer science. Methods that rely on heuristics, for example, have been effective, but only to a limited extent. “Heuristic solvers are often guided by good intuition,” says Richard Peng, a graduate student in the CMU computer science department and a member of the new algorithm team. “However, the critical missing pieces of understanding make them unreliable, especially with the large and complicated systems that we face today.”

### The Spielman-Teng Solver

The path to developing a more effective method than heuristics for finding a good preconditioner dates to the early 1990s and a series of ideas that suggested viewing SDD systems as combinatorial graphs. Research projects in spectral graph theory and numerical analysis developed these ideas in a string of new theories that, in 2004,

emerged as what was widely considered to be a breakthrough proof by Daniel Spielman and Shang-Hua Teng. Spielman and Teng were able to prove that every SDD matrix has a good and discoverable preconditioner.

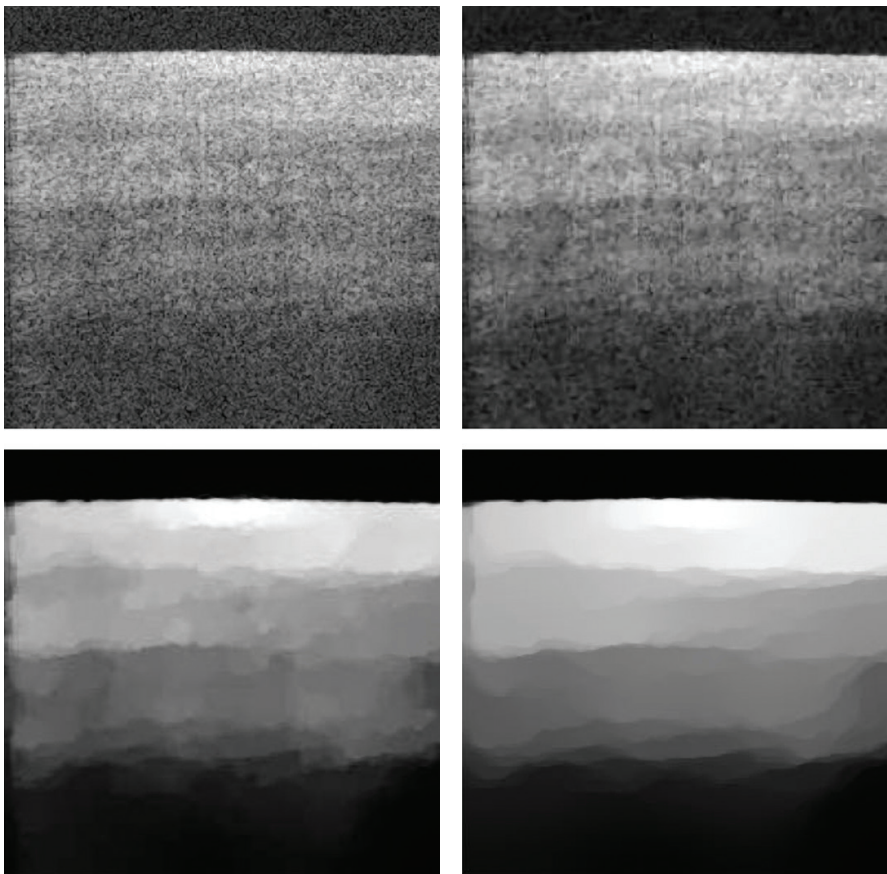
To put the idea into electrical terms, Spielman and Teng showed that for a given electrical network, there will be one that uses fewer resistors while having the same reliability and energy-consumption properties as the original. “The Spielman-Teng solver is asymptotically much faster than everything that was known before,” says Peng. “It is faster than previous solvers for all systems larger than a fixed size, and that difference in speed increases as the system becomes larger.”

Building on Spielman and Teng’s work, the CMU team developed their new algorithm that, from a mathematical point of view, is more concise, taking only five pages to detail instead of 50. “It’s nearly optimal,” says Ioannis Koutis, the third member of the CMU team and now a professor of computer science at the University of Puerto Rico, Rio Piedras. “We know that we can’t do much better, if that’s possible at all.”

Due to its simplicity, along with its promise of significant speed improvements over earlier algorithms, the new solver made headlines when it was introduced last October at the IEEE 51st Annual Symposium on Foundations of Computer Science (FOCS). With an optimized implementation, the researchers say, the algorithm would be some 10 to 20 times faster than other solvers for current problems. (The technical details of the algorithm are in the FOCS paper; see I. Koutis, G.L. Miller, and R. Peng, “Approaching Optimality for Solving SDD Linear Systems.”)

Spielman, a professor of applied mathematics and computer science at Yale University, says the CMU algorithm represents a significant improvement for solving SDD systems. “It is the first algorithm for this problem that is both provably fast in an asymptotic sense and that could be fast in practice,” he says.

