Implementing comprehensive offender behaviour in a realistic agent-based model of burglary

Nick Malleson, Linda See, Andy Evans and Alison Heppenstall

Abstract
Explaining and modelling crime patterns is an exercise that has taxed policy-makers, criminologists, social reformers and the police ever since the first crime patterns were recorded. Crime is a particularly difficult phenomenon to model because of its inherent complexity; crime patterns are built up from a multitude of human-human and human-environment micro-interactions that ultimately lead to individual crime events. Commonly used modelling techniques, such as regression, struggle to fully account for the dynamics of the crime system. They work at aggregate scales thereby disregarding important individual-level variation and also struggle to account for the effects of different types of human behaviour. Furthermore, important concepts from environmental criminology – such as individual offender awareness spaces or heterogeneity in offender decision-making – cannot be included directly when working at a resolution above that of the individual.

This research addresses the drawbacks associated with traditional mathematical crime models by building an agent-based simulation with a unique offender behavioural model. Through use of the PECS framework for modelling human behaviour, agents are endowed with needs and motives that drive their behaviour and ultimately lead to the commission of crime. As the model uses real-world environmental data, it can be used to make predictions in existing cities. The paper demonstrates that use of this framework, in combination with an agent-based model, can replicate patterns and trends that are supported by the current theoretical understanding of offending behaviour.

Keywords
agent-based modelling, crime, human behaviour, simulation

1. Introduction
Explaining and modelling crime patterns is an exercise that has taxed policy-makers, criminologists, social reformers and the police ever since the first patterns were recorded. One of the difficulties in simulating this system is its complexity. City-wide crime patterns are an emergent phenomenon; the individual crimes that constitute a city-wide pattern result from a multitude of interactions between people and their environment. Modern criminology theories – such as routine activities theory and crime pattern theory – suggest that to be able to predict the occurrences of individual crime events (and hence the larger patterns) it is imperative to consider the individual people and objects that determine whether or not a crime will occur. Taking the crime of residential burglary as an example, these individuals include the burglar(s) who might commit the crime (along with their personal psychology and motivational state), the victim/house which will be the subject of the burglary (as well as the surrounding urban environment) and any potential guardians or passers-by who might directly or indirectly affect the burglary.
Potentially, the most productive crime model will be the one that is able to account for all of these factors and model them directly at the level of the individual.4 

Because many crime models work at aggregate scales, they omit important individual-level variation. Research has found that burglars choose individual homes based on their individual characteristics; making assumptions about the homogeneity of a community when assessing burglary risk is therefore invalid.5 Furthermore, it has been demonstrated that patterns in street-level crime vary considerably so that aggregate-level modelling of even the smallest areas would hide important information.6,7,8 Moreover, such models face difficulties in accounting for the effects of different types of human behaviour. Important behavioural concepts from environmental criminology – such as individual offender awareness spaces, motivational states and heterogeneity in offender decision-making – can only be included directly when individual people are simulated. However, accurately modelling human behaviour is an extremely difficult objective. Humans exhibit “soft” psychological factors such as seemingly irrational behaviour and complex decision making,9 which are highly challenging to simulate in a computer model.

To address some of these drawbacks, agent-based crime models have started to emerge that simulate the behaviour of individual people or objects and attempt to incorporate criminology theory. Notable sources are the recent book entitled “Artificial Crime Analysis Systems: Using Computer Simulations and Geographic Information Systems”10 and a special issue of the Journal of Experimental Criminology.11 For a full review of recent approaches, the reader is directed to.12 One of the most promising avenues for the improvement of recent approaches is the incorporation of a behavioural framework that allows for rich, dynamic human psychology, producing agent behaviour that closely represents the underlying criminology theory. Current approaches either include only simple behavioural frameworks that allow for limited (or non-existent) dynamic behaviour, or no framework at all.13–15 This research improves upon the current approaches by using an advanced cognitive framework, the PECS model, to create a more realistic model of offending behaviour. The focus of the paper is on the crime of residential burglary and the study area, at present, is the city of Leeds, UK.

The layout of the paper is as follows. Section 2 will explore some of the available cognitive frameworks in detail, ultimately concluding that the PECS framework is the most appropriate for this application. Section 3 will then discuss how PECS has been adapted to create a behavioural model for virtual burglar offender agents. Section 4 will outline how the theoretical model has been constructed using the available crime and environmental data. Section 5 will then compare the simulation results to the real crime data and discuss how closely the results match current theoretical understanding about offender behaviour, followed by conclusions in Section 6.

2. Overview of behavioural frameworks

In agent-based modelling (ABM), an agent’s (or individual’s) architecture determines how the functionality of the agent is organised and how the agent replicates human or biological traits.16 Creating accurate architectures to model human behaviour is one of the most challenging aspects associated with ABMs of social systems. Fortunately, a number of architectures (or cognitive frameworks) have been proposed to address how these traits should be mimicked. Three of the architectures are reviewed here: (i) Beliefs Desires Intentions (BDI); (ii) Behaviour Based Artificial Intelligence (BBAI); and (ii) Physical Conditions, Emotional State, Cognitive Capabilities and Social Status (PECS).

2.1. Beliefs desires intentions (BDI)

The BDI architecture17 is currently the most commonly used architecture and has been used in a wide variety of applications including geo-political conflict simulations,18 air traffic management systems19 and frameworks for models of crime reduction.20,21 The framework consists of three major components: beliefs, desires and intentions. Beliefs represent the agent’s internal knowledge of the world. The agent has a memory of past experiences and the current state of the environment. Desires are all the goals that an agent is trying to achieve. These can include short term goals such as ‘eat food’ or more complex, long term goals such as ‘raise children’. As some goals might be contradictory, intentions represent the most important goals which the agent will try to achieve first. Intentions are sometimes viewed as a subset of goals, while at other times they are viewed as the set of plans to achieve the set of desired goals.16 Goals will change with time depending on external inputs and the agent’s internal state. A level of caution can be integrated into a BDI agent by specifying how eager the agent is to change its intentions.

