Research Article

‘Situating’ Simulation to Model Human Spatio-Temporal Interactions: An Example Using Crime Events

Elizabeth R Groff
Department of Criminal Justice
Temple University

Abstract

Many social phenomena have a spatio-temporal dimension and involve dynamic decisions made by individuals. In the past, researchers have often turned to geographic information systems (GIS) to model these interactions. Although GIS provide a powerful tool for examining the spatial aspects of these interactions, they are unable to model the dynamic, individual-level interactions across time and space. In an attempt to address these issues, some researchers have begun to use simulation models. But these models rely on artificial landscapes that do not take into account the environment in which humans move and interact. This research presents the methodology for ‘situating’ simulation through the use of a new modeling tool, Agent Analyst, which integrates agent-based modeling (ABM) and GIS. Three versions of a model of street robbery are presented to illustrate the importance of using ‘real’ data to inform agent activity spaces and movement. The successful implementation of this model demonstrates that: (1) agents can move along existing street networks; (2) land use patterns can be used to realistically distribute agent’s homes and activities across a city; and (3) the incidence and pattern of street robberies is significantly different when ‘real’ data are used.

1 Introduction

The importance of incorporating space and time into research on human behavior has long been recognized in a variety of disciplines (Engel-Frisch 1943, Hawley 1950, Chorley and Haggett 1967, Harvey 1969, Horton and Reynolds 1971, Hagerstrand 1973, Miller 1991, Sampson 1993). Addressing spatio-temporal interactions among individuals and
their environments is a challenging undertaking with two major hurdles. First, it is not enough to be able to capture the multitude of individual decisions that occur within unique contexts; the method must also be able to accommodate dynamic changes in the characteristics of individuals and situations that influence the outcome of subsequent interactions. Second, individual-level data that can support these types of studies must be obtained. Although there has been tremendous growth in the availability of micro-level data describing places, data regarding the daily activities of individuals (e.g. the type, location, and duration of activity) remains sparse. While these data are essential to modeling the spatio-temporal convergence and interaction of individuals, they are unlikely to become available due to privacy concerns (O’Sullivan 2004).

Given the large quantity of micro level environmental data and mature software packages, researchers frequently turn to geographic information systems (GIS) to model human behavior (Miller 1991, Kwan 1998, An et al. 2005). GIS provide a powerful tool for collecting, managing and displaying the multitude of spatially explicit data available on places but they are unable to model the dynamic, individual-level interactions across time. The inability of GIS to accommodate time is a well-known issue that remains unsolved despite a great deal of attention (Peuquet 1994, 2002; Brown et al. 2005; Albrecht 2007). Physical scientists address this issue in their process models by preparing their data in a GIS and then analyzing them in a dynamic model (for examples of this strategy see the GIS & Environmental Modeling conference proceedings at http://www.ncgia.ucsb.edu/).

Simulation modeling offers a method capable of capturing dynamic interactions among individuals taking place at the micro level and their relationship to macro level patterns. All models, simulation or statistical, involve the creation of a simplified representation of a social phenomenon (Gilbert and Terna 1999). Simulation modeling has three main advantages over statistical models. First, it allows heterogeneity among individuals that more closely approximates the variety found in life. Second, it is able to accommodate the non-linear relationships present in dynamic and complex interactions (Epstein and Axtell 1996, Gilbert and Terna 1999, Dibble 2003). Third, simulation modeling can be used in situations where little or no empirical data are available. Statistical models require data, either empirical or simulated.

Agent-based modeling (ABM) is one type of simulation that employs a bottom-up approach in which agents are imbued with unique characteristics and general behavioral rules (Epstein and Axtell 1996, Gilbert and Terna 1999, Gilbert and Troitzsch 1999). An agent is an “autonomous goal-directed software entity” (O’Sullivan and Haklay 2000) that most often represents a person but can also represent organizations, neighborhoods, etc. The decisions of the individuals determine the outcome of interactions with other agents and subsequently, those decisions dynamically change the characteristics of agents. ABM agents are removed from their ‘real-word’ situation and placed in an artificial world (O’Sullivan and Haklay 2000). The use of artificial landscapes which do not take into account the impact of the environment in which individuals move and interact represents one significant drawback to ABM.

In order to model individuals in a non-artificial environment, an approach is needed that combines the strengths of ABM and GIS (Albrecht 2005, An et al. 2005, Brown et al. 2005). A combined ABM/GIS simulation model integrates the advantages of autonomous agents found in agent-based modeling with the spatial explicitness of a GIS. This allows agents to interact on city streets in a particular environment. It also enables their activities during the simulation to be informed by the distribution of opportunities
for housing, employment, shopping, and recreation across the urban backcloth. The result should be patterns of agent spatial behavior that are more representative of actual human activity patterns than when agents are created with and interact on artificial landscapes. By moving away from virtual landscapes, human-environment interactions can be quantified and tested. For example, if a street is closed off how will that impact the outcome pattern? Does it matter if the street is in a commercial area that is thriving or one that is run-down?

The STREETS model of pedestrian movement recognizes the importance of the street configuration and the locations of activity attractors along the network (Haklay et al. 2001). Similar to the models of physical systems mentioned earlier, STREETS uses information from a GIS to assign geographically-informed properties and schedules to agents which are then used in the dynamic simulation. The focus in STREETS is on the way finding behavior of pedestrians in a street center and how well the model can produce ‘plausible’ movement patterns.

A natural application for combined ABM/GIS simulation model is toward achieving a better understanding of the crime event in its situational context. Some researchers have already begun to explore the use of simulation for capturing the dynamic interactions taking place at the micro level and their relationship to macro level patterns (Brantingham and Brantingham 2003, 2004; Gunderson and Brown 2003; Brantingham and Groff 2004; Eck and Liu 2004; Liu et al. 2005). This research extends those efforts by developing a model of street robbery using Agent Analyst software (i.e. an integrated ABM/GIS). The GIS component of the software enables the creation of realistic activity spaces and the movement of agents along vector street networks. The ABM component controls the temporal elements of the simulation and the interaction of agents with one another. By combining the two, the situational elements of the convergence of offender and victim at a specific place and time are simulated.

