Optimising an Agent-Based Model to Explore the Behaviour of Simulated Burglars

Alison Heppenstall [1], Nick Malleson[1], Linda See[2,3] and Andy Evans[1]

[1]School of Geography, University of Leeds, LS21 1HB

Tel. (+44 113 343-3330) Fax (+44 113 343-3308)

n.malleson@leeds.ac.uk, http://www.geog.leeds.ac.uk/people/n.malleson

[2]International Institute of Applied Systems Analysis, Laxenburg, Austria

[3]Centre for Applied Spatial Analysis, University College London (UCL), London, UK

KEYWORDS: agent-based model, crime, behaviour, optimisation

1. Introduction

Social systems are incredibly complex due to the large number of interacting elements and many underlying processes that are simply not understood. Moreover, these processes are generally non-linear such that small changes in system parameters can have large effects on the outcomes of the system as a whole. Agent-based models (ABMs) have been developed as one technique for modelling complex systems where the individuals or 'agents' of the system are explicitly represented. Agents are independent entities that are capable of interacting with each other and with their environment. The agents make assessments of their situation over time (or during each iteration of the model) and then make decisions in response to these assessments (Bonabeau, 2002). By providing realistic environments and rules that are based on observed and expected patterns of human behaviour, it is possible to create models that can simulate real world systems (Moss and Edmonds, 2005).

Although ABMs represent a way to capture complexity in social systems, they have issues related to parsimony, i.e. they contain a potentially large number of parameters. Some parameter values can be determined through expert knowledge or can be derived from field measurements or social surveys. However, many others are unknown and therefore require a method to determine their values. This need to calibrate a model is not limited to ABMs and many different methods of search and optimisation are available. However, classical search methods are not effective in finding large numbers of parameters so other methods such as genetic algorithms (GAs) are needed.

GAs are intrinsically suited to optimisation when the fitness landscape is complex, changes over time or has many local optima. Through inherent parallelism, they are able to simultaneously explore numerous potential solutions (Holland, 1992; Mitchell, 1998; Goldberg, 1989). GAs operate as follows: a population is first initialised and the objective functions are then set. The fitness of each individual is assessed and on the basis of this, the fittest in the population are selected for reproduction via crossover. This continues over many generations or iterations until predefined criteria are satisfied, e.g. a certain threshold value for the objective function has been reached. For a more detailed overviews of GAs, the reader is referred to Goldberg (1989), Davis (1991), Michalewicz (1992), Bäck and Schwefel (1993) and Eiben and Smith (2003).

This paper briefly presents the optimisation of an ABM through the application of a GA for exploring the behaviour of simulated burglars.

1. The Agent-Based Burglary Model

The model utilised here attempts to provide a detailed representation of burglary at the city scale that includes a) detailed offender drivers, decision making, and behaviour; b) realistic victim distributions and attributes, including daily variations in household occupancy; and c) a realistic environment including a full transport network and reasonable levels of guardianship. A detailed description of the model design and data preparation is given in Malleson (2010a, b). The experiments presented here were run on the city of Leeds, UK.

The motivation behind the model is to simulate the spatio-temporal locations of burglaries at the city scale and, ultimately, to provide a framework for modelling and testing our understanding of the criminal system. The model runs for a fixed length of simulated time -- sufficient to reach dynamic equilibrium -- so does not predict the actual number of crimes. Instead, we focus here on the values of the behavioural parameters that drive the behaviour of the agents to determine what these tell us about the behaviour of burglars in the real world.

2. Optimising Parameters Using a GA

There are 7 different variables that determine where burglar agents will start searching for targets and which houses, in particular, they will actually victimise (see Table 1). Each variable has a weight associated with it, and it is these weights that will have their value optimised by the GA.

This can help to determine which parameters have the most substantial influence on the model, and the values may eventually inform our understanding of the behaviour of burglars in the real world. For example, the majority of offenders in the area might be less concerned with the distance to travel (*Distance*), but much more concerned with the potential returns (*Atrractiveness*) – the GA will help to illuminate this.

Table 1 The parameters that influence an agent's burglary decision.