The process which is used to determine how an agent will react to some input from the environment is termed the “actor loop”. Each action is determined by the use of the BDI architecture, i.e. no action is performed without some form of deliberation.22 Therefore the
Although the BDI architecture has been widely used, it has difficulties. Beliefs, desires and intentions are difficult to observe directly unless viewed in a controlled laboratory setting, which might not relate to real situations. In addition, the use of the three components (beliefs, desires and intentions) is criticised by both classical decision theorists, who criticise them for being overly-complicated, and researchers in sociology, who find them too restrictive. However, the chief problem for crime modellers is that the architecture assumes rational decision making. This makes it difficult to justify because humans rarely meet the requirements of rational choice models.

2.2. Behaviour based artificial intelligence (BBAI)

BBAI is a modular behavioural architecture which has been used in another agent-based model of crime. Originally designed to control autonomous robots, it can also be applied to software agents. The basic structure consists of a number of hierarchical layers of increasing behavioural complexity. All layers act as individual controllers of the agent, and they operate independently and simultaneously. Therefore it is the purpose of a 'suppression mechanism' to determine which layer should have overall control at a particular time.

The advantage with this approach is that the agent can work towards different goals simultaneously and no early decision needs to be made about which goals to pursue. Having separate and autonomous layers also provides robustness (i.e. if a high level fails, the lower behavioural levels will ensure that the agent continues to function) and efficiency (i.e. there are no communication overheads between layers). However, a new layer must be created to implement the basic functionality that would otherwise be provided by the lower layers leading to considerable model complexity and attempts to implement intelligence using BBAI have not proved as successful as alternative, hand-designed systems. Although the trade-off of added complexity for extra robustness is appropriate for implementing simple behaviour in physical robots that may encounter unexpected objects, virtual agents do not need very robust behaviour because the environment is wholly specified by the researcher. The advantages of robustness are counteracted by the difficulties required to implement complex ‘human-like’ behaviour. Moreover, an architecture that is specifically designed to model high levels of human intelligence is more appropriate for socio-economic research and, subsequently, the BBAI architecture is not widely used in social modelling.

2.3. PECS

The PECS architecture is founded on the basis that human behaviour can be modelled by taking into account a person’s physical conditions, emotional states, cognitive capabilities and social status. The authors of the architecture believe that all aspects of human behaviour can be modelled using these components. Personality is incorporated into the agents by adjusting the rate that internal state variables change and also how these changes are reflected in agent behaviour. Using a modified example suggested by Schmidt, consider a person in a shop who is considering purchasing some goods. They might experience physical needs (such as hunger), emotional states (such as surprise at the available goods), cognition (such as information about current prices) and influences associated with social status (which will affect how the agent reacts to the shop assistant). The framework is modular, with separate components controlling each aspect of the agent’s behaviour.

To compare the strength of all the different types of behaviour which might be acting upon an agent simultaneously, PECS uses the concept of ‘motives’. Motives can be drives (with drive-controlled behaviour), emotions (with emotionally-controlled behaviour) or acts of will (with constructive behaviour). Motives can be compared from different behavioural systems (e.g. comparing the drive to eat food with the act of will of studying for an exam) using what are called ‘intensity functions’. The motive with the highest intensity becomes the ‘action guiding motive’. Once the action guiding motive is known, then the agent can behave accordingly, whether this is to instinctively react to a stimulus or to create a complex action plan to pursue a constructive goal. Figure 1 provides an illustration of different motives and motive selection.

The PECS framework distributes all behaviours into two main categories: reactive and deliberative. Reactive behaviour classifies actions which are largely instinctive and can be modelled using a set of rules without deliberation on the part of the agent. The agent does not consider why it is behaving the way it is, e.g. they are not aware that looking for food is a task which ultimately ensures survival. Deliberative behaviour, on the other hand, involves the conscious pursuit of goals. The agent can deliberate over its current goal(s), form action plans to achieve a goal and break a larger goal into smaller sub-goals. Table 1 summarises the different types of behaviour as outlined by Schmidt and also provides the corresponding intensity functions.
Since rational decision making is not required and the framework is not restricted to the factors of beliefs, desires and intentions, PECS represents an improvement over the BDI architecture. For these reasons, the PECS framework appears to be the most appropriate for the development of more realistic burglar agents.

The documented use of the PECS framework, however, is currently limited, especially when compared to other behavioural models such as BDI. Both and have used the PECS framework to build emotions into a virtual learning environment. The authors incorporate non-verbal communication in the form of emotional facial expressions to improve the relationship between a human learner and a computer-controlled tutor. In the field of health care, the framework has been used to improve a simulation of disease screening. The use of PECS allowed for the incorporation of individual behaviour, which is an important determinant of a patient’s attendance at a screening session, and is absent from the majority of other models in this area. There is no currently published use of PECS for agent-based modelling of crime other than initial work by and.

3. Adapting the PECS framework for burglar agents

In environmental criminology, routine activities theory stipulates that a crime event comes about as a result of the spatio-temporal convergence of a motivated offender and a victim in the absence of a guardian to prevent the crime. Therefore, to predict the spatio-temporal locations of crimes it is necessary to model the daily activities of the offenders and victims as these will determine when the crime convergence takes place. In a similar vein, crime pattern theory notes that as a person navigates their environment – travelling between ‘anchor points’ such as work, home and friends’ houses – they build up an awareness space of the areas that they know well. Crime is thus more likely to occur where a person’s awareness space overlaps with a criminal opportunity. Modelling this theory requires a consideration for what a person’s anchor points are and the ability to include an individual awareness space built up over the course of a simulation.