Specifically, this study develops and tests a method for successful implementation of: (1) agent movement on a GIS street network; and (2) the creation of unique routine activity spaces for individuals that reflect the distribution of population, jobs and retail/services/recreation opportunities in Seattle, Washington. In contrast with the STREETS model, this research does not focus on reproducing intra-city movement by individuals but rather on demonstrating how that movement is related to a specific social outcome, street robbery. The crime of street robbery is a natural choice for this type of model because it stems from the interaction of individuals in a public area. In addition, it is an instrumental crime (for economic gain), and thus more likely to be the result of a rational decision than an expressive crime.

The remainder of this paper is organized as follows. First, the theoretical basis for the model is covered and a conceptual model presented. Next the implementation details including model rules and input data are provided. Lastly, the results from the same model run on a grid versus a street network are described and the implications for future research are discussed.

2 Theoretical Basis for a Street Robbery Model

Both criminological and geographical knowledge are important to the development of a conceptual model of street robbery events. Opportunity theories of crime, ones that address the elements of the situation in which the offender makes the decision to offend,
form the basis of the model. Specifically, routine activity theory (Cohen and Felson 1979) provides the structure of the model and rational choice theory (Clarke and Cornish 1985, 2001) is used to guide offender decision making. The structure and timing of agent’s activities and movement along the street network is informed by the major theoretical perspectives that address routine activity spaces.

2.1 Criminological Theory

Cohen and Felson’s (1979) original formulation of routine activity theory identifies the key to increases in crime as the shift of routine activities away from home. The theorists hypothesize that as individuals spend more time away from home, crime will increase. As originally conceptualized, routine activity theory identifies the convergence of motivated offender, suitable target, and the lack of a capable guardian at a particular place and time as the core elements necessary for a crime to occur. They emphasize that crimes occur when the normal everyday activities of offenders and victims intersect with no guardian present.

The theorists recognize the importance of routine activities in influencing when and where victims and offenders converge but they do not directly address the details of human mobility. They view routine activities as the key dynamic element in determining aggregate crime rates because it affects the three other elements necessary for a crime, motivated offender, suitable target, and guardianship. Changes in routine activities directly impact the frequency of convergence among these elements which in turn, increase or decrease overall crime rates. Thus, the theorists neatly tie the interaction of clearly defined elements of a crime to societal level crime rates. These four elements of offender, target, guardian, and routine activities form the main constructs of the model.

As previously mentioned, routine activity theory pays little attention to the source of the offender’s motivation and assumes a supply of motivated offenders. Consequently, the model developed here relies on rational choice theory for the specifics of offender decision-making (Clarke and Cornish 1985). Rational choice theory is based on the economic principle of expected utility where each individual’s decisions are predicated upon balancing projected benefits against projected costs of activities. The theory does not assume people have perfect knowledge but rather recognizes that offenders make the decision to commit a particular offense based on the characteristics of the specific situation using bounded rationality (i.e. imperfect knowledge) and taking into account three factors: the suitability of the situation, the presence of a viable target and the level of guardianship. Rational choice theory also assumes that offender spatial behavior is essentially similar to that of non-offenders so all non-police agents in the model have identical movement rules.

2.2 Activity Spaces

One of the core concepts in routine activity theory involves the necessity of the convergence of victims and offenders in space and time. The specific ‘where’ and ‘when’ of convergence stems from the routine behavior patterns of each actor involved. Thus representing the spatio-temporal aspects of human behavior that facilitate convergence is a critical element in modeling street robbery events since it is the interactions between humans and their environment that serve as the source of explanation of observed spatial

A large quantity of research is available to inform agent movement and routine activities in the model (see Groff 2007c for a review). In general, people tend to have an area within which they conduct their daily activities. Some researchers term this area an activity space (Horton and Reynolds 1971), others call it a domain (Hägerstrand 1970, 1975) or a potential path area (Miller 1991). This area encompasses both the locations that are visited and the paths taken among those locations. Different perspectives have their own terms for these locations and paths. Locations that are visited are called stations (Hägerstrand 1970, 1975), nodes (Lynch 1960; Brantingham and Brantingham 1981, 1993; Miller 1991), or anchorpoints (Golledge 1978, Golledge and Stimson 1997). These are the places where the majority of human interaction occurs. The particular routes taken among the locations are termed paths (Lynch 1960; Hägerstrand 1970, 1975). None of these elements are static, for example, the shape and size of areas (i.e. activity spaces) can change as people change jobs (i.e. nodes) or as their circumstances change (Hägerstrand 1970).

Regardless of the terminology, home tends to be the dominant place in any activity space. Travel tends to be concentrated along certain routinely frequented paths. Frequently traveled paths may be important factors in determining aggregate crime patterns because they bring offenders and victims together in space and time. Individual’s travel patterns are influenced by constraints (i.e. temporal, economic and spatial) on their ability to take advantage of opportunities for housing, employment, recreation, etc. Together this body of research provides a strong basis for conceptualizing routine activity spaces of individuals as a set of places and the paths between those places.

2.3 Conceptual Model

The preceding review of research identifies the basic elements represented in the conceptual model (see boxes in Figure 1). The conceptual model identifies two classes of agents representing people, civilians and police. Civilians have activity spaces and can take on different roles (i.e. offender, victim, or guardian) depending on the particular situation. Police exist only as agents of formal guardianship. Civilians with criminal propensity can potentially take on any one of three roles, offender, victim, or guardian. Civilians without criminal propensity can be either victims or guardians. In addition to criminal propensity, each civilian in the model has a unique set of characteristics that include wealth and employment status.

Two other spatial elements are important to convergence of people in a model of street robbery. One is the activity spaces of the people and the other is the network of streets available for travel. The size and form of activity spaces is influenced by the distribution of residential housing, jobs, schools, retail, and services. Each civilian has a unique activity space reflecting the places they visit. Once convergence occurs, factors such as guardianship and suitability of target are considered by the offender when making the decision whether or not to commit a robbery.

3 Model Implementation and Testing

This section describes a methodology for ‘situating’ simulation models including software, data, movement and activity space formulation. The model of street robbery developed
is based on the core elements of routine activity theory: a motivated offender, suitable target, and the lack of a capable guardian. The spatial and temporal aspects of agent's routine activities are important in determining the frequency with which those elements converge in space-time. These concepts form the basis for three of the agent classes in the model: place, civilian, and police officer (Figure 2). The fourth class, active nodes, is generic and serves as a computational device to identify which nodes have agents present at each tick of the model. The active node class improves the performance of the model by restricting the set of places that have to be checked each minute of the model. The remainder of this section describes the software package, data, agents, and agent behavior in the model as implemented.