Variable	**Description**
Decision where to *start* searching (the individual house to travel to)	
1. Distance	The distance from the agent's current location
2. Attractiveness	The affluence of the target area
3. Social Difference	The social similarity between the target and the agent's home
4. Previous Successes	The number of previous successful burglaries

Decision whether or not to burgle a house as the agent passes it	
5. Community Cohesion	How cohesive the surrounding community appears.
6. Accessibility	How easy the house would be to enter.
7. Visibility	How visible the entrances are to neighbours / passers-by

3. Results

A population of 20 model configurations were used and the GA was tracked over three iterations. Computational requirements limited both the size of the population and the number of iterations – these will be improved in the future. The fitness of each model configuration (or 'chromosome') is provided in Figure 1, which is plotted against the iteration number. The GA is able to identify which model configurations result in the lowest error and, hence, which should be used to generate the configurations in the next iteration. Accordingly the model error decreases with each subsequent iteration. Figure 1 also highlights some clustering of fitness values after the initial (random) population undergoes an evolution. This illustrates that the algorithm is fine-tuning the ABM in different parts of the parameter space that have the lowest error.

Table 3 provides the values for each of the parameters for the models with the lowest error after each iteration. The GA appears to converge very quickly to an optimal configuration, which is found after the first iteration and does not change substantially over the next few iterations, although a marginally different configuration is found in iteration 3. This implies that the algorithm has found a global maximum, or alternatively, the model may need to run many more generations before a further improvement is found.

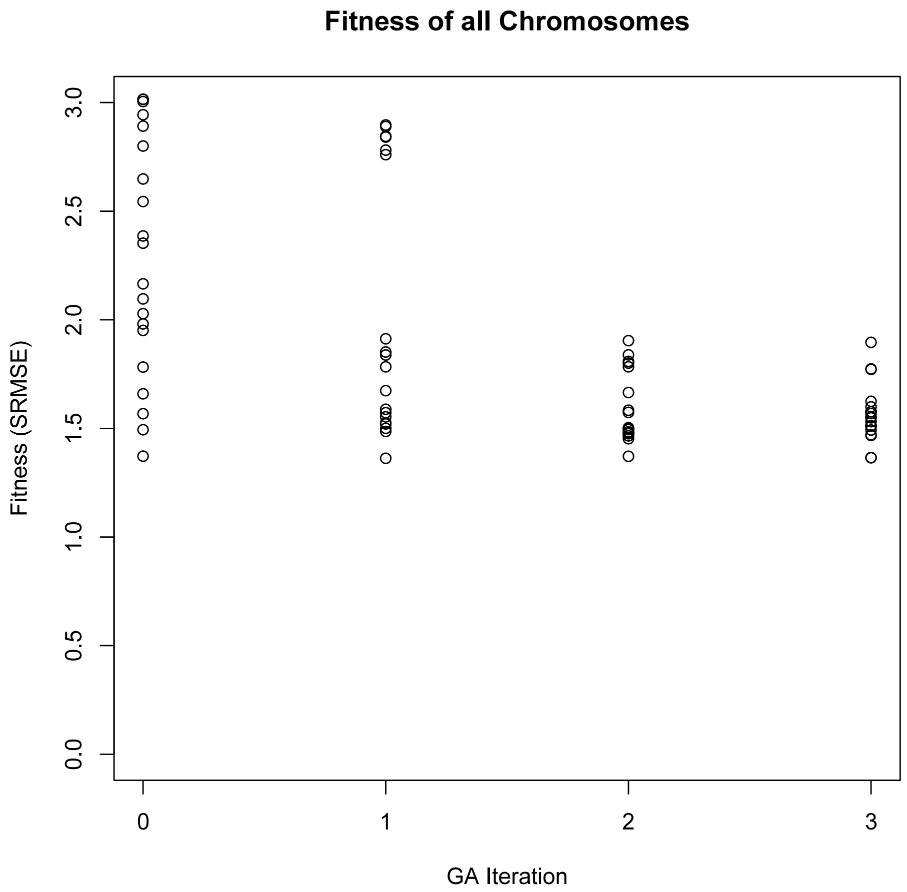

Figure 1. Fitness of all the chromosomes by GA iteration

Table 3. Values of the parameters after each iteration

Iteration	Fitness	w1	w2	w3	w4	w5	w6	w7
0	1.372	0.719	0.668	0.736	0.683	0.541	0.291	0.984
1	1.362	0.719	0.668	0.736	0.683	0.541	0.291	0.984
2	1.372	0.719	0.668	0.736	0.683	0.541	0.291	0.984
3	1.365	0.689	0.717	0.781	0.727	0.510	0.241	1.000

