Simulation Modeling using Protégé

Henrik Eriksson
Magnus Morin
Joakim Ekberg
Johan Jenvald
Toomas Timpka
Models of community structures

- Geographical
- Logistical
- Social
- Cultural

Mixing-group approach
People get infected where they meet other people
Number of secondary cases (Ro)

Results from 1000 simulation runs: Frequency of secondary cases

In most cases, we get no outbreak

Super spreaders
Simulator requirements

- Fast turn-around time (<24h)
  - Updated simulation as more information become available
- Transparent, user-friendly models
- Pluggable models
  - Interchangeable disease model
  - Alternative community models (e.g., actual and randomized)
- Scalable computational environment
- Visualization
Pandemic modeling and simulation
Simulation environment—architectural layers

- Community model
- Disease model
- Intervention model
- Population generator
- Intermediate specification format (XML)
- Simulator engine

Protégé with extension

Custom simulator (C++)
Simulation Architecture

Protégé

Protégé extensions

Simulation model (scenario ontology)

Scenario developer

XML-based simulation parameters

Simulation engine (Computational environment)

Results

Report generator
Simulation model

Ontology-based simulation model
Editing of community definition
Protégé tab extension for scenario editing

Submodels used in the scenario
Protégé tab for simulation job specification

Job parameters

List of scenarios to simulate
Simulation engine and computational environment

- **Requirement**
  - Interactive simulation
  - Dynamic scaling

- **Problem**
  - Supercomputers are fast, but using them takes too long time (job queues)

- **Solution**
  - Separation of modeling and execution environments
    - Protégé (Java) versus custom simulator (C++)
  - Condor
    - Pool of machines (basic resource)
    - Rent additional machines as needed (Amazon EC2)
  - Web application for managing simulation nodes and simulation jobs
    - Google Web Toolkit (GWT)
Computational environment

- The Condor platform
  - System for managing clusters of dedicated compute nodes
  - Workload management system for compute-intensive jobs
  - Batch system with a job queueing mechanism, scheduling, resource monitoring and management, etc.
  - Matching of resource requests (jobs) with resource offers (machines)
  - Developed by University of Wisconsin-Madison (UW-Madison)

- Condor components/actors
  - Condor manager
    - Collects information about the pool of machines
    - Manages the job queue
    - Dispatches tasks to workers
  - Condor workers
    - Machines that execute tasks
  - Storage system
    - Storage of input and output data
Tasks parallelized at the level of
- alternative scenarios
- different randomized communities
Computational environment (cont.)
Multiple Condor pool environment

Computational environment

- Internal Network
- Condor worker
- Condor worker
- Condor worker
- Condor worker
- Condor worker

Other Cloud

- Condor worker
- Condor worker
- Condor worker
- Condor worker

Amazon Elastic Compute Cloud (EC2)

- Condor worker
- Condor worker
- Condor worker
- Condor worker
- Condor worker
- Condor worker
- Condor worker
- Condor worker

Web client (GWT)

- submit jobs
- manage results
- rent machines (GWT-RPC)

Web server (Tomcat)

- Condor Manager

Protégé (+plugins)

- store models
- store simulator executables

Storage System

AWS management
Web interface for Amazon EC2 management

- Bootable images (with Condor)
- Running machines
Condor job queue

Two jobs running

<table>
<thead>
<tr>
<th>ClusterId</th>
<th>ProcId</th>
<th>Status</th>
<th>Args</th>
<th>RemoteHost</th>
<th>JobPrio</th>
<th>RunTime</th>
<th>Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>197</td>
<td>0</td>
<td>Running</td>
<td>run1_A09_A2B1_Intervention-1.xml.tar.l</td>
<td><a href="mailto:slot2@ec2-79-125-40-194.eu">slot2@ec2-79-125-40-194.eu</a></td>
<td>0</td>
<td>0:00:09</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>1</td>
<td>Running</td>
<td>run1_A09_A2B1_Intervention-2.xml.tar.l</td>
<td><a href="mailto:slot1@ec2-79-125-40-194.eu">slot1@ec2-79-125-40-194.eu</a></td>
<td>0</td>
<td>0:00:09</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>2</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-3.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>3</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-4.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>4</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-5.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>5</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-6.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>6</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-7.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>7</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-8.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>8</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-9.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>9</td>
<td>Idle</td>
<td>run1_A09_A2B1_Intervention-10.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>10</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-1.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>11</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-2.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>12</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-3.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>13</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-4.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>14</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-5.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>15</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-6.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>16</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-7.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>17</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-8.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>18</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-9.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
<tr>
<td>197</td>
<td>19</td>
<td>Idle</td>
<td>run1_A09_A1B2_Intervention-10.xml.tar.l</td>
<td></td>
<td>0</td>
<td>0:00:00</td>
<td>5/14 11:02:47</td>
</tr>
</tbody>
</table>
Simulation results

New and completed jobs

Results overview
Discussion and Conclusion

- **Summary**
  - Simulation of influenza outbreaks benefits from a clear separation of modeling and implementation
  - Ontologies provide a suitable representation scheme for such epidemiological models
  - Ontology management during a factual pandemic outbreak is supported by the maintenance of a scenario library with a collection of instances representing the scenarios

- **Implementation**
  - Modeling and simulation environment based on ontology models in Protégé
  - Corresponding cloud-based execution environment

- **Continued work**
  - Improved flexibility of simulator engine
  - Submitting simulation jobs from Protégé
  - Controlling Amazon EC2 and Condor from Protégé
  - Visualization of results