Dynamic Adaptive Disaster Simulation: Developing A Predictive Model of Emergency Behavior Using Cell Phone and GIS Data

Francis Chen, Zhi Zhai, Greg Madey
Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN

George Mason University, January 31st, 2011

1The research presented in this paper is based in part upon work supported by the National Science Foundation, CISE/CNS-DDDAS, Award #0540348.
Outline

- Introduction
  - Motivation
  - Previous Work
  - Contributions

- Our approach
  - Modeling Process
  - Calibration

- Validation, Results, and Discussion

- Conclusions and Future Work
Why Model Populations?

- Hurricane Katrina
  - No comprehensive information on population movement
    - 70,000 left in New Orleans
  - Resources distributed inefficiently
    - High ground areas (Superdome)
    - “fascinating phenomena”

http://www.nerdylorrin.net/jerry/Katrina/KatrinaSuperdome.html
Why Model Populations? (cont.)

- London attacks on royal couple
  - Civil disorder was not reliably tracked
    - 100-person mob allowed to disrupt motorcade
    - Alternative routes were possible

Existing Methods of Population Modeling

- Agent-based modeling
- Flow/continuum-based modeling

Challenges in Disaster Modeling

- Restricted to pre-programmed scenarios
- Based on speculations and assumptions
  - 25-40% difference in predicted evacuation times
- Online validation and data incorporation are difficult
- Dynamic Data-Driven Application Systems (DDDAS)
  - Better for real-time, adaptive predictions (Darema 2006)

The WIPER Project

- Wireless Phone-based Emergency Response System
  - Cell phones used as dynamic data source

- Simulation and Prediction
  - Pedestrian and vehicle agents
  - Basic movements: flee, flock, jam

- More work is needed (model complexity, adapt to scenarios)
Contributions

- Developed Dynamic Adaptive Disaster Simulation (DADS)
  - Proof-of-concept
  - DDDAS concepts
    - Adapts to specific scenarios
    - Continuously refines predictions
  - Can incorporate data
    - Geographic Information System (GIS)
    - Streaming real-time cell phone location data
    - Tested on synthetic cell phone data
  - Netlogo language and modeling environment, version 4.1.1
    - Used GIS extension
Outline

- Introduction
  - Motivation
  - Previous Work
  - Contributions

- Our approach
  - Modeling Process
  - Calibration

- Validation, Results, and Discussion

- Conclusions and Future Work
System Architecture

Real data Available?

Yes

No

Program synthetic agent movements to resemble disaster movement

Synthetic agents report positions as if using cell phones

Trace cell user movements

Cell phone location data set
Agent Types

- **Synthetic agents**
  - Move as people would during a disaster
  - Represent real cell phone users
    - Movements generate synthetic cell phone data
    - Necessary because of nondisclosure agreements (NDA) with cell phone company—I am a minor

- **Predictive agents**
  - Move according to inferences drawn from cell phone data
  - Represent predictions of future population movement
    - Attempting to adaptively model disaster evacuation
    - Emergency management can be conducted based on agents’ predictions
Place names removed to maintain anonymity
Cell Phone Data

- Networks must be able to constantly track cell phones
  - Call Data Records (CDR)
  - Accuracy varies
- Phone-integrated GPS technology

```
20070127 000400 @6f19d5 @fafd42 10004
20070127 000600 @69a50b @fafd42 10004
20070127 000600 @31f919 @fafd42 10004
20070127 000700 @570f5c @fafd42 10004
20070127 000700 @e940a6 @fafd42 10893
20070127 000800 @8e97cd @fafd42 10343
20070127 000900 @a620f5 @fafd42 10005
20070127 000900 @687ae0 @fafd42 10011
20070127 001000 @2297d7 @fafd42 10011
```

http://tuberose.com/Graphics/cell%20tower.jpeg

http://googlephonetracking.com/
Modeling Approach

- Dynamic Potential Fields or elevation fields (Park 2009)
  - Agents move from high to low potential
  - Conceptually portrayed as a terrain of varying elevations
  - Used for both synthetic data and DADS itself

- Use fluid-like agents (Helbing 2002)

Example

Implementing Modeling Approach

- Elevation field represented as matrix (Wilensky 2006)
  - Each element represents a patch of ground
  - Convolve the matrix with kernels:
    
    $\begin{bmatrix}
    1 & 1 & 1 \\
    0 & 0 & 0 \\
    -1 & -1 & -1
    \end{bmatrix}$  
    
    $\begin{bmatrix}
    -1 & 0 & 1 \\
    -1 & 0 & 1 \\
    -1 & 0 & 1
    \end{bmatrix}$

  - For each of two gradient matrices:
    - Calculate aspect: $a(x, y) = \arctan(y/x)$
    - Done in Netlogo, with GIS Extension

- Agents continuously set headings to match aspect of patch
Generating Synthetic Data

- Synthetic elevation field
  - Types of regions in a scenario
    - Disaster (fixed)
    - Dangerous (random)
    - Safe (random)
    - Roads (fixed)

- 3200+ synthetic agents
  - Realistic pedestrian speeds

- Random scenarios
  - Example

Disaster location  Dangerous location  Safe location  Road
Conducting Inference on Data