The focus of the behavioural model in this research is, therefore, on creating realistic daily behavioural patterns where potential burglars travel around the city at sensible times, go to realistic places and build up accurate awareness spaces. This can be achieved through the simulation of drive controlled behaviour as relatively simple drivers can be used to determine agents’ current actions. It was decided that a more complex behavioural model that includes forms of deliberative behaviour is unnecessary at this stage. Only when the model itself is much more complex and includes a greater variety of behaviour (such as developing long-term life plans for example) must deliberative behaviour be explicitly modelled. Nevertheless, the following sections will demonstrate that the agent behaviour is actually more complex than simple reactive behaviours stipulated by PECS, and the framework has been adapted accordingly. For example, the use of awareness spaces are consistent with learned behaviour in that the agents remember the neighbourhoods that they have visited, which influences their future choices of burglary targets. Moreover, whilst the agents do not know why they need to satisfy their goals (which is a reactive trait), the methods that they use to satisfy them are complex and involve the “conscious pursuit of goals” with intermediate stages (i.e. a deliberative trait) such as obtaining money to satisfy a drug addiction.
3.1. State variables, motives and actions

Agents have a number of motives which vary in strength; the strongest motive at a particular time drives the agents’ behaviour. The intensity of the motives can be calculated using intensity functions which are based on internal factors (or ‘state variables’) as well as external influences. For example, an agent’s level of energy could be a state variable that affects their hunger motive. The first decision to make, therefore, is which state variables should be included in the model. As suggested by crime theories, burglars commonly become aware of potential burglary victims either by actively searching, or by passing them on journeys that are otherwise unrelated to burglary. Therefore it must be decided what legitimate (non-burglary related) behaviours should be included. The crime literature reveals that the most common drivers for burglary are the need to sustain a drug addiction\(^{36,37,38}\) or maintain ‘high living’ (i.e. socialising).\(^{36,39,40}\)

Therefore, the state variables will be (i) Drugs: the level of drugs in an agent’s system; (ii) Socialising: a measure of how much the agent has socialised; and (iii) Sleep: a measure of the amount of sleep an agent

---

**Table 1. Different types of PECS reactive behaviour**

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instinctive behaviour</td>
<td>An automatic reaction to stimulus such as a parent reacting instinctively to a child’s cry. Instinctive behaviour can be modelled relatively easily using pre-defined rules which are called up in certain circumstances.</td>
</tr>
<tr>
<td>Learned behaviour</td>
<td>Similar to instinctive behaviour but with rules that are learnt dynamically.(^{26}) cites the example of a car driver who will instinctively brake if they see a child running across the road.</td>
</tr>
<tr>
<td>Drive controlled behaviour</td>
<td>This type of behaviour is directed by internal drives to satisfy needs. These range from basic needs required to preserve life (such as the need for food or safety) to social needs and finally to intellectual needs.(^{26}) defines the function to determine drive intensity, (T), as (T = f(N, E, X)) where (N) is the agent’s personal preference for the need; (E) represents environmental influences; and (X) denotes other influences. For example, a drug addict will have a strong drive to take drugs if the need, (N), is high because they have gone without drugs for some time. However, the environment, (E), must also be taken into account: the drive might be strong if they are surrounded by other addicts who are also using drugs even if the need, (N), is not great.</td>
</tr>
<tr>
<td>Emotionally controlled behaviour</td>
<td>Emotions are similar to drives because, if they are strong enough, they will affect the behaviour of the agent. Unlike drives, however, they are stimulated externally, not by internal state changes.(^{26}) notes that the intensity of emotions, (E), are very hard to model, but defines the following formula: (E = g(I, A, X)) where (I) represents the importance of the event which has generated the emotion, (A) is the agent’s personal assessment of the event and (X) represents other influences.</td>
</tr>
<tr>
<td>Deliberative behaviour</td>
<td>26) discusses how an organism which is able to perform constructive behaviour is able to build an internal representation of its environment and also construct and deliberate over plans of action which should allow it to satisfy goals. Goals assembled in this manner are associated with acts of will, the organism “wants” to achieve the goal.(^{26}) In a similar fashion to reactive forms of behaviour which have a “need” associated with them, constructive behaviours have an “importance” attached to them by the agent which will influence their intensity. For example, one agent might attach a higher importance to the pursuit of gaining knowledge than another. In addition, the closer a goal is to completion, the higher the will associated with the goal.(^{26}) defines the following function, (h), to calculate will intensity, (W): (W = h(I, D, X)) where (I) is the importance of the goal, (D) is the distance from completing the goal and (X) are other influences.</td>
</tr>
<tr>
<td>Reflective behaviour</td>
<td>Representing the highest level of behaviour, reflective action relates to the ability to monitor and control one’s own thought processes. Also, in addition to a model of their environment, reflective organisms have a model of self which can lead to the most advanced forms of emotion such as an inferiority complex and jealousy.(^{26}) To model this type of behaviour, the PECS agent will have another entire PECS model of itself within its cognitive module.(^{26})</td>
</tr>
</tbody>
</table>
has had. Although this is a simplification, it provides sufficient variety to create realistic daily behaviour. It also represents an improvement over existing agent-based models of crime in terms of encapsulating agent behaviour.\textsuperscript{13,14} State variables can be thought of as internal energy levels, e.g. in the same way that a person with a low energy level will be hungry, an agent with low drugs/social level/sleep will want to behave in such a way as to increase the value of the state variable by taking drugs, socialising, or sleeping.

External variables will also affect behaviour. To account for this, the motive associated with each state variable can include external influences and, through the use of intensity functions, motives can be compared to each other thus establishing what the agent’s current action should be. The external influences are specific to each motive and will be discussed in detail below. In general, the intensity of a motive, $m$, is inversely proportional to the size of its state variable, $s$, i.e. the more $s$ an agent has, the lower their motive to collect or enact it will be. Since the population of agents does not need to be homogeneous, which is one of the benefits of agent-based modelling, different agents can be affected by state variables and motives differently. This feature is incorporated into the model by including a personal parameter, $p$, that affects motive intensities such that:

$$m \propto \frac{1}{ps}$$

(1)

For example, an agent with a large value of $p$ will be more strongly affected by a particular motive than an agent with a low value even if both agents have the same state variable level. This can be used to change the importance that agents place on particular behaviours. If $p = 0$ then $m$ will be zero and the agent will be unaffected by $s$. In the absence of external influences or personal preferences, Figure 2 depicts how motive intensity varies with state variable size. An exponential function is used so that as $s \rightarrow 0$, $m \rightarrow \infty$ and motives with the lowest associated state variable are likely to be the strongest. Figure 3 graphically illustrates how state variable levels are combined with personal preferences and external factors to determine the strongest motive (termed the action-guiding motive).