3.1 Agent Analyst: GIS/ABM Integration

The method uses a new software package, Agent Analyst, which integrates GIS and ABM to provide a platform for the dynamic modeling of individuals across space and time. This package follows the middleware approach in which the temporal relationships are handled by the agent-based modeling software and the topological relationships are managed by the GIS (Brown et al. 2005). Agent Analyst combines two of the most popular packages for ABM and GIS; the Recursive Porous Agent Simulation Toolkit (Repast) (North et al. 2006) and ArcGIS (ESRI 2005). To make the software easier to use, Agent Analyst is built using the rapid development version of Repast called Repast for Python Scripting (RepastPy) which has a graphical user interface that automates
much of the programming to create the framework of a model. Agent Analyst is designed to be added into ArcGIS as a toolbox. Once the toolbox is added in ArcGIS, individual models can access shapefiles allowing: (1) individual agents to become spatially aware, and (2) the visualization of agent movement and decision outcomes (e.g. locations of crimes).

The integration of GIS and ABM enables the exploration of how individual decisions by heterogeneous agents translate into aggregate rates of street robbery. The combination of ABM and GIS offers several advantages over either used separately. ABMs permit the researcher to: (1) collect data about the characteristics of each individual present during an interaction; (2) randomly assign characteristics to agents greatly reducing the possibility of systematic bias; (3) allow agents to make independent decisions within behavioral guidelines; and (4) systematically vary one attribute while holding all others constant to undertake controlled, repeatable experiments (Epstein and Axtell 1996, Gilbert and Terna 1999, Dibble 2003). GIS make it possible to take into account how the characteristics of the real environment (i.e. street network, distribution of homes, jobs and activities) impact the activity spaces of agents. In addition, it provides the ability to explore the role of routine activities in facilitating the space-time convergence of a motivated offender and a suitable target, without a capable guardian present.

3.2 Data, Parameters, and Landscapes

This research uses both synthetic and empirical data to create two landscapes and to inform the activity spaces of the agents. The synthetic data consists of a uniform grid with approximately the same number of intersections \( n = 15,975 \) as Seattle’s street

\[\text{Figure 2} \quad \text{Classes in the street robbery model}\]
network \((n = 16,035)\). Empirical data describing conditions in Seattle are collected to inform the activity spaces of agents in the model: (1) total population; (2) total employment; (3) total potential activities and (4) streets. Block group level population data are used to describe the distribution of residences across Seattle (U.S. Census Bureau 2000). Employment data are used to describe the number of employees per zip code area. Total potential activity locations are quantified through the use of retail and service establishments (e.g. grocery stores, convenience stores, dry cleaners, gyms, etc.). The final input data set is the street network which is derived from the King County Street Network Database (SND) file and is used to structure the agent’s movements.

Two landscapes are created, one from a uniform grid and the other from the street network of Seattle, Washington. Two classes in the model have to do with landscape, places and active nodes (Figure 2). The place class of agents in the model represents the street nodes or grid nodes depending on the landscape. Each place has attributes that are updated during the model run (e.g. total robberies, total visits, etc.). As a vector agent class, places are directly linked to shapefiles representing street intersections/grid intersections and provide the only mechanism for visualization of the model while it is running and after a model year.

In addition to the input data describing Seattle, a number of parameters are set prior to the model run (Table 1). There are 1,000 civilian agents in the model because a literature review of earlier models indicated that 1,000 agents were a computationally acceptable number. The number of cops, while inflated compared to typical police to citizen ratios, was picked to increase the chances of a cop being present at some convergences across the 16,035 potential places a crime could occur. The unemployment rate of six percent is based on the 2002 unemployment rate for Seattle (Bureau of Labor Statistics 2003). Given that 20% of the population has committed a crime, 20% of citizens are assigned criminal propensity (Visher and Roth 1986). To ensure that each agent has an equal chance of being selected, a uniform distribution is used to select the agents who are assigned criminal propensity and a status of unemployed. Civilians who commit a robbery must wait at least one hour before they can re-offend. Again, since this is the initial implementation, a very short interval was chosen to ensure there would be convergences among the elements necessary for a crime.