- Uses “vision cone” (Torrens 2007)
- Used as “cone of inference”
  - Patches inside the cone are inferred to be attractive
    - When a synthetic agent moves, decrease predictive elevation of patches
    - Generate a field of predictive elevations
  - DADS predictive agents move on predictive elevation field
    - Represent prediction of future locations of cell phone users
- Example
Problem becomes that of “reconstructing” a reasonable predictive elevation field

- Must accurately capture factors influencing movement
Summary of Methods

- Generate synthetic elevation field
  - Synthetic agents move on it to produce synthetic data

- Conduct inference as location data streams in
  - Generate predictive elevation field

- Predictive agents move on predictive field
  - Represent predictions of population movement
  - Example
Measuring Simulation Quality

- Manhattan distance metric (Schoenharl 2008)
  - Compare synthetic vs. predictive agents

  - $p_i$ and $q_i$ represent numbers of each agent type at each cell tower

- Smaller Manhattan distance = more accurate simulation
Experimental Setup and Results

- Identified optimal values for vision cone angle and range
- Multi-resolution approach
  - Coarse, then finer parameter sweeps
  - Compared predictions of all possible parameter pairs
    - Evaluated in three random scenarios

Best angle: 45°
Best range: 200m
Outline

- Introduction
  - Motivation
  - Previous Work
  - Contributions

- Our approach
  - Modeling Process
  - Calibration

- Validation, Results, and Discussion

- Conclusions and Future Work
Validation

- Verifies that a model “is a reasonably accurate representation of the real world” (Xiang et al. 2005)
  - **Internal Validation**
    - Measures stability
  - **Predictive Validation**
    - Measures predictive accuracy
  - **Other Validation Methods**
    - Online validation is an “open research question” (Schoenharl 2007)
Internal Validation

- 500 runs; measured final predictions (75 minutes in advance)
  - Different randomly generated scenario each time

- Mean correct percentage: 77.30%; standard deviation: 3.87%
Predictive Validation

- An average of 77.30% of predictive agents made correct predictions, 75 minutes in advance.

- Disadvantage of quantitative validation:
  - Predictions are only correct if in the correct serving cell.

- Qualitative validation is necessary.
Qualitative Predictive Validation

- Compare paths taken by synthetic/predictive agents

**Diagram:**
- Synthetic
- Predictive
Qualitative Predictive Validation

- Dangerous location
- Safe location
- Disaster location
- Road

- Shapes of attractive and repulsive regions
  - Can indicate disaster type
  - Can use to refine simulation
  - Future Work

- Road networks
  - Usable for traffic simulations
Predictions of population locations at 10:20 a.m.
- Green lines: at 8:05, red lines: at 8:35, blue lines: at 9:05

Reflects the DDDAS concept of dynamically updating simulations
Discussion

- Assumptions—“a model is only as good as the assumptions on which it is based”
  - Homogeneous agents
  - No crowd dynamics
  - Synthetic data
  - No restrictions on agent vision or movement
Outline

- Introduction
  - Motivation
  - Previous Work
  - Contributions

- Our approach
  - Modeling Process
  - Calibration

- Validation, Results, and Discussion

- Conclusions and Future Work
Summary

- **DADS** uses streaming cell phone location data to simulate and predict population movement in disasters
  - Makes use of emergent intelligence
  - Can analyze historical data
    - Study tool for past disasters
  - GPS will further increase utility

- **Demonstrates DDDAS**
  - Adapts to specific scenarios and constantly improves

- **Validated on synthetic data**
  - Predictively and internally valid
  - Provides useful inferences in situations like Katrina
    - Helpful in evacuations, even if disaster disables cell service
Important Considerations

- Distinguish evacuees from responders
  - Potentially misleading data
  - Heterogeneous agents?

- Cell phones as a data source
  - Limited power supply
    - Could be problematic in the long term
  - Cell towers vulnerable
    - Earthquakes
    - Volcanic ash
  - GPS technology
  - How quickly can we draw accurate inferences?

---

http://www.wirelessestimator.com/t_content.cfm?pagename=Hurricane%20Ike%20telecom

Future Work—DADS Itself

- Test on real cell phone location data
  - Allow for adjustment of data reception
    - DDDAS concept—sensor adjustment

- Further assess modeling techniques
  - Increased realism
    - Agent heterogeneity
    - Crowd dynamics

- More sophisticated methods of parameterization

- Explore more ways to use cell phone data
  - Examine call volume, distribution, location, etc.
Future Work—Population Simulation

- **Large-scale**
  - Modeling citywide or global movement patterns in other situations

- **Small-scale**
  - Modeling individual behavior
  - Depicting movement and/or evacuation in a building

- **Tool for study as well as prediction**
Future Work—Adaptive Simulation

- Simulations designed to adapt to streaming data
- Modeling landslides in China
  - Caused by dams, mining, and deforestation
- Better sensor networks enable this sort of technology
  - DADS is an example

Ethical Concerns

- “The Tradeoff of Confidentiality and Access” (NRC 2007)
  - Must sacrifice precision for privacy
    - WIPER—aggregated by voronoi cell
  - How much precision is needed?
    - DADS appears to require precision
    - “Naïve realism”

- Possible solutions
  - Opt-out (Johnston 2010) or opt-in?
    - How many cell phones are needed?
  - Aggregation and other tactics
  - “Data enclaves” and legal or licensing systems
Thank you!

Questions?

A paper describing this research has been accepted for SpringSim/ADS 2011, and can be found at the following URL: http://www.nd.edu/~dddas/Papers/papers.html