The levels of the state variables will deteriorate over time, such as a person becoming hungry as their energy level drops after a meal. For the purpose of the current model, the rate that state variables deteriorate can be used to configure the amount of time that an agent should spend, on average per day, satisfying the motive.

Burglary is commonly a response to a drug addiction\textsuperscript{36,37,38,41} so addiction must therefore form part of the model. In particular, the process of going to purchase drugs will have important influences on the agent’s cognitive map, making geographical areas with (and on the way to) large numbers of drug dealers more susceptible to burglary. Drug taking, in this instance, is clearly a drive-regulated behaviour as outlined in Section 2.3, being dependent on the current drug level in the agent’s system. In contrast to the general form of drive-regulated behaviour, the motive associated with drug use does not depend on other factors such as the surrounding environment. Therefore the strength of an agent’s drug motive, $m_{dr}$, can be calculated from their personal preference for drugs, $p_{dr}$ and the drug’s state variable level, $s_{dr}$:

$$m_{dr} = \frac{p_{dr} \cdot s_{dr}}{sd}$$

(2)

The second activity is socialising. All agents have the need to socialise which includes visiting friends’ houses, pubs, etc. The choice of areas that an agent is likely to visit is very important because it will strongly influence the size and shape of the agent’s cognitive map. The social motive, $m_{soc}$, is dependent on the time of day, $g(t)$, the agent’s personal preference for socialising, $p_{soc}$, and the size of the social state variable, $s_{soc}$:

$$m_{soc} = \left(\frac{1}{s_{soc}}\right) \left(\frac{g(t) + p_{soc}}{2}\right)$$

(3)

The shape of $g(t)$ is illustrated in Figure 4; agents will desire to socialise more in the evenings than during the day.
The final activity is sleep. All agents are set to sleep eight hours per day on average, which is a generally accepted healthy amount.\textsuperscript{42} We assume that this is the length people would sleep if not interrupted, and will ultimately have to make up. The time of day also affects the strength of the motive so that the desire to sleep is stronger at night than during the day. The shape of this function is illustrated in Figure 5. Although this type of sleep pattern might reflect most people’s habits, it should be noted that the chaotic lifestyles of many potential burglars might result in very different sleep patterns. In general, an agent’s sleep motive intensity, $m_s$, can be calculated from the sleep state variable level, $s_s$, the time of day function, $f(t)$, and their personal preference for sleep, $p_s$, as follows:

$$m_s = \left( \frac{1}{s_s} \right) \left( f(t) + p_s \right)$$

(4)

As discussed previously, it is possible to vary the rate that the Sleep state variable deteriorates and the

---

**Figure 3.** How state variables, $s$, personal preferences, $p$ and external factors (e.g. the time of the day, $t$) are used in intensity functions to determine the action-guiding motive. In this example, the agent’s social level is very low (the agent has not socialised in some time) and this is the strongest motive.

**Figure 4.** How the social intensity motive varies with time of day, assuming a constant value for the social state variable.
amount that it increases while an agent is sleeping. In this manner, the model can be configured such that an agent will sleep for 8 hours per day on average.

Each motive has an associated goal which the agent can accomplish to increase the value of the appropriate state variable and subsequently lower the strength of the motive. For example, accomplishing the goal of ‘taking drugs’ will increase the value of the Drugs state variable, reduce the size of the motive and cause another, larger, motive to start controlling the agent’s behaviour. To satisfy a goal, however, there are often numerous sub-goals that must be accomplished first. For example, with socialising and buying drugs, wealth is required that must be sought through burglary (a complex action which requires many sub-goals itself as discussed in Section 3.2). Flowcharts illustrating how goals can be accomplished via the use of sub-actions are presented in Figure 6.

3.2. The process of burglary

Finding a property to burglar is the most intricate of all the agent actions, and consists of numerous goals and sub-goals. It is also the most important as it will have the greatest effect on final citywide burglary patterns. The burglary process can be broken down into three distinct actions:

1. Deciding where to start looking for victims;
2. Searching for victims;
3. Deciding upon a suitable target.

The advantage of this modular approach is that it enables different types of burglar to be simulated simply by replacing one action with another (e.g. how the agent searches for a target). The main drawback with this approach, however, is that it does not allow for purely opportunistic burglary, which is common. For example, using the above scheme, it is not possible for an agent to notice an open door or window and choose to burgle immediately even if they were not otherwise considering burglary. They have to make a conscious decision to start the burglary process regardless of their surrounding environment. Including more opportunistic types of burglary is recommended as an avenue for future research.

3.2.1. Deciding where to start the search. Burglars act as ‘optimal foragers’ when they choose target areas because their decision is based on an analysis of potential rewards against risks. The model here works in a similar manner. When deciding where to start searching, agents consider the communities of which they are aware and assign a likelihood, \( l \), to every area, \( a \) in their cognitive map, relative to their home, \( h \), and current location, \( c \):

\[
l_a = w_1 \ast \left( \frac{1}{\text{dist}(ca)} \right) + w_2 \ast \text{attractiveness}(h, a) + w_2 \ast \text{socialDiff}(h, a) + w_4 \ast \text{prevSucc}(a) \tag{5}
\]

where

- \( \text{dist}(c,a) \) represents the distance (in travel time) to the target from the agent’s current position. Research has shown that agents are unlikely to travel far from their homes, so farther areas are less attractive. Here we use a linear decay function. \( \text{Attractiveness}(h,a) \) represents the abundance of attractive goods of the potential target relative to the agent’s home. Using relative attractiveness provides for the finding that affluent communities are at most risk of burglary when they are close to deprived communities.

- \( \text{socialDiff}(h,a) \) represents the difference between the agent’s home area and that of the potential target (where values of 1 indicate similarity and 0 indicates dissimilarity) as offenders are more likely to target
areas which they know well and where they feel safe. The difference is calculated as the Euclidean distance between all the socio-demographic variables from the 2001 UK census in each area.