The weights: w1 = distance; w2 = attractiveness; w3 = SocialDifference; w4 = PreviousSuccess;
w5 = CollectiveEfficacy; w6 = Accessibility; w7 = Visibility

To review the results spatially, Figure 2 presents maps of the burglary counts generated by the three best model configurations after GA iterations one ('Model 1), two ('Model 2') and three ('Model 3'). The results are presented in two forms. The maps on the left hand side of Figure 2 present results that have been spatially aggregated to the geography of the communities and the maps on the right hand side present point density estimates produced using the Kernel Density Algorithm (KDE). The use of KDE arguably presents a more accurate picture of the underlying point patterns and is commonly used by police analysts (Chainey and Ratcliffe, 2005).

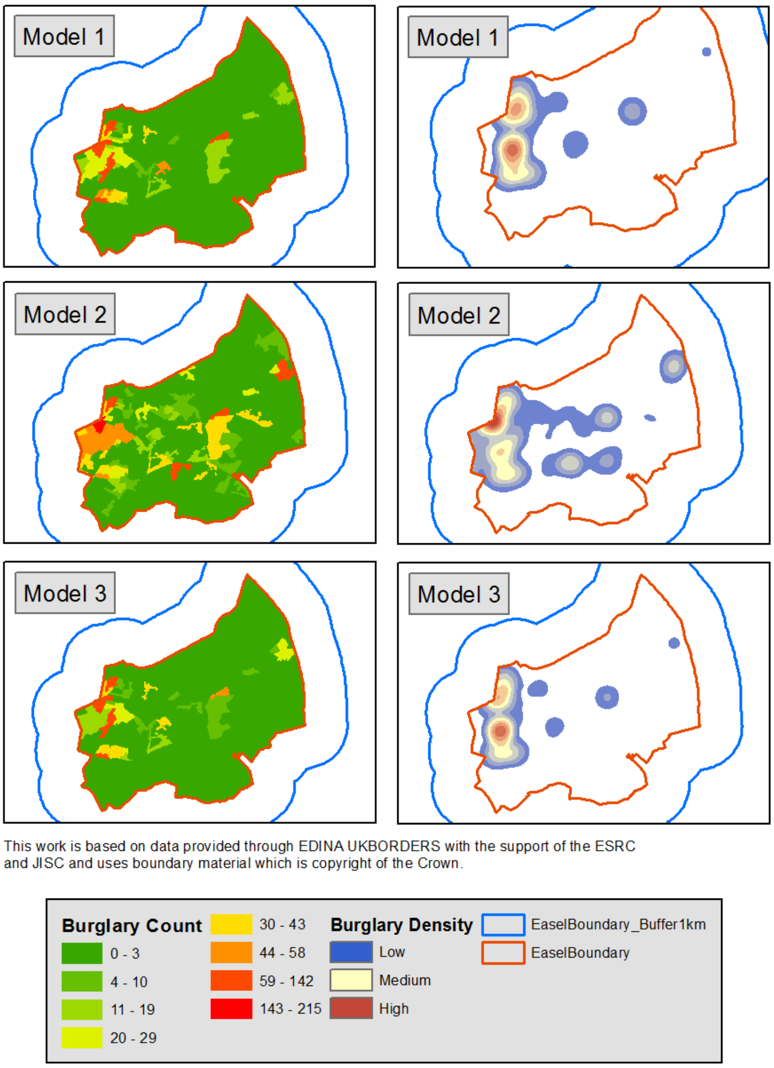

Fig. 2. Results from the three best model iterations after 1, 2 and 3 iterations

The model results show consistently high numbers of burglaries on the western side of the study area, which matches the general pattern exhibited by the observed data. Interestingly, these larger scale patterns are similar regardless of the differences in configuration, which suggests that small changes to any of the agents' behavioural parameters have little effect on the model results. Some small differences can be seen in the centre of the study area where some small hotspots are picked up differentially between the three models. Any discrepancies are likely to be a result of the probabilistic nature of the model, although there would be scope to investigate what might be generating these differences. However, it is encouraging that small parameter changes have little effect on the model results because, were this not the case, it would be more difficult to be confident in the robustness of the results. This represents another considerable advantage of the application of an optimisation algorithm to this model.