Where characteristics are assigned to agents, the distributional aspects of the values are also specified. For example, in this initial implementation the time spent away from home is assigned based on a normal distribution with a mean of 432 minutes (for the 30% condition) and a standard deviation of 10% of the mean \((sd = 43)\). Initial wealth is assigned with a mean of 50 and a standard deviation of 20 units. This yields a classic normal curve but does not accurately represent the wealth distribution. The primary goal in the initial model is simplicity but subsequent tests of the model should attempt to better represent actual wealth distributions in Seattle (Axelrod 2006). The same drawback exists for using a static amount of wealth received each payday and for assuming that the same amount of wealth is taken in each street robbery. However, these parameters provide a starting point from which to build more complex, and more representative models. There is no empirical basis for the initial wealth, amount of payday wealth or the wealth exchanged during a street robbery. The two parameters that are important to the decision to commit a robbery, guardianship and suitability of target will be described in detail in the section on agent behavior. Both values are used to condition the situational characteristics of a situation and in doing so, represent the perception of the offender.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Society Level</strong></td>
<td></td>
</tr>
<tr>
<td>Number of Agents = 1,000</td>
<td>Represents a balance between ensuring there are enough agents so that interactions can occur and the computational overhead from using more agents.</td>
</tr>
<tr>
<td>Number of Police = 200</td>
<td>Chosen to ensure that police agents would be present at some of the convergences that occur across the 16,035 places in Seattle.</td>
</tr>
<tr>
<td>Unemployment Rate = 6%</td>
<td>The unemployment rate of six percent is based on the 2002 unemployment rate for Seattle (Bureau of Labor Statistics 2003).</td>
</tr>
<tr>
<td>Rate of Criminal Propensity = 20%</td>
<td>Given that 20% of the population has committed a crime, 20% of civilians are assigned criminal propensity using a uniform distribution (Visher and Roth 1986).</td>
</tr>
<tr>
<td>Time To ReOffend = 60</td>
<td>Parameter value chosen as a starting point since the author could find no empirical data on which to base time to reoffend.</td>
</tr>
<tr>
<td>Random Number Seed = 100</td>
<td>An explicit random number seed based on the Mersenne Twister (MT) algorithm is used as the basis for all random number distributions used in the model. MT is currently considered to be the most robust in the industry (Ropella et al. 2002).</td>
</tr>
<tr>
<td><strong>Agent Level</strong></td>
<td></td>
</tr>
<tr>
<td>Societal Time Spent Away From Home = 30%</td>
<td>Assigned based on a normal distribution with a mean of 432 minutes (for the 30% condition) and a standard deviation of 10% of the mean (sd = 43).</td>
</tr>
<tr>
<td>Initial Wealth = 50</td>
<td>Initial wealth is assigned with a mean of 50 and a standard deviation of 20 units.</td>
</tr>
<tr>
<td>Amount of wealth received each payday = 5</td>
<td>No empirical evidence available.</td>
</tr>
<tr>
<td>Amount of wealth exchanged during robbery = 1</td>
<td>No empirical evidence available.</td>
</tr>
<tr>
<td><strong>Situation Level</strong></td>
<td></td>
</tr>
<tr>
<td>Guardianship Perception = U(−2,2)</td>
<td>The guardianship perception value can add or subtract zero, one or two guardians from the actual number present. This represents the stochastic element in the offender’s perception of the willingness of a guardian to intervene.</td>
</tr>
<tr>
<td>Suitable Target Perception = U(−1,1)</td>
<td>The value in suitable target can increase or decrease the suitability or leave it unchanged. This enables the offender to sometimes decide a target is not suitable even when they have more wealth.</td>
</tr>
</tbody>
</table>
Another type of parameter used in the model is random number seeds. An explicit random number seed is used as the basis for all random number distributions used in the model. The random number seed is based on the Mersenne Twister algorithm which is currently considered to be the most robust in the industry (Ropella et al. 2002). This provides the ability to replicate the model behavior over subsequent runs and is essential to using simulation as a laboratory for experimentation (Axelrod 2006).

In a simulation model, the modeler controls the data that are collected and how frequently they are written to a file. There is a computational cost each time the program writes to a file that must be balanced with the need for information about the model run. Here the outcome data from the simulation are collected daily for individual civilians and for society as a whole but only at the completion of each model run for street nodes/places. These data are written to two types of files, text files and shapefiles.

3.3 Movement in the Model

Two types of agent movement are implemented in the model, random and directed (i.e. among a predefined set of locations). Each requires different strategies for implementation but both rely on nodes rather than lines within Agent Analyst. This is necessary because Agent Analyst does not support connections to a geodatabase or a network dataset so there can be no dynamic routing of directed agent travel in the model. The alternative strategy implemented here uses GIS functionality to identify the shortest paths among activity places, converts each path into a set of nodes, and then reads each set of nodes into Agent Analyst as an activity space. This strategy requires that the set of nodes representing street/grid intersections be created from the street file and then used to represent the places among which citizens travel and at which a crime can occur. The use of point locations to represent places among which agents can travel enables both predefined activity spaces and dynamic random movement in the model.

Directed movement in the model is the more complex type of movement and requires the definition of activity spaces for the agents. Activity spaces consist of four activity nodes and a list of path nodes. The list of path nodes describes the complete set of nodes to be traversed to visit all four activity nodes. Movement takes place from street intersection/node to a connected street intersection/node (hereafter referred to as street nodes). Since routing in ArcGIS uses the streets, identifying the street nodes that are traversed in the course of visiting all four activity nodes required the creation of a custom program. The output of the program provides a list of street nodes that are traversed while traveling the shortest path among the activity nodes. Civilian agents in the model always travel among their activity nodes in the same order each day (i.e. home, main, activity one, activity two, home). As seen in the left panel of Figure 3, directed agent movement occurs from node to node along a pre-defined path. In the example, the agent starts at home (node 107). From there the agent moves to 110, 122, 124, and so on.

In addition to the directed movement by civilians going about their daily activities, dynamic random movement is also implemented. Random movement is used by the police agents in all three versions and by the civilians in two versions of the model. Random movement is implemented using the same two-step process for both the grid and the street landscapes; identification of neighboring nodes and random selection of the target node. As part of the preprocessing of data done before running the model, a set of adjacent nodes is identified for each node and written to a file. The creation of
this file is achieved through a series of topological queries (i.e. select node, select streets/grid lines adjacent to selected node, select nodes that intersect selected streets/grid lines). In this way a file of neighboring nodes is created for each travel node and then used as the basis for random movement.

When traveling in a random fashion, the agents follow a ‘random walk’ where the agents move one randomly chosen node each minute of the model (Chaitin 1990). During the model run, the file of node neighbors is used by each agent as they travel. When an agent stops at a node, the list of neighboring nodes for that street node is checked and one node is randomly chosen from the list. The agent is then associated with the new node. This process is repeated for each agent who is traveling at each minute of the model. The right panel of Figure 3 shows a simplified travel movement. The agent is at node 134 and could potentially move to any of the following nodes: 102, 120, 121 or 133. A uniform random number is generated giving each node an equal chance of being selected. The agent then moves to the selected node and the cycle repeats.

3.4 Creating Activity Spaces

As discussed earlier, theory from both geography and criminology holds that the travel behavior of individuals is influenced by the street network and the locations at which opportunities for employment, recreation, retail and services exist. The structure of individual travel, in turn, plays a large role in the convergence of offenders, targets, and guardians. An individual’s temporal schedule (i.e. the amount of time spent at each activity) is impacted by the distances between the activities and the speed of travel. The more time spent traveling the less is available to spend at an activity. Activity spaces consist of nodes (places) and paths (list of places traversed).