- \( \text{preSucc}(a) \) is the number of previous successful burglaries which the agent has committed in area \( a \); numerous successful burglaries are likely to encourage the offender to return to the same area. This ‘near repeat’ phenomenon has been found in numerous crime research studies and also informs burglary-reduction policies in Leeds.

An extremely important element of Equation 5 is the variables \( w_1 \) to \( w_4 \). These are weights, discussed in further detail below, which will influence how important the different factors are to the different burglars and depend on the burglar type along with their internal variables. For example, a ‘professional’ burglar with access to a car might be unconcerned with the distance that they need to travel (i.e. a low value for \( w_1 \)) compared to a ‘chaotic’ burglar with a serious drug addiction and limited access to transport (i.e. a high value for \( w_1 \)).

Before applying the weights and calculating \( l \), each variable must be normalised so that they all have the same magnitude. To determine the area to which the offender will travel, roulette wheel selection is used so that the decision is probabilistic. An agent calculates \( l \) for every area, \( a \), in their memory and then chooses an area such that the probability of being chosen, \( P(a) \), is proportional to its share of the sum of all \( l \):}

\[
P(a) = \frac{l_a}{\sum_{i=0}^{n} l_i}
\]

This probabilistic component makes the decision more realistic; it is unlikely that a human will always reach an identical decision even when faced with similar input information.

### 3.2.2. Searching for a victim

Research has shown that burglars do not search randomly for burglary targets but exhibit identifiable search patterns. For example, has identified the tear-drop and bulls-eye patterns. The bulls-eye pattern suggests that the burglar’s search expands outwards from their home (or another anchor point) such that houses close to the start of the search have the largest risk. The tear-drop pattern, on the other hand, stipulates that the journey to the start of the search is included so that the overall search pattern looks like a teardrop connecting the home and location that marks the start of the search. The search implemented in the model combines both of these ideas. As the previous section discussed, the offender chooses an area to start their search. They then travel to their chosen location and start a bulls-eye like search. On the route to their chosen start location
the offender observes properties for burglary suitability (discussed in the following section) so that if the chosen area is close to their starting position then the final search pattern will resemble a bulls-eye, otherwise it will resemble a tear-drop.

Once the offender reaches a suitable area and begins a search, the amount of time spent on the search depends on the characteristics of the offender, with desperate offenders spending the shortest amount of time and often victimising the first empty house they find.\(^{49}\) This search time variable can be varied in the model. Once it is reached, the burglary process is repeated; the agent chooses a new start location, travels there and commences a new search.

### 3.2.3. Choosing a suitable victim.

Once the agent decides to commit a burglary, they start to examine all the houses they pass in order to determine their suitability for burglary. This happens on the way to the target location as well as while they are actually performing the search, as stipulated in the literature.\(^ {39}\) Table 2 illustrates the variables that determine the suitability of each property. Overall household suitability is then calculated by summing the individual components:

\[
\text{suitability} = \frac{w_5 \cdot CE + w_5 \cdot TV + w_7 \cdot Occ + w_8 \cdot Acc + w_9 \cdot Vis + w_{10} \cdot Sec}{\sum_{i=5}^{10} w_i}
\]

where \(w_5\) to \(w_{10}\) are weights that can be applied to each of the variables in order to personalise the calculation depending on the burglar type. For example, a ‘professional’ burglar might be less deterred by security than an ‘amateur’ and will subsequently have a lower value for \(w_{10}\). The suitability is normalised to the range 0 to 1 where the most suitable properties will have a value near 0 and the least suitable near 1.

Equation 7 will return an absolute value for the suitability of a house for a particular agent. The final step is to determine whether or not the agent is desperate enough to commit the burglary. ‘Desperateness’ is based on the intensity of the motive which is currently driving the agent’s behaviour: if the suitability of a property is higher (i.e. less suitable) than the intensity of the motive, then the agent will not attempt a burglary. If, however, the suitability value is lower than the agent’s motive intensity then the agent might attempt a burglary. A random element is included such that the agent is more likely to burglar if the difference between the suitability of the house and the agent’s motive intensity is large. This is important because otherwise the agent’s decision becomes somewhat deterministic and they are also likely to burglar the first house they find once they become active (i.e. their next-door neighbour), which is not always realistic. Crime studies have shown that although burglars are unlikely to travel far, they will not usually burglar too close to home for fear of being recognised.\(^ {50}\) A gradient of \(y = x^2\) is chosen so that, with only small differences between an agent’s motive intensity and house suitability, they are unlikely to burglar.

### 3.3. Different types of burglar agent

Although each burglar uses the same ‘burglary template’, by varying weight values it is possible to represent a wide range of burglar behaviour and, therefore, simulate different types of burglary. Aggregate techniques traditionally employed in crime modelling, such as hotspot modelling\(^ {51}\) or regression analysis,\(^ {52}\) cannot incorporate this level of behavioural complexity to simulate the effects of heterogeneous burglar behaviour on city-wide crime rates.

However, understanding the different ‘types’ of burglar behaviour is non-trivial. Although there is a large body of literature that attempts to classify criminal propensity,\(^ {53}\) there is less research that attempts to classify burglar behaviour. Most burglaries were found to be committed by low income amateur offenders for relatively low returns.\(^ {54}\) This is consistent with the findings of\(^ {55}\) and\(^ {56}\) who found that the majority of

