Crime Science
a SpringerOpen Journal

RESEARCH   **Open Access**

Modelling patterns of burglary on street networks

Toby P Davies[1,2*] and Steven R Bishop[1]

Abstract

A fundamental issue in crime prevention is the efficient deployment of resources and the effective targeting of interventions, both of which require some form of prediction of future crime. One crime for which this is feasible is burglary, the distinctive spatio-temporal signatures of which can be exploited to inform predictions. Mathematical models in particular are capable of both encoding concisely the theoretical foundations of criminal behaviour and allowing the quantitative analysis of specific scenarios, and their capacity to reproduce the general patterns of burglary suggests that the approach has considerable potential. Previous models, however, are situated on simplified representations of space and do not reflect realistically the built environment in which crime takes place; specifically, they do not incorporate urban street networks. Such networks are fundamental to situational theories of crime, in the sense that they determine the configuration of urban space and, therefore, shape those human activity patterns which are thought to give rise to crime. Furthermore, streets are the natural domain for many policing activities, and their structure is determined by planning decisions, so that insight into their relationship with crime is likely to be of immediate practical use. With this in mind, this paper presents a mathematical model of crime which is explicitly situated on a street network. After discussing theoretical considerations and specifying the model itself, examples of typical networks are explored.

Keywords: Burglary; Street network; Applied mathematics; Modelling

Background

Residential burglary has been, and remains, a significant criminal problem, and consequently has been the subject of academic research for some time. As with most analytical work concerning crime, the objectives of such research are both theoretical, in improving the understanding of the process leading to the crime, and practical, in using the insights gained in order to prevent future crime. With respect to the latter, this typically entails some form of prediction, either at the fine spatio-temporal granularity which, for example, might be required for a strategy of 'hotspot policing' (Chainey and Ratcliffe 2005), or in the more generalised terms which might be used to inform long-term policy. Previous work has shown predictive policing based on statistical analysis to have considerable potential (Johnson et al. 2009a), and the desire to build

on this is well-aligned with the desire of administrators to carry out policing more efficiently; a principle which is also a central theme of the field of crime science (Laycock 2005). Modelling has much to offer here, through its ability to distil theoretical mechanisms to formal expressions and to afford both quantitative and rigorous analysis of hypothetical scenarios via well-established techniques. Indeed, this potential, along with the increased availability of geographical data and the development of a wide array of tools for the analysis of complex social systems, has inspired significant recent interest within the modelling community. Within this domain, attempts to model burglary range from mathematical approaches (Berestycki and Nadal 2010; Pitcher 2010; Short et al. 2008) to those employing agent-based simulation (Birks et al. 2005; Groff 2007a; Johnson 2008; Malleson et al. 2009), with these playing complementary roles. This divergence is common to crime modelling as a whole: although the possibility of generalised analysis means that mathematical models might offer greater insight, they have so far failed to match agent-based approaches in terms of their scope for the

*Correspondence: toby.davies@ucl.ac.uk
[1] Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, UK
[2] Department of Security and Crime Science, University College London, Gower Street, London, WC1E 6BT, UK

Springer

incorporation of detailed individual-level behaviour, with only some examples occupying a middle ground (Davies et al. 2013; Short et al. 2008).

Criminological theory

In the case of mathematical approaches, much previous work has focussed on one of two aims. The first of these is to explore whether hypothesised mechanisms for burglary, when encoded mathematically, are sufficient to generate the generalised patterns observed empirically (Berestycki and Nadal 2010; Pitcher and Johnson 2011; Short et al. 2008), and this appears to indeed be the case in simplified mathematical representations. This general approach is also characterised by its versatility, and has been used as a starting point for more realistic extensions, such as those incorporating policing activity (Pitcher 2010), which lend themselves to the exploration of simple policy questions. On the other hand, with a view to producing effective tools for short-term prediction, research elsewhere has sought to make specific predictions about the location of future risk, given a particular set of current burglary data (Mohler 2011). Common to both approaches, however, is the intention to leverage the patterns observed in empirical burglary data in order to make predictions. Such patterns have been studied widely, and can generally be summarised as forms of clustering, though acting at varying scales and dimensions. Looking first at space, it has long been established that the locations of crimes are clustered (Shaw and McKay 1969; Sherman et al. 1989), and the importance of place in the study of crime has been emphasised accordingly (Eck and Weisburd 1995). The importance here of scale, though, is considerable, since patterns can be subject to significant heterogeneity within units of analysis: previous studies have shown area-level crime rates to be driven largely by those of only a few streets (Weisburd et al. 2004) and burglary risk to vary significantly between individual houses (Bowers et al. 2005). In line with this, the importance of retaining fine spatial granularity in crime analysis has been stressed (Brantingham et al. 2009).