The complete process of developing agent activity spaces is detailed in Figure 4 which also describes the entire data flow from input through output. The first stage, uses GIS to link the locations of the street nodes to the polygon layers; this process
associates a particular node with the area in which it falls (e.g. street node 1 is in blockgroup 201). The outcome from the process is a list of nodes and their associated blockgroup number. In stage 2, the distribution of homes, jobs and retail/service/recreation activities across Seattle is calculated. These distributions are then used to assign agent homes, jobs and activities in the same proportion as they are found in Seattle (e.g. if 10% of the population lives in a particular blockgroup then 10% of the agents are assigned to that blockgroup). This process produces two files. One file contains the activity node number and the blockgroup in which it is located. The other file contains the blockgroup and number of agents to be assigned a home node from that blockgroup. The same basic methodology is then repeated to assign work places and activities.

Stage 3 uses the two files just described in a Java program that randomly selects and assigns agent homes, work places and activities in the same proportion as they are found in Seattle. Four activity nodes are selected representing a home node, main node (e.g. work, school, etc.) and two additional activity nodes (e.g. a retail store, gym, coffee shop). Two thousand path files are written out; one for each agent when employed and another for each agent when unemployed.

The final stage in creating directed movement paths involves finding the shortest path among the nodes. The shortest path among the activity nodes is calculated using ArcGIS Network Analyst via a custom Visual Basic program that generates a list of nodes that are traversed while traveling the shortest path and writes them out to agent path files. Two paths are created for each agent; one describes their activity space when employed and the other when unemployed. Three of the four nodes remain the same between the two activity spaces, home, activity node 1 and activity node 2; only the main node changes. When employed, the agents’ main node is assigned in the same proportion as employment; when unemployed it is assigned from the distribution of activities in Seattle. The 4,000 output files describing the activity nodes (N = 2,000) and

Figure 4 Data inputs/outputs related to agent activity and outcome data
activity paths (N = 2,000) for each agent are then ready to be used to define directed civilian agent movement in the model.

At this point, a temporal schedule is assigned within the street model using the following steps. First, the time spent at home is randomly assigned to each agent so that the societal average time spent away from home meets a preset value. Next the number of nodes traversed is counted and subtracted from the total time away from home. The larger the geographic extent of an individual’s activity nodes the greater the time required to travel among them. The remaining time is randomly allocated to the Main, Activity 1 and Activity 2. Upon completion, each agent has two potential activity spaces consisting of four nodes, times to spend at those nodes, and the path to travel among them.

3.5 People in the Model

Regardless of version, two agent classes operationalize people in the model, civilian and police. The civilian class represents the general population of Seattle. The three roles that agents can take on during a crime event are encompassed in the civilian agent class; civilians can be offenders, targets, or agents of informal guardianship. The particular role a civilian agent takes is driven by their characteristics and the contextual dynamics of the specific interaction. Police are the agents of formal guardianship. Both police and civilian agents are assigned a type of movement that is constant over a particular model run. Only civilians have additional attributes that are used in the model. Police are described first because of their relatively simple role and characteristics.

Police agents have only one role, that of a formal guardian. In the model, the presence of a police agent prevents a crime from occurring. At the start of the simulation, police agents are randomly distributed across the nodes. To accomplish their mission of crime prevention police agents follow a random movement pattern in which they move one node at a time and only to an adjacent node. Police never commit crimes in this model and they are never targets.

Civilian agents in the model are assigned three characteristics that are integral to the decision to commit a street robbery: (1) a time to spend at home; (2) criminal propensity; and (3) wealth. In the Street Directed version agents also have an employment status. A final characteristic of agents is their mode of movement and accompanying temporal schedule which vary by version of the model; random or spatio-temporal (these were discussed in the previous section). All civilian agents are assigned a time to spend at home that is static over a model run. Civilians in the Street Directed version have assigned places where their activities occur and assigned times to be at those places. When not at home, civilians in the Grid Random and Street Random versions travel randomly and those in the Street Directed version follow the shortest path among their activity nodes.

Criminal propensity is used to differentiate agents who evaluate situations and make the decision to offend from other agents in the model. In all other ways, civilians with criminal propensity are exactly the same as those without. While only agents with criminal propensity can make the decision to offend, it is the particular constellation of individual and situational factors that determines whether a crime is committed. In this way, patterns of offending and victimization are allowed to emerge from decisions made by individuals in particular contexts.

The characteristic of employment status is added to civilian agents in the Street Directed version of the model. This characteristic has two important impacts in those versions of the model. First, it changes the amount of time spent at the three activity
nodes (but not the overall time spent away from home). Incorporating employment status enables the model to reflect its influence on the temporal and spatial aspects of routine activity schedules. Second, the attribute impacts the wealth of the civilian agents. Employment and wealth are linked in the model. Those who are employed receive regular but static infusion of wealth every two weeks over the model year. Civilians who are unemployed do not get paid. Every month, 3% of unemployed agents become employed and are replaced by a new random selection of employed agents who become unemployed. It is important to note that the employment status is assigned independently of the criminal propensity indicator; civilians with criminal propensity can be employed in the model, as they are in life.

3.6 Agent Behavior

Major actions at each minute/tick of the model are as follows. Only those nodes with at least one agent present (i.e. an active node) are evaluated. Active nodes meeting the following criteria continue to be evaluated: (1) no police present; (2) at least two civilians present; and (3) at least one of the civilians has criminal propensity. If there is a cop at the node, the situation is not evaluated further. In this way, the presence of formal guardians always deters the commission of a crime. Then an active offender agent is selected from the agents at the active node. If there is only one offender at the node, they automatically become the active offender. Otherwise, the active offender is randomly selected from the list of agents with criminal propensity who are at the node. Random selection is necessary to ensure the same agent is not selected to be active each time the model is run. Offender agents who are not selected are at risk of becoming victims.

Each active offender agent considers the following aspects of a situation at their node: (1) level of guardianship; and (2) existence of a suitable target. Informal guardianship is evaluated by counting the number of other citizen agents at the node and subtracting two. This is necessary since neither the potential victim nor the offender can act as an informal guardian. An offender’s perception of the effectiveness of the particular informal guardians present at a place is incorporated into the formula through the error term \( P \) which is drawn from a uniform random distribution that ranges from \(-2\) to \(2\). The error term either adds or subtracts as many as two informal guardians from the guardianship determination. The decision regarding guardianship is represented as:

\[
G = (N_A - 2 + P_G)
\]

where \( G \) = guardianship, \( N_A \) = number of agents at node, and \( P_G \) = perception of capability of guardians who are present (uniform random number between \(-2\) and \(2\)).