<table>
<thead>
<tr>
<th>Environment variable</th>
<th>How it affects burglar behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collective efficacy (CE)</td>
<td>High levels of CE make the area less attractive to burglars because the community appears cohesive and neighbours/passers-by are likely to notice someone acting suspiciously.</td>
</tr>
<tr>
<td>Traffic volume (TV)</td>
<td>High levels of traffic volume make the houses on a road less attractive because it is difficult to access a property without being seen by passers-by.</td>
</tr>
<tr>
<td>Accessibility (Acc)</td>
<td>Houses with few possible entrances are more difficult to enter without being seen by others.</td>
</tr>
<tr>
<td>Occupancy (Occ)</td>
<td>Houses which are likely to be occupied are less attractive.</td>
</tr>
<tr>
<td>Visibility (Vis)</td>
<td>Houses which are highly visible to neighbours/passers-by are more difficult to enter.</td>
</tr>
<tr>
<td>Security (Sec)</td>
<td>High levels of security can present problems to potential burglars.</td>
</tr>
</tbody>
</table>
crimes were not committed in a planned, systematic way by ‘professionals’. At the simplest level, burglary behaviour can therefore be divided into two classes: amateurs and professionals. However, being an ‘amateur’ does not necessarily mean that no planning goes into a burglary, and not all burglaries in these situations are purely opportunistic.\textsuperscript{57} For example,\textsuperscript{58} differentiate between ‘opportunist’, ‘search’ (an offender becomes active and then searches for an immediate victim) and ‘planned’ (the offender returns to a previously found opportunity at a later time/date) offences and argue that over half the offences were planned. Similarly,\textsuperscript{57} propose the labels ‘professional’, ‘journeyman’ and ‘novice’ when referring to burglar types. However, it is highly probable that, when interviewed after the event, a burglar might rationalise the burglary to a much greater extent than they did at the time.\textsuperscript{57}

To support the crime literature, and gain an insight into the specific behaviour found in the study area of Leeds, local crime-reduction practitioners at Safer Leeds were interviewed. They suggested that the different types of burglar behaviour were very similar to those already outlined in the literature. Table 3 provides the types recommended by Safer Leeds and the literature.

### 4. Building the theoretical model

The previous section has outlined a means of incorporating realistic burglar behaviour into an agent-based model. To test the design, a model was built using the Java programming language and the Repast agent-based modelling toolkit (available at http://repast.sourceforge.net/ (2008)). The model is extremely computationally expensive and, therefore, was executed using high-performance computing provided by the UK’s National Grid Service (NGS).\textsuperscript{59} For more information about the model and the scenario which is discussed below, see.\textsuperscript{12}

#### 4.1. Crime data and model setup

Although the model could be applied to any urban area, due to data availability it is presently based in the city of Leeds, UK. Crime data were provided by Safer Leeds which contains, along with other information, a dataset of people who have somehow been involved with a crime event (termed “nominals”). It must be noted that involvement does not necessarily mean the person was ultimately convicted of the crime – they could be suspects or simply be wanted for questioning. Therefore it is inevitable that the dataset will contain people who were not actually involved in the crime at all.\textsuperscript{60} Nevertheless, the dataset is the best empirical data for offending activities available and it will be used to validate the behaviour of agents in the model by comparing the behaviour of nominals to that of simulated agents.

Each entry in the dataset contains the home postcode of the nominal’s given address when they were added as well as the postcode in which the associated crime was committed. Because the simulation only covers part of the entire city (as discussed in the following section), the nominal data set was filtered to contain only people who lived within the simulation boundary and were associated with an offence within the boundary.

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chaotic</td>
<td>An opportunist who is desperate to generate wealth in any means possible (i.e. through burglary or other types of crime) in order to fund a drug addiction. This high level of drug addiction is reflected by a chaotic lifestyle. Often unemployed and with limited means of transport, they are unlikely to travel far to search for a victim. Also will want to avoid confrontation.</td>
</tr>
<tr>
<td>Local opportunist</td>
<td>Again an opportunist but will not commit crime if it is too difficult (they will be deterred by security and guardianship). Not necessarily a “full time” burglar, they are often younger, do not need to support a drug habit, occasionally have legitimate employment and will only burglar if a good opportunity presents itself. Their limited access to transport means that they will only travel short distances.</td>
</tr>
<tr>
<td>Travelling chaotic</td>
<td>Similar to the chaotic type but more aware of the best opportunities and will travel further to reach them. For example, travelling to an insecure car park to commit theft from a motor vehicle. Otherwise very similar to the chaotic type.</td>
</tr>
<tr>
<td>Organised/professional</td>
<td>The most organised type, there is a considerable amount of planning involved in a burglary. The type is characterised by low levels of drug addiction, occasional legitimate employment, skills to evade security precautions and will travel the furthest distances for the greatest rewards. Confident in ability to “blend in” with community and tackle security precautions.</td>
</tr>
</tbody>
</table>
Offender agents in the model are generated directly from entries in the nominal dataset by creating one agent per entry. Agents are assigned to a house (chosen at random) from within the output area surrounding the nominal postcode centroid. Creating one agent per nominal means that prolific offenders in the data (those who commit a number of burglaries) are actually represented by a number of different agents who live in the same area (although not in the same house). Therefore there is no concept of prolific offending in the current model configuration; every offender in the model is identical with respect to their behaviour. An avenue for future work is to experiment with heterogeneous behaviour and create different types of burglar agent as discussed in Section 3.3. This initial configuration of agents is unlikely to be an accurate picture of offending in the city and will therefore limit the potential of the model to predict overall burglary rates. However, it does not detract from the ability of the model to simulate burglar behaviour. In fact, it makes the task easier because the behaviour of nominals in the crime data can be compared directly to the behaviour of simulated burglars.

4.2. Environmental data

The scenario itself is based in a part of east Leeds which is home to a large regeneration project called EASEL (East and South East Leeds) as illustrated in Figure 7. This area was chosen because it is a useful candidate for
predictive modelling. With respect to environmental data, the following sources were used to construct the required layers of the virtual environment:

- **Ordnance Survey MasterMap data.** This highly detailed spatial dataset consists of three separate layers:
  - The *Topographic Area* layer contains building boundaries and can be used to construct virtual houses (these are the places people live and where the burglar agents can choose to burgle).
  - The *Address Layer* which can be used to distinguish between different types of building (i.e. offices, house, garages etc);
  - The *Integrated Transport Layer* contains different types of roads such as motorways, major/minor roads, alleyways, pedestrianised streets etc. This can be used to build the transport network which the agents use to travel around the environment.
- **Output Area and Super Output Area** boundaries. The Output Area (OA) is the smallest geographical area boundary used to provide a spatial reference to the 2001 UK census and contains approximately 100 households. This is used to represent demographic community factors that might influence the behaviour of a potential burglar. Super Output Area (SOA) boundaries are built by combining a number of OA boundaries and are used later for data aggregation when validating the model (see Section 5).
- **The Output Area Classification.** A classification scheme that groups individual output areas into distinct types based on the values of their socio-demographic census variables. This can be used to distinguish between different types of community.