Spielman explains that when he and Teng created their initial approach to this problem in 2004, their algorithm was guaranteed to find solutions in near-linear time. However, this guaran-



**An application designed to improve the quality of optical coherence tomography images for an automated cartilage health-assessment routine. The top images represent the input. The bottom images, enhanced with a linear system designed to smooth the optical coherence tomography images, show striations in the cartilage that are indicative of unhealthy tissue.**

**“The main point of the new algorithm is that it is guaranteed to work and to work quickly,” says Gary L. Miller.**

tee was only theoretical and for what he calls “impractically huge” system matrices. The new CMU solver, says Spielman, fixes this problem. “The resulting algorithm is theoretically sound, probably correct, and reasonable in practice,” he says. “Now, they just need to optimize their implementations.”

Spielman says he expects it will be another decade before understanding of the CMU algorithm fully matures, and that it remains to be seen whether the algorithm will be put to use by developers in the near term. “There are still many reasonable ways of varying their algorithm,” Spielman says. “I expect particular applications will benefit from different optimizations.”

For the algorithm to be useful in practice, these optimizations must be done so the new algorithm can accommodate the massive sets of data that are the norm for machine-learning problems, materials modeling, image processing, and other applications whose computational results often benefit from better input-data quality. One way to accommodate large amounts of input data while still achieving acceptable computation speed is, of course, to parallelize, but parallelization remains an ongoing challenge in itself for algorithm designers.

“People want better answers, but the algorithms cannot practically handle the larger data sets unless they are fast,” says Spielman. “The best we can hope for is algorithms whose running time scales linearly with their input size.”

Koutis, Miller, and Peng are working on such optimizations, including parallelization, and are testing their ideas in several practical implementations, such as a new approach to medi-

cal imaging developed with David Tolliver, also at CMU, and researchers at the University of Pittsburgh Medical Center. The idea is to use a technique called spectral rounding, driven by SDD systems, to improve the quality of retinal image segmentation. Koutis says that, so far, he and his colleagues have achieved imaging results whose quality, in many cases, has far exceeded that of previous methods.

“This medical imaging application highlights the utility of fast SDD solvers as a primitive operation in building more involved software systems,” says Koutis, who predicts that an increasing number of researchers will realize they can formulate some of their problems to benefit from SDD solvers. “We’re excited about the possibility that some of our ideas will have a positive impact in the future.”

#### Further Reading

Koutis, I., Miller, G.L., and Peng, R. Approaching optimality for solving SDD linear systems, *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, Las Vegas, NV, Oct. 23–26, 2010.

Spielman, D.A. and Teng, S-H. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems, *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, Chicago, IL, June 13–16, 2004.

Koutis, I., Miller, G.L., and Tolliver, D. Combinatorial preconditioners and multilevel solvers for problems in computer vision and image processing, *Proceedings of the 5th International Symposium on Advances in Visual Computing: Part I*, Las Vegas, NV, Nov. 30–Dec. 2, 2009.

Blelloch, G.E., Koutis, I., Miller, G.L., and Tangwongsan, K. Hierarchical diagonal blocking and precision reduction applied to combinatorial multigrid, *Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis*, New Orleans, LA, Nov. 13–19, 2010.

Teng, S.-H. The Laplacian paradigm: Emerging algorithms for massive graphs. *Lecture Notes in Computer Science 6108*, Kratochvíl, J., Li, A., Fiala, J., and Kolman, P. (Eds.), Springer-Verlag, Berlin, Germany, 2010.

Kirk L. Kroeker works in communications and has written extensively about the impact of emerging technologies.

© 2011 ACM 0001-0782/11/09 \$10.00

#### Security

## Small Companies Targeted

Cybercriminals have expanded their reach beyond large corporations and are increasingly attacking small businesses, according to *The Wall Street Journal*.

While break-ins at Sony and other well-known corporations have recently attracted widespread media attention, the boom in small business hacking has gone largely unnoticed.

Last year, the U.S. Secret Service and Verizon responded to a combined total of 761 data breaches, with 63% of them involving small businesses. In 2009, they responded to 141 data breaches, of which only 27% involved small businesses.

As small companies have grown increasingly reliant on computers in recent years, they have started to store more business-critical information online, including credit card information and other financial data. While large organizations usually employ rigorous security measures to safeguard sensitive data, small businesses’ relative lack of IT sophistication has made them easy prey. The wide array of small business systems now in place has created ample opportunities for hackers to develop new techniques for compromising those systems.

In one common ploy, hackers pilfer money by gaining access to companies’ online bank account login information. *The Journal* reports that Lease Duckwall of Abilene, KS, saw \$63,000 disappear from his company’s bank account when a hacker added nine fictitious employees to the company’s payroll. By the time Duckwall spotted the discrepancy and notified his bank to freeze the accounts, the hackers had already withdrawn \$22,000. To this day, Duckwall has no idea how the hackers gained access to his account information.

Small business hacking is becoming a “prolific problem,” Dean Kinsman, a special agent in the Federal Bureau of Investigation’s cyberdivision, told *the Journal*. “It’s going to get much worse before it gets better.”

—Alex Wright