

RTI International and University of Pennsylvania Model the Spread of Epidemics Using MATLAB and MathWorks Parallel Computing Tools

Public health officials often struggle to determine how best to prevent the spread of infectious disease. For livestock, they can institute quarantines or culling policies; for humans, they can issue travel advisories and provide immunizations. Until recently, officials relied on research based on heuristics and trial-and-error approaches to decide when and where to implement these policies. Today, sophisticated mathematical models make use of data from past outbreaks.

University of Pennsylvania (Penn) researchers use MATLAB® to develop models of epidemics among animals. Research Triangle Institute (RTI) extends those models to simulate infectious disease outbreaks among human populations.

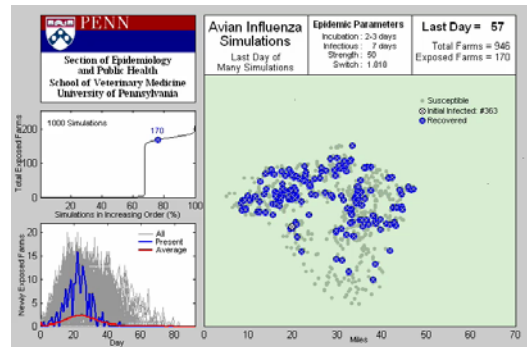
RTI uses MathWorks tools to run millions of simulations of the animal and human models in parallel on computer clusters.

The analyses of outbreaks among humans are part of the Models of Infectious Disease Agent Study (MIDAS), sponsored by the National Institutes of Health.

“Using MathWorks tools we can develop sophisticated computational models and leverage the massive computing power available today to more completely describe how epidemics spread and how they can be controlled,” says Chris Rorres, lecturer in epidemiology at Penn.

THE CHALLENGE

A catastrophic foot-and-mouth disease outbreak in 2001 in the U.K. provided a wealth of epidemiological data, including the size and location of infected farms and the date when they became infected. To analyze this data, Penn researchers needed



MATLAB simulation of an avian influenza epidemic.

to build discrete-time, discrete-space, stochastic models, which could be adapted to simulate the spread of other diseases. They also needed flexible tools for rapidly testing ideas, visualizing and animating simulation results, and communicating with non-technical stakeholders.

RTI researchers needed to scale up the simulations without placing a burden on their programmers. They needed software that would use their computer clusters in a way that was efficient, transparent, and easy to implement.

THE SOLUTION

Rorres and his colleagues at Penn used MATLAB to model and simulate the spread of disease among animals. With Parallel Computing Toolbox™, RTI researchers accelerated the simulations on a 64-node Linux-based computer cluster with 128 processors.

Working with data from approximately 1000 farms infected in the U.K. outbreak, Rorres developed a strategy for modeling the spread of foot-and-mouth disease.

THE CHALLENGE

Predict and control the spread of infectious diseases

THE SOLUTION

Use MATLAB to model epidemics and MathWorks parallel computing tools to accelerate the processing of millions of Monte Carlo simulations on a computer cluster

THE RESULTS

- Code reused, reducing development time
- Simulations completed 250 times faster
- Public health application demonstrated

He wrote MATLAB algorithms that calculated the probability that a farm would contract the disease. On each day-long time step of the simulation, MATLAB determined the probability that a single farm would be infected.

Rorres conducted thousands of Monte Carlo simulations using the same initial conditions and fine-tuned the contagiousness parameters until the results approximated an actual epidemic. He then simulated epidemics starting at other locations and tested the effectiveness of culling, vaccinations, and other control policies.

“We used MATLAB to create movies that revealed patterns in how an epidemic develops and that helped nontechnical audiences visualize our findings,” says Rorres.

Rorres adapted the foot-and-mouth disease model to simulate the spread of avian influenza. Working with a MathWorks consultant, RTI researchers used Parallel Computing Toolbox to parallelize Rorres’ model and used MATLAB Distributed Computing Server™ to execute the simulations on their computer cluster.

Diglio Simoni, senior computational scientist at RTI, found that under some initial conditions and parameter settings the epidemics died out very quickly, while under other conditions they were prolonged. To solve the load-balancing problem this trend created, RTI researchers wrote MATLAB scripts to programmatically identify the simulations that were likely to require relatively few computational resources.

Using the National Science Foundation’s TeraGrid infrastructure, RTI is now developing an agent-based model for

“

Using Parallel Computing Toolbox

we added four lines of code and wrote some simple task management scripts. Simulations that took months now run in a few days. MathWorks parallel computing tools enabled us to capitalize on the computing power of large clusters without a tremendous learning curve.

”

Diglio Simoni, RTI

MIDAS to simulate the spread of an epidemic—naturally occurring or released in a bioterrorist attack—throughout the U.S. population.

THE RESULTS

▪ **Code reused, reducing development time.** RTI reduced development time for the MIDAS model by 80% because researchers were able to reuse MATLAB code created by Rorres for previous animal models.

▪ **Simulations completed 250 times faster.** “What would have taken months in C, we did in just a few days with Parallel Computing Toolbox,” says Simoni. “We parallelized the application with a few lines of code, enabling us to complete epidemic simulations 250 times faster than before.”

▪ **Public health application demonstrated.** “One of our veterinary graduate students developed a standalone MATLAB application showing what an epidemic in poultry populations in Lancaster County, Pennsylvania, might look like,” says Rorres. “Emergency response personnel used the application to analyze what-if scenarios and evaluate policies.”

Learn more about RTI: www.rti.org

Learn more about the University of Pennsylvania: www.upenn.edu

APPLICATION AREAS

- Biotech
- Academia
- Algorithm development
- Statistics and data analysis
- Parallel computing

PRODUCTS USED

- MATLAB®
- Parallel Computing Toolbox™
- MATLAB Distributed Computing Server™

www.mathworks.com

 The MathWorks™

© 2009 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See www.mathworks.com/trademarks for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.

91705v00 04/09