The locations of potential drug dealers in the model were established directly from the crime data, creating a virtual dealer address for every point in the data set where a dealing-related crime had been recorded. Creating places where the agents are likely to socialise, however, is considerably more problematic. In light of the lack of empirical evidence to guide the choice of social address, broad assumptions must be made. At this stage, the National Land Use Database (NLUD) code is used to estimate where burglar agents might go to socialise. Social places are considered to be U093 (“Restaurants and cafes”) and U094 (“Public houses and bars”). It is important to note that none of these decisions are fixed and it would be a trivial operation to change social-place addresses in the future as new data becomes available. An immediate area for future work should be to estimate where friends of the offender might live, rather than solely using restaurants and pubs for socialising places.

It is also necessary to decide, from the set of all available locations, *which one* an agent chooses to travel to if they need to visit a drug dealer or to socialise. This is one of the most difficult features to estimate as there are very limited data available to assist in making the assumptions. With regards to drug dealers, it was decided that the agent is assigned a drug dealer at random and always uses the same one. It is likely that in reality a person builds a preference for certain dealers but often travels to different addresses depending on the abundance of supply, but this avenue of exploration is beyond the scope of this work. With regards to social locations, it is assumed that an agent is more likely to travel to a social location that is in a community of a similar type to their own. Again this is likely to be too simplistic but can be investigated further in the future.

Calibration is still required for some variables because data or expert opinions on their values are not available. These variables are presented in Table 4. While an automated calibration using a genetic algorithm would have been ideal this is not practical with current run-times, even with the model parallelised (as is necessary for a single run). The model was therefore initialised using 2001 census and offender data and calibrated by hand against 2001 victim data (details of these variables and the calibration can be found in12). Because the simulation is probabilistic, each run will produce slightly different results. For this reason the simulation was executed 50 times and the results collated to reduce variability. This means that there are many more burglary events in the simulated data compared to the real data, although this is not a problem in the following analyses as the geographical distribution of relative proportions will be used rather than absolute crime counts.

5. Results

The model was calibrated against 2001 victim data (Figure 8a), and the best result is shown in Figure 8b. The reasonable match between these two datasets suggests the model can, structurally at least, replicate real aggregate crime figures in space. To validate the predictive power of the model it was run with 2004 offender data, to predict 2004 victim data (Figure 8d). There is an inherent difficulty in the prediction in that in some cases the absence of a 2004 census dataset meant that the 2001 census had to be used during initialisation. As can be seen (Figure 8c), the results from the model are nonetheless broadly good at identifying crime-susceptible areas, though the most intense areas of crime have shifted. This suggests the model, and the theories it is founded on, do reasonably well at replicating the spatial patterns of the aggregate crime density. Further comparisons, including a multi-scale statistical
comparison, can be found in. Here, however, we would like to concentrate on some of the detailed behaviour, reflecting on the extent to which the model replicates individual behaviour and the areas that still need development in both the model and theory.

While the model match with victim data at the aggregate scale is fairly good, this makes no assessment of whether the correct offenders are committing the correct crimes, or, more generally, whether there are the right levels of offender flows between areas. A more thorough examination of the offending behaviour can be obtained by generating origin-destination matrices from the real and simulated data which illustrate how many people travelled between each area to commit a burglary. The matrices can be compared using traditional goodness-of-fit statistics that quantitatively illustrate how similar they are. If the model is accurately reflecting the behaviour of people in the real dataset, then the matrices will be similar. Before generating the matrices, however, it must be determined how the burglary points will be aggregated. Using larger areal boundaries is more likely to generate matrices that are similar but provides a lower resolution assessment of the success of the model. Similar studies (e.g.63) have used ward area boundaries, although these are deemed too large for use here (there are only nine wards that cover the simulation boundary). Instead, Lower Super Output Area (SOA) boundaries will be used as these are smaller than wards (there are 58 SOAs covering the study area). Using the $R^2$ statistic to compare the simulated and observed origin/destination matrices results in a similarity of 0.22 (Table 5; i.e. the model is able to explain 22% of the variation in the observed data). Although this value is not substantially lower than other studies – .64 for example, note that an $R^2$ value of only 0.36 is generally on par with other criminal justice studies – it is lower than expected. It is highly likely, however, that by using administratively-defined areal boundaries the analysis is susceptible to the modifiable areal unit problem,65,66 mis-representing

---

**Figure 8.** a) Real crime data, 2001; b) Results of best model calibration to 2001 data; c) Model prediction for 2004; d) Real crime data for 2004.
boundary-agnostic short-distance flows. Aggregating to a square grid instead might mitigate these problems. Therefore the data were also aggregated to a square grid with cells of a size equal to the mean of the SOAs that cover the simulation area (0.42km²). This resulted in an $R^2$ value of 0.30 which, although more accurate than the SOA-aggregated data, is still relatively low.

Insight into the location of the errors can be gained by examining the distance between each offender’s home and the crime site in the two datasets. Figure 9 illustrates a histogram of the distances travelled and Table 6 provides the mean and standard deviation of the distances. It is apparent that, on average, agents in the model are likely to travel further to their crime site than nominals in the real data. This may mean that the burglar agents are both more mobile and willing to travel further in the simulation than in the real world.