In reality, the perception of the presence of capable guardians is most likely evaluated along a continuum using a variety of criteria including presence of: place managers (security guards, parking lot attendants, etc.), number of other people on the street, characteristics of neighborhood, etc. By incorporating a stochastic element, \( P \), in the potential offender’s decision-making process in situations where there is some informal guardianship the model is able to more realistically represent uncertainty in how guardianship is evaluated.
Next, the active offender agent considers whether there are suitable targets at the node. Prior to the calculation of a suitability value the following actions are taken. First, the status of each citizen agent is checked. Only the wealth levels of citizen agent’s who are at risk of being a victim of street robbery are evaluated. Second, the agent with the highest wealth is chosen from the subset of civilians who have more wealth than the active offender and evaluated using the following formula:

\[ S = (W_T) - (W_A) + P_S \]  

(2)

where \( S \) = perceived suitableness of target, \( W_T \) = wealth of target, \( W_A \) = wealth of active agent (potential offender), and \( P_S \) = offender perception of target suitability (uniform random value between \(-1\) and \(1\)). The error term \( P \) represents the uncertainty inherent in judging the relative suitableness of a target; its value is randomly assigned using a uniform distribution between \(-1\) and \(1\), \( U(-1, 1) \).

Returning to the agent decision scenario, if \( G < 1 \) and \( S = True \), then there is both a suitable target and a lack of capable guardians so the decision is to rob the suitable target that was identified. If \( G = 1 \) and \( S = True \), then the level of guardianship is in a gray area and the decision is made randomly.

If a crime occurs, the wealth level of the offender is increased by one unit and the wealth level of the victim is decreased by one unit. Both the citizen as offender and the citizen as victim have the appropriate counters representing offending and victimization, increased by one. In addition, offender agents must wait an hour before they can re-offend so their time counter is started. Once each citizen with criminal propensity evaluates their situation, all agents move and the decision structure is repeated.

3.6 Comparing Three Versions of Activity Space and Movement

To test the impact of different conceptualizations of activity spaces and movement on street robbery patterns requires the implementation of three versions of the base street robbery model. Two of the versions are identical except for the landscape on which the model is run (Table 2). In both the Grid Random and Street Random versions civilian agents are assigned only a time to spend at home. All agents are randomly distributed at the start of the simulation and then travel randomly until the end of the day. Their

<table>
<thead>
<tr>
<th>Civilian Movement</th>
<th>Grid Random</th>
<th>Street Random</th>
<th>Street Directed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police Movement</td>
<td>Random</td>
<td>Random</td>
<td>Defined Activity Space</td>
</tr>
<tr>
<td>Civilian Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criminal Propensity</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wealth</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Activity Space</td>
<td>No</td>
<td>No</td>
<td>Spatio-temporal</td>
</tr>
<tr>
<td>Multi-faceted Risk Status</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Employment Status</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2 Implementation versions of the conceptual street robbery model
next day begins at the node where the previous one ended. Since they are at risk of being robbed whenever they are not at home, civilians in the Random versions of the model have the highest level of risk.

The third version, Street Directed, uses the pre-defined activity places and routes that reflect the distribution of activity places and the road network in Seattle to incorporate a more complex notion of activity space and risk. In this version of the model civilian agents are not at risk when they are at home or at work; only when they are at other activities or traveling. This representation of risk is in keeping with the crime being studied. By definition, street robbery happens only on the street or in public places; not in a home or inside a workplace.

4 Results of Implementing the Model

This successful implementation of a theoretically-based, geographically-aware model of street robbery capitalizes on the recent development of Agent Analyst which allows a simulation model to be ‘situated’ on an extant rather than an artificial landscape and in doing so provides a more realistic context to the model behavior. However, the effect of the ‘situating’ simulation is an open question. This research quantifies the impact of realistic landscapes by comparing the differences in the incidence and pattern of street robberies when the same model is run on a uniform grid versus an actual street network. Finally, the effect of distributing agents and their activities based on the distribution of population, jobs, and retail is then tested via a third version of the model.

Beginning with the comparison of the same model on a grid landscape versus a street network, the results clearly demonstrate that the street network itself has a measurable effect on the number and pattern of street robberies. When situated on a street network, society experiences roughly 10% more convergences and robberies than when it exists on a uniform grid (Table 3). This outcome is most likely due to the funneling effect of the street network on human activity; it increases the number of times people converge. In addition, the patterns of street robberies produced by the two landscapes are markedly different. The crime pattern generated from the grid landscape is very dispersed with about 94% of all nodes experiencing at least one street robbery as compared to the only about 83% of street nodes. Kernel density maps confirm that street robberies are more dispersed under the Grid Random version than the Street Random version and that the spatial pattern to a large extent reflects the underlying street pattern (Figure 5). Where the street network is denser, the likelihood of people converging and a crime happening is higher.

A comparison of the Street Random version to the Street Directed version reveals even more striking differences. This comparison specifically examines the effect of considering the distribution of opportunities via routine activity spaces on street robbery. In the Street Directed travel version, the number of convergences increases by 27% (Table 3). Agents are coming into contact more frequently because their activities reflect the clustered nature of opportunities and because those activities are routine rather than random. However, both the number of street robberies and percentage of civilians who are robbed decrease in the Street Directed version. This seeming anomaly is due to the following process, as citizens with criminal propensity get more wealth than those without, fewer crimes occur because there are fewer suitable targets.

