

A Sensitivity Analysis for Deriving Dynamic and Evolutionary Rules in an Artificial Immune System-Cellular Automata Model

B. Curtis¹, C. Willy¹, and J. Bischoff¹
¹Dept. of Eng. Management & Systems Engineering, George Washington University, DC, USA

Abstract - There have been several studies advocating the need for, and the feasibility of, using advanced techniques to support decision makers in urban planning and resource monitoring. One such advanced technique includes a framework that leverages the use of remote sensing and geospatial information systems (GIS) in conjunction with cellular automata (CA) to monitor land use / land change phenomena like urban sprawling. However, little research has been performed to analyze these frameworks’ sensitivity to the input data (e.g. imagery). New technology is promising better data more frequently; all with an associated price tag. Understanding sensitivity provides decision-makers and analysts the necessary information to procure just the right amount, and type, of data. Our research focuses on arming analysts and decision makers with this information.

Keywords: Sensitivity, GIS, Remote Sensing, Cellular Automata

1 Introduction

Urban environments are complex systems presenting dynamic spatial and temporal features along with emergent and non-linear growth behaviors. Understanding and predicting these dynamic phenomena can be difficult, however useful. One example of a concrete problem at the nexus of engineering and land use are the issues associated with urban sprawl – the migration, or expansion, of our populations away from city centers outward towards low density, residential, and usually highly automobile-dependent areas. The research of Ewing, et al. demonstrated a connection between urban sprawl and an epidemic affecting the United States today, obesity, stating that "residents of more compact counties have lower BMIs and lower probabilities of obesity and chronic diseases" [1]. The CDC also published a report [2] attributing this epidemic to local policies and our physical environments in which we live, including the lack of physical activity due to residential zoning strategies requiring people to drive, vice walk, to work and school simply because it is too far. Ultimately, according to research conducted by Ogden, et al. [3], more than one-third of adults and 17% of youth are considered obese in the United States today.

As a result of these findings, several studies have been published advocating the need for, and the feasibility of, using advanced techniques to model and simulate urban sprawl and provide decision makers the tools necessary for urban planning and resource monitoring. One such advanced technique includes a framework that leverages the use of remote sensing and geospatial information systems (GIS) in conjunction with a cellular automata (CA) to monitor land use / land change phenomenon like urban sprawling. While this technique has been shown to be a viable solution to simulate the complex nature of urban sprawl, little research has been conducted analyzing the sensitivity to input data (i.e. imagery). Therefore, this research seeks to analyze the relationship between a GIS-CA model and the imagery which feeds it; looking at frequency, imagery resolution, and fragmentation (i.e. partial coverage). A case study simulating urban sprawl for the city of Albuquerque, New Mexico will be used to analyze the relationship between GIS-CA and the data which feeds it.

2 Methodology

The use of geospatial information systems (GIS) as we know them today has been around for decades; over 180 years in its purest form of spatial analysis. In the simplest of terms, GIS is a system used to manipulate, analyze, and visualize all different types of geospatial information (e.g. land use, roads/streets, bodies of waters, elevation information, etc.). When used in raster form, GIS information is stored as a layered grid of data; each layer is a two-dimension matrix representing a specific piece of information as it relates to the geographic area being studied.

The use of cellular automata (CA) to simulate the evolution of complex systems, spatially and temporally, has been widely applied to urban sprawl research. CA-based models use a ‘bottoms up’ approach where a simple set of transition rules govern the interactions between cells. They have the ability to represent non-linear, spatially dependent, stochastic processes and simulate the evolution of the complex systems [4]. According to Liu, et al. [5], “this ‘bottoms-up’ approach coincides with complexity theories stating that a complex system comes from the interactions of simple subsystems.” The fact that a CA-model is cell based – or a
two-dimensional matrix – makes it perfect to couple with raster GIS and remote sensing data for studying the complex nature of urban sprawl.

One method for deriving the dynamic transition rules of a CA-model takes its inspiration from nature; an intelligent computational technique capable of solving complex problems known as artificial immune systems (AIS). Liu et al. [5] first used an AIS-based CA model to determine policy impacts on land use and found its ability to adapt, learn, organize, and memorize new information was extremely promising for complex geographical problems. Much like natural systems, AIS uses the concept of ‘antigens’ and ‘antibodies’ to derive the transition ‘rules’ for the CA-model. More specifically, cells needing to be classified are the ‘antigens’ and the classifiers which will assign the proper land use (e.g. urban) to such cells are the ‘antibodies’. He, et al. [6] proposed a process to calculate the evolution probability for the urban CA-model through a standard recycling process: defining antigens, generating an initial set of antibodies, calculating antibody/antigen affinities, clonal selection, and the mutation and updating of the antibodies. The high-level process flow of the AIS-based CA model is described in Figure 1. The crux of our research focuses on this model’s ability to be repeatedly updated when new remotely sensed imagery becomes available; dynamically adapting and learning to the introduction of new antigens (i.e. information) into the simulation.

A hypothetical example for our temporal frequency based scenario would include using historical data to seed the antigen library and then build the antibody library. Year sets could be spaced by 10 years (e.g. 1990, 2000, 2010), 5 years (1990, 1995, … , 2010), and 1 year or less (depending on available data). Once the transition rules have been obtained, a simulation forward to the most recent available imagery (e.g. 2015) could be used to measure the accuracy of each year set. To measure the findings between simulation results and actual situations, we use a cell-level comparison analysis (i.e. pixel by pixel) and apply a ‘figure of merit’ (FoM) metric [7]. This ‘FoM’ is a ratio by which we measure the number of cells correctly simulated as urbanized cells (numerator) divided by the total number of instances.

3 Conclusions

New technologies are promising to deliver a radical change in access to information. However, our research indicates a lack of understanding into how the complex and dynamic models used to monitor sprawl can leverage this new technology. Therefore, this research will answer the questions if more, finer, and partial remote sensing data can be used to improve an AIS-based CA model’s accuracy.

4 References