Table 4. Variables used in the calibration

<table>
<thead>
<tr>
<th>PECS variables</th>
<th>House and Community parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorkGain</td>
<td>Accessibility importance</td>
</tr>
<tr>
<td>SleepGain</td>
<td>Visibility importance</td>
</tr>
<tr>
<td>SocialGain</td>
<td>Security importance</td>
</tr>
<tr>
<td>DrugsGain</td>
<td>Traffic volume importance</td>
</tr>
<tr>
<td>CostSocialise</td>
<td>Collective efficacy importance</td>
</tr>
<tr>
<td>CostDrugs</td>
<td>Occupancy importance</td>
</tr>
<tr>
<td>ConstTravelTime</td>
<td>Attractiveness importance</td>
</tr>
<tr>
<td>DeteriorateAmount</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. The goodness-of-fit (measured using the SRMSE and $R^2$) between origin-destination matrices for the real and simulated data

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>SRMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>5.81</td>
<td>0.22</td>
</tr>
<tr>
<td>Grid</td>
<td>6.60</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 6. The mean and standard deviation of the distances travelled by agents in the simulation and nominals in the real data

<table>
<thead>
<tr>
<th>Data</th>
<th>Mean commute distance (m)</th>
<th>Standard deviation (3sf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real data</td>
<td>862</td>
<td>833</td>
</tr>
<tr>
<td>Simulated data</td>
<td>1630</td>
<td>1090</td>
</tr>
</tbody>
</table>

Figure 9. A histogram comparing the distance between the home location and the crime location for agents in the model and nominals in the offender data set.
A first-sweep solution to this would be to cap the distance travelled to a more realistic value to better capture this phenomenon. However, ultimately this may reflect the absence of the most opportunistic burglars, and those burglars desperate enough to target their neighbours.

As Section 4 discussed, the nominal data used here is unlikely to fully reflect the real system even if we had information on short-distance burglars, and so a more pertinent verification of the model would be to see if the behaviour of individual agents replicates the behaviour we understand for specific criminal types for which we
do have data. Major environmental criminology theories such as routine activity\(^2\) and crime pattern theory\(^3\) as well as numerous empirical studies\(^{39,69}\) suggest that an offender becomes aware of opportunities as they travel around the city on potentially legitimate business. Burglaries are then likely to occur in the areas that the offenders know well, which are commonly around their home and other anchor points such as friends’ houses, drug dealers etc. The model here acts in a similar manner; as the agent travels around their environment they remember the houses and communities that they have passed and these form their overall awareness space. Figure 10 depicts such an awareness space for an agent at the end of a simulation as well as the locations of the burglaries that they have committed, the places that they visited to socialise and the address of their drug dealer. It is apparent that, as stipulated by crime pattern theory, the agent’s awareness space is built up around their anchor points and the routes between them. This then determines where they are the most likely to commit burglary, i.e. burglaries occurring in the places that they are the most familiar with. Overall this demonstrates that the model is closely replicating theory and filling a gap significant in current modelling practice.

Interestingly, however, there are no burglaries committed by the observed agent in the vicinity of the agent’s drug dealer. This seems counter-intuitive.
because one might expect that, as the offender will be very familiar with the drug dealer’s neighbourhood, they are likely to commit crimes there (assuming there are available crime opportunities). However, as Figure 11 demonstrates, the simulated agent in this case commits a burglary along routes between their home and other anchor points (in this case their drug dealer). The process is as follows:

1. The agent decides to take drugs but first requires money.
2. They travel to a nearby area that they know well in order to search for a victim (in this case the target is on the route to their drug dealer; this type of behaviour is stipulated by crime pattern theory\(^3\)).
3. Once they have completed the burglary they travel to the drug dealer and then back home.

This demonstrates a close correspondence with criminology theory, which is only possible when individual virtual offenders are equipped with a realistic behavioural framework.

To determine how the development of the agent’s knowledge of their environment compares to their offending behaviour, Figure 12 illustrates how the agent’s awareness space develops over the course of the simulation. As the agent visits a larger number of activity nodes they become aware of a greater number of opportunities and are more likely to travel further afield for burglary. This factor has also been found by quantitative studies; offenders who are older and have access to transport have greater awareness spaces and are likely to travel further to commit crime than younger people who have narrower awareness spaces.\(^41,70,71\)
To complete the comparison of simulated behaviour to criminology theory, Figure 13 illustrates the search behaviour of a simulated burglar. The burglar leaves their home, travels to a chosen location and then begins a search in the area. As they travel to the location they observe the suitability of the houses that they pass, so if an opportunity presents itself, then they will take it. If they reach the desired area without having found a victim, they will begin a search of the area and, in this case, the agent is able to find a suitable target in the search area. As Section 3.2.2 discussed, this type of searching behaviour is consistent with understanding of real and theoretical criminal movement in the literature.48

6. Conclusions

This paper presents a first attempt to replicate the individual behaviour of those engaging in burglary using a psychologically realistic framework. It details the PECS framework, current criminological theory that might be modelled better utilising such a framework, and describes an individual-based model of burglary that enacts this potential. The model takes into account the daily routines of criminals, their drivers, and the opportunities presented by their environment. Current aggregate mathematical models cannot take individual-level behaviour of this type into account, and yet criminological theory centres on its absolute importance in
understanding crime. As such, the model presented represents a significant step towards a criminological model with a greater verisimilitude and therefore greater potential, both as a predictive tool and as a tool for assessing and understanding how individual behaviour builds to a regional crime pattern.

At an aggregate level the model does a reasonable job of modelling the spatial distribution of the crimes. However, when we look at the match between individual behaviour in the real world and the model, we still have some ground to cover. The matching of the aggregate statistics despite the omission of key criminal types and datasets indicates that there are broad equifinality issues with crime data, and that a range of theories may replicate the aggregate patterns within the error ranges expected of aggregate modelling. It is only by comparing behaviour and statistics that track individual movements more closely, like journey-to-crime paths, that we may distinguish whether particular theories are an appropriate explanation of aggregate data. Given this, it is plain that models such as this, which concentrate on the individual drivers, responses, and behaviour of offenders, are the only mechanism by which regional spatio-temporal patterns of crime can be reliably related to criminological theory. Future developments of the model will represent a greater diversity of criminal types, particularly distinguishing between opportunistic and professional burglars with greater detail; however it should be clear from these initial results that there is considerable potential for models enacting a more realistic, individual-level behavioural framework, both in terms of crime prediction, and in terms of theory testing and development.

This research was funded through a PhD studentship provided by the Economic and Social Research Council (ESRC).

Acknowledgements

The authors would like to thank Safer Leeds for their invaluable data and professional expertise as well as the National Grid Service (NGS) for the use of high-powered compute resources.

References