There are also significant differences in the spatial pattern of street robberies between the two versions. The Street Directed version of the model produces more
Table 3  Societal-level model outcomes

<table>
<thead>
<tr>
<th></th>
<th>Grid Random</th>
<th>Street Random</th>
<th>Street Directed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEOPLE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time spent away from home* (minutes)</td>
<td>432.3</td>
<td>432.2</td>
<td>431.1</td>
</tr>
<tr>
<td>Total Robberies*</td>
<td>49,819</td>
<td>54,374</td>
<td>30,524</td>
</tr>
<tr>
<td>Total Convergences*</td>
<td>1,298,080</td>
<td>1,447,812.6</td>
<td>1,836,638</td>
</tr>
<tr>
<td>Total Robberies Deterred by Police*</td>
<td>1,147.4</td>
<td>1,577.2</td>
<td>1,066.4</td>
</tr>
<tr>
<td>Percentage of civilians who were robbed</td>
<td>78.0%</td>
<td>78.3%</td>
<td>72.9%</td>
</tr>
<tr>
<td>Percentage of civilians who were repeat victims of street robbery</td>
<td>66.2%</td>
<td>65.9%</td>
<td>62.8%</td>
</tr>
<tr>
<td><strong>PLACES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of nodes with a robbery*</td>
<td>93.9%</td>
<td>82.8%</td>
<td>9.0%</td>
</tr>
<tr>
<td>Percentage of nodes with repeat robberies*</td>
<td>79%</td>
<td>68.2%</td>
<td>6.9%</td>
</tr>
</tbody>
</table>

* Averaged across five different random number seeds.

Figure 5  Comparison of street robbery outcome patterns across model versions
clustered crime patterns. Much smaller proportions of nodes have a single robbery (9 versus 83%) or multiple robberies (7 versus 79%) in the Street Directed version versus the Street Random version (Table 3). Once again, the kernel density maps confirm the change in the pattern of clustering between versions. The pattern from the Street Directed version is extremely clustered along high volume pathways that represent areas of greatest routine activity. Both the changes in incidence and pattern revealed by the results are in line with empirical findings related to the characteristics of crime patterns.

Previous research has clearly established that crime is clustered rather than dispersed in space (Sherman et al. 1989, Eck et al. 2000, Weisburd et al. 2004). Violent crime tends to be even more concentrated in even fewer areas. These empirical patterns reflect the uneven distributions of opportunities for transportation, housing, recreation, and employment. Given the above findings, the outcome of the Street Directed version of the model most closely matches what we would expect based on empirical data. The other versions also produce plausible results. That is we would expect random movement on a uniform grid to produce a more dispersed pattern of crime than random movement on a street network just by virtue of varying densities of streets in different parts of the city. Finally, a landscape that clusters opportunities and structures human activity on a street network would be expected to produce the most clustered and thus most realistic pattern of street robbery events. Thus the results of the model confirm the importance of ‘situating’ simulation when modeling crime events.

4.1 Limitations of the Implementation

Although this research provides significant advances, there are some limitations to the current model and to simulation modeling more generally that deserve mention. There are two drawbacks to the current implementation of random movement. First, the random selection of the next node is accomplished from a list that only considers adjacent nodes. Thus, agents are limited to moving only one node per minute. Second, there is no prohibition against backtracking by agents. The same node that an agent just came from is included in the list of adjacent nodes for the current node which means it can be selected as the goal node. Together these implementation decisions may lead to smaller, less realistic activity spaces for agents that are moving randomly. Future implementations should consider giving agents who are moving randomly, the same ability to move more than one node as the agents who have directed movement.

The choice of parameter values is a critical aspect of all models that deserves special attention because it impacts the external validity of the model. Parameter validity is the measure of how well the parameter values used in the model matched reality (Carley 1996). As noted in the data section, every attempt is made to use realistic model parameter values. However, in some cases there was no evidence available and in others a simplified representation was chosen to establish a baseline (e.g. wealth distribution) (Epstein and Axtell 1996, Axelrod 2006). In these cases, the validity of the parameters is unconfirmed and their impact on the model results needs to be thoroughly investigated.

5 Conclusions

This article documents the implementation and testing of a new method for ‘situating’ simulation within Seattle, Washington. It makes use of a recently released software
package that integrates GIS and ABM enabling simulated agents to be ‘situated’ within an empirical context. A methodology for enabling agent movement on a street network and creating geographically informed activity spaces is detailed and provides the foundation for the development of three versions of a model of street robbery events. This new procedure offers three advantages over the current practice. First, the method allows agents to move along an existing street network rather than being restricted to an abstract grid space or to moving along linked grids to mimic a street network. Second, researchers can use a street network directly without having to convert the network to a grid for use in a simulation model. Third, it demonstrates how the land use patterns of a city can be used to inform the spatial behavior of agents.

To examine the effect of using ‘real’ landscapes as compared to synthetic ones, the same model is implemented on a uniform grid landscape and on a real street network. Further, the results of those runs are compared with a more realistic model that incorporates the distribution of opportunities and the notion of routine activity spaces. The outcomes from the three models clearly demonstrate the importance of ‘situating’ simulation. The incidence and pattern of street robbery events generated when the same model is run using a uniform grid landscape are significantly different from those generated from the street network landscape. The street network landscape concentrates the travel of agents which increases convergences (i.e. the intersection of at least one civilian with criminal propensity and one other ‘at risk’ civilian at the same place and time) and consequently increases the potential for street robbery to occur. Additional evidence of the concentration of activity due to the street network is revealed by kernel density maps which show increased clustering of street robbery from the Street Random version. Finally, when routine activity spaces that incorporate the distribution of opportunities for housing, employment and recreation are used to structure agent travel the pattern of street robbery becomes significantly more clustered. Although this effort implements only simplified versions of activity spaces, the methodology developed provides a guide for future research extensions.

The results of this initial implementation of a street robbery model have important implications for the use of simulation to elaborate theory (Albrecht 2005, Eck 2005) and to conduct experiments (Schultz and Sullivan 1972). Previous attempts to test routine activity theory, although generally supportive, have produced mixed results (Messner and Blau 1987, Kennedy and Forde 1990, Sampson and Lauritsen 1990, Miethe Terance and McDowall 1993). None of those tests were able to sufficiently address the spatio-temporal structure of routine activities, satisfactorily deal with measurement issues or effectively capture the dynamic nature of interactions at the micro level. The methodology and model implemented here address all three of those issues. In addition, agent-based software allows single aspects of the behaviors of agents to be manipulated while all others are held constant. This and other recent research illustrate how these capabilities can be used to conduct controlled experiments to test the core axioms of routine activity theory (Groff 2007b) and to separate the effects of space and time on the convergence of the elements necessary for a street robbery to occur (Groff 2007a, c).