5. Discussion

This paper presented some very preliminary attempts at using a GA to estimate the parameters of an ABM of burglary. These initial findings indicated that the weight associated with the visibility parameter was consistently high. This means that with the models that closely matched the observed crime data, the agents were more likely to burgle houses that were well hidden from their neighbours and passersby.

However, one major issue with the GA approach as utilised in this example was the large amount of computational time required to run the model, even with only three iterations per GA run. The ABM itself is extremely computationally expensive. Even after some simplifications from the original configuration (Malleson *et al.*, 2010b), a single model run still required approximately 10 hours to complete on a normal desktop machine.

The implications of using such an approach when scaling up an ABM to a much larger area, with larger numbers of individuals and with a much larger number of parameters is clearly evident from these preliminary experiments. Ambitious ABM projects such as modelling the entire economy of the United States (Farmer and Foley, 2009) will clearly face major computational challenges in the future. But without methods like GAs, the task of parameter estimation would render such modelling approaches infeasible.

References

Bäck, T. & Schwefel, H.-P. (1993) An Overview of Evolutionary Algorithms for Parameter Optimization. Evolutionary Computation, 1(1), 1-23.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences, 99, 7280–7287.

Chainey, S., & Ratcliffe, J. (2005). GIS and Crime Mapping (1st ed.). Chichester: John Wiley and Sons.

Goldberg, D. (1989). Genetic Algorithms: in Search, Optimisation and Machine Learning. Addison Wesley, Crawfordsville.

Eiben, A.E. & Smith, J.E. (2003) Introduction to Evolutionary Computing. Springer.

Farmer, J.D. & Foley, D. (2009). The economy needs agent-based modelling. Nature, 460, 685-686.

Holland, J. (1992) Genetic algorithms. Scientific American, July 1992, 66-72.

Malleson, N., Heppenstall, A. J., Evans, A. J., and See, L. M. (2010a). Evaluating an agent-based model of burglary. Working paper 10/1, School of Geography, University of Leeds, UK. January 2010. [Online]. Retrieved on 16 June 2011 from: http://www.geog.leeds.ac.uk/fileadmin/downloads/school/research/wpapers/10_1.pdf

Malleson, N., Heppenstall, A. J., and See, L. M. (2010b). Crime reduction through simulation: An agent-based model of burglary. Computers, Environment and Urban Systems 34, pp. 236–250.

Michalewicz, M. (1992). Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag.

Mitchell, M. (1998). An Introduction to Genetic Algorithms. MIT Press

Moss, S. and B. Edmonds (2005). Towards good social science.Journal of Artificial Societies and Social Simulation 8(4).

Biography

Dr Nick Malleson is a Lecturer in Geographical Information Systems in the Centre for Spatial Analysis and Policy at the School of Geography, University of Leeds. Dr Malleson's primary research interest is in developing spatial computer models of social phenomena with a focus, in particular, on crime simulation.

Dr Alison Heppenstall is a lecturer in Geocomputation at the School of Geography, University of Leeds. AH has a background in development of novel computational tools, in particular, agent-based modelling. Her current work is focused on integrating agent-based models with microsimulation and behavioural frameworks. She was recently an editor on the book Agent-based Models of Geographical Systems.

Dr Linda See is a Research Scholar at the International Institute of Applied Systems Analysis (IIASA) in Austria where she works with Dr. Steffen Fritz on geo-wiki.org, land cover validation and food security issues within the 'Landspotting', 'FarmSupport' and 'GEOSAF' projects funded by the Austrian Agency for the Promotion of Science and the European Space Agency. She was a recent co-editor on the book Agent-based Models of Geographical Systems.

Dr Andy Evans is a Senior Lecturer in Geocomputation who has worked extensively with agent-based models of socio-economic and ecological systems. His particular interests at the moment are in error propagation and constraint in complex system models. Dr Evans is a member of the Centre for Spatial Analysis and Policy, at the University of Leeds.