Researchers Plan to Simulate Movements of 300 Million Americans

December 9 2008, by Lisa Zyga

By developing an extremely detailed simulation of the US population, researchers are hoping to understand how contagious diseases spread through society.

(PhysOrg.com) -- Researchers from Virginia Tech are developing a computer simulation that matches the movements of all 300 million people in towns across the US. The team hopes that the model will help them understand the spread of contagious diseases, fads, and traffic flows.

Currently, the researchers' model consists of about 100 million Americans, and they expect to be able to simulate the movement of all 300 million US residents in the next six months. To achieve this, the researchers use large amounts of publicly available demographic data, mostly from the US Census. Each synthetic American possesses as many

as 163 variables, which describe characteristics such as age, education level, occupation, and whether one lives with a family or alone.

The software, called EpiSimdemics, can provide an accurate simulation of the demographic attributes of groups composed of 1500 people or more. Based on the data, the software generates individuals to populate real US cities, giving them real street addresses and real jobs or schools within a reasonable distance from their address. Individuals are also matched to local grocery stores and shopping centers, which are identified through a database from Navteq, a digital mapping company.

One of the first applications for compiling all this data will be studying how contagious diseases, such as a flu epidemic, might spread through different regions. The software infects a few simulated individuals with the flu, and tracks them as they go about their daily lives. The model gives each person a different probability of responding to the virus, derived from the individual's data, such as age and general health.

Using data from all the interactions between infected individuals and others, the algorithm determines the number of new infections. The software treats each person and location as a separate set of calculations, so that many parts can be computed in parallel on a supercomputer. By breaking up the problem in this way, the researchers could significantly speed up the calculations.

By showing the path that a virus takes through a population, the simulation can help researchers implement effective public health intervention programs. The simulation can also determine when the infection peaks, representing the biggest burden on a city's health system, and preparing officials.

"The vision is for a Google-like interface, where you approach the system and ask it a question," says Christopher Barrett, who works on

the project and is the director of Virginia Tech's Network Dynamics and Simulation Science Laboratory. "The framework is there, and now we're pushing the system to larger and larger scales."

via: IEEE Spectrum