On the whole, the methodology is relatively straightforward and the results demonstrate the importance of including more realistic representations of landscapes in agent-based models but there is room for further enrichment. In the case of agent movement, one enhancement would be to take into account barriers. For example, in the case of police, they typically are assigned to patrol within designated areas. Subsequent implementations could make use of those types of barriers and in doing so provide a step
toward more realistic police agent behavior. Barriers are also important to civilians and may be physical (e.g. streams, limited access highways) and/or perceptual (e.g. edges of neighborhoods, etc.). Another enhancement would be to use speed limits and one-way streets in the development of the routes among activity places.

The future use of this methodology is both facilitated and limited by the software packages available for implementation. Currently, Agent Analyst offers a unique opportunity and some challenges to ‘situating simulation’. Since it is still under development, its ultimate form is still evolving so all these statements apply to the beta version only. On the opportunity side of the ledger, Agent Analyst offers the most straightforward option for non-programmers who are interested in developing their own spatially-aware models. It frees the new programmer from the numerous details involved in developing the model framework and learning the syntax of Java. On the challenges side, the most serious drawback is that the current version is only able to read shapefiles, not the network datasets that would enable dynamic routing. In addition, the debugging tools are extremely limited. Finally, would-be modelers must become familiar with a unique subset of Python syntax and any Java classes that are used. Addressing these issues would decrease the difficulty of programming models, speed development time, and increase the realism of agent travel and activity spaces in the models.

Acknowledgements

Thanks to Mary Jo Fraley, Jochen Albrecht, and Nick Collier for their suggestions regarding the technical aspects of the implementation of the model. Ned Levine and the anonymous reviewers provided illuminating comments on earlier drafts of this paper. This research was supported in part by Grant 2005-IJ-CX-0015 from the National Institute of Justice.

Notes

1 Comprehensive data describing both the activity type and location are not collected. Surveys of small samples of a population over short time periods are used in transportation and accessibility modeling (Kwan 1998, Kwan et al. 2003) but are not adequate for this approach. The need for data describing the activities of individuals is noted in the literature that emphasizes the opportunity perspective on crime (e.g. Meier et al. 2001, Sampson and Lauritsen 1995, Wilcox et al. 2003).

2 Simulation modeling is widely used in a variety of disciplines. One mature example is in economics (Schelling 1971) and now more recently in criminology where the goal is to allow the exploration and elaboration of theory (Eck 2005). Simulation is also widely used in travel demand modeling (see for example – the TRansportation ANalysis and SIMulation System (TRANSIMS) application at http://tmip.fhwa.dot.gov/transims/). Although based on real data, the goal of these efforts is to accurately forecast travel patterns rather than the generative perspective underlying this research (following Epstein and Axtell 1996).

3 For a differing view please see Wu (1999) who makes a nice argument for the utility of GIS-based simulation that employs cellular automata rather than ABM.

4 Space constraints prohibit a detailed examination of extensions to routine activity theory or even related opportunity theories pertinent to micro level modeling. Rational choice theory is addressed because it provides for bounded rationality in the decision to offend. Several books offer a good overview of opportunity theories (e.g. Cullen and Agnew 1999, Akers 2000, Vold et al. 2002).
This simplistic view of human spatial behavior does not consider the role of trip purpose in determining timing or mode of travel. Mode of travel is an important determinant of exposure (i.e. risk of victimization) and should be incorporated in future models.

This research focuses on using GIS data about real places to inform agent behavior in the model. For another example of empirical data as an input to an agent-based model see Heppenstall et al. (2005).

Agent Analyst was developed from a partnership between ESRI and Argonne National Laboratories. Agent Analyst is free and available for download at http://www.institute.redlands.edu/agentanalyst/. Documentation and software code for the model versions discussed here is also available at that site.

Employment data are from the 2002 County Business Patterns dataset. Employment data are not available by blockgroup, only by zip code. There far fewer zip codes ($n = 56$) than blockgroups ($n = 570$) in Seattle which means the employment data is less precise than the blockgroup data. Specifically, the areal units to which jobs are aggregated are much larger than blockgroups and thus the potential for allocating agents in a way that is not reflected by the actual distribution of employment is higher. This is a minor limitation since the goal is to distribute job locations of agents proportionately, not exactly.

ESRI Biz data 2003 contain a total of 18,024 services or retail establishments in Seattle.

There are two types of agents in the RepastPy GIS Model, vector and generic. Vector agents are associated with a shapefile and can thus be displayed on a map while generic agents cannot.

While research has shown that the shortest topological path is frequently not the fastest path, it offers a standardized and convenient heuristic for this initial model. The custom program was created in Visual Basic and added to the ArcGIS 9.1 session to identify the street nodes traversed by each agent. The author gratefully acknowledges the assistance of Mary Jo Fraley who wrote the code and is making it available via the author.

A python script is used in ArcGIS to automate the identification of the neighbor nodes for each street node and to write those node IDs out to a file.

Please see Groff (2007b) for more information of the use of this model to test routine activity theory by systematically varying the time spent away from home over a set of experiments.

Because of the large size of agent activity spaces, the number of nodes traveled per turn is determined via a random normal distribution (mean = 6, sd = 1). This solution more closely approximates travel since people rarely stop at each street intersection.

Other empirical research has suggested that offender spatial behavior has both criminal and non-criminal components (Rengert 1980, Brantingham and Brantingham 1991, Ratcliffe 2006). However, this research focuses on creating a model based on routine activity theory and thus follows its assumptions.

For a more in-depth discussion of tests of the parameter values please see Groff (2007a, b) who found no change in the overall relationship between time spent away from home and the number of street robberies even when five of the parameters were systematically varied and five different random seeds were used.

References


Brantingham P L and Brantingham P J 1993 Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. Journal of Environmental Psychology 13: 3–28
Chaitin G 1990 Information, Randomness and Incompleteness. Singapore, World Scientific
Chorley R and Haggett P (eds) 1967 Models in Geography. London, Methuen
Cullen F T and Agnew R (eds) 1999 Criminological Theory: Past to Present. Los Angeles, CA, Roxbury Publishing Company
ESRI 2005 ArcGIS 9.1. Redlands, CA, Environmental Systems Research Institute


Harvey D 1969 Explanation in Geography. London, Edward Arnold


Lynch K 1960 The Image of the City. Cambridge, MA, MIT Press


U.S. Census Bureau 2000 *Census 2000: Summary Tape File 1 (SF1)*. Washington, DC, U.S. Census Bureau