As well as a location in space, however, all crime occurs at some point in time, and clustering is also evident here, most obviously manifested as daily and seasonal cycles. While valuable insight can be gained by considering this, however, of greater interest is the interaction between spatial and temporal dimensions; indeed, to ignore this may lead to incomplete understanding of each individual factor. Studies have consistently demonstrated that crime clusters in both space and time (Grubesic and Mack 2008; Johnson et al. 2007; Townsley et al. 2003); that is, that the clustering is greater than would be expected if it were simply occurring independently in each dimension. In the case of burglary, this is exemplified by the phenomenon of *repeat victimisation*, whereby victimised properties, for a period after the initial event, are subject to a rate of further victimisation over and above that which would be expected by chance (Farrell 2005). Indeed, the temporal component of this is particularly distinct, with risk appearing to decay exponentially with time after an initial event (Johnson et al. 1997). In addition, the concept can be extended to that of *near-repeats* (Morgan 2001), whereby properties close to an initial event also experience elevated risk for some period afterwards. These phenomena, which represent the primary drivers of burglary hotspots in urban environments, appear to be ubiquitous in data from several countries (Johnson et al. 2007; Townsley et al. 2003).

The near-repeat patterns observed empirically imply that burglary victimisation cannot be understood by considering properties in isolation, and that they must instead be considered in the context of their wider neighbourhood. In trying to account for repeat victimisation, much attention has focussed on two (non-mutually-exclusive) hypotheses, both of which seek to provide explanations based on factors acting at the level of individual properties. The first of these is the concept of *risk heterogeneity*, also known as the *flag* hypothesis, which suggests that patterns may arise due to differences in the time-stable risk of burglary at individual properties (Pease 1998). This 'risk' refers to any non-varying factor which may influence the probability of victimisation, such as the type of property, presence of security features, affluence or location. That repeat victimisation might arise from this can be understood by a simple statistical argument: even if burglaries occurred randomly, some repeat victimisation would occur by chance, and the fact that some properties are more attractive than others simply biases this process. The preferential victimisation of more attractive properties necessarily implies that the time between their victimisations will be shorter, and that more repeats will therefore occur on this basis. This can be extended easily to near-repeats by considering that nearby properties are likely to be of similar attractiveness.

The other explanation which is typically invoked is the *boost* hypothesis, which states that, for some period after an initial event, the risk to nearby properties is temporarily elevated (Pease 1998). The natural explanation for this is that any repeat offence is likely to be the work of the same offender (since it is unlikely that another offender would have knowledge of the first incident). The reason for the elevation can be understood by considering the decision process of a rational offender (Cornish and Clarke 1986): the commission of the first offence affords knowledge of both how to successfully burgle the property and the potential rewards available, making it a more attractive proposition than an 'unknown quantity' when identifying a future target. This is lent further credence by interviews with offenders (Ashton et al. 1998; Cromwell

et al. 1991; Summers et al. 2010) and police detection data (Bernasco 2008; Johnson et al. 2009b), both of which support the identification of repeat incidents with the same offender. Again, this idea is easily extended to near-repeat victimisation: knowledge of one house is likely to offer insight into the characteristics of its neighbours (such as the layout and location) and, furthermore, the journey to and from the initial crime affords an opportunity to evaluate nearby properties. The question of the relative strength of these effects - flag and boost - has been considered in recent work using both simulation-based (Johnson et al. 2009b) and statistical approaches (Short et al. 2009), in which evidence was found in support of both hypotheses.

Both of these arguments can be cast in the light of more general theories of 'environmental criminology' (Brantingham and Brantingham 1981); an approach which focuses explicitly on the criminal act itself, and the circumstances which give rise to it, rather than the characteristics of the offender. Central to this is 'routine activity theory', which builds on simple observations regarding the conditions under which a crime takes place. A fundamental idea is that a crime can only occur under the convergence in space-time of three factors - a motivated offender, a suitable target, and the absence of a capable guardian (Cohen and Felson 1979) - and the question therefore naturally shifts to how such a concurrence might arise in a realistic setting. A popular explanation is built on the hypothesis that the majority of crime is essentially opportunistic; that offenders encounter targets whilst going about non-criminal activities and that crime patterns are simply a manifestation of heterogeneities in this target awareness. Conditions suitable for crime might therefore be best conceptualised as a result of the cumulative activity patterns of the public.

Building on this, 'pattern theory' (Brantingham and Brantingham 1993a) seeks to add more detailed geographic considerations to these explanations. It is based on the idea of situating the above concepts in a realistic urban environment, considering both how activity patterns are shaped by the configuration of space (*i.e.* the 'urban form') and how the physical characteristics of certain areas can influence the decisions of a potential offender. The term 'urban backcloth' is used in this context to refer to the layout of the built environment, with particular emphasis on those elements which relate to common activities (*e.g.* homes and workplaces), and the inferred activity patterns can then be reconciled with levels and types of criminality. Going further, and somewhat anticipating the use of network theory in crime analysis, the notions of "nodes, paths and edges" (Brantingham and Brantingham 1993b) are introduced as a means of encoding the urban backcloth and the significant features within it.

The role of the street network

The definition of the urban backcloth as the configuration of urban space immediately invites consideration of the street network, since this is the primary means by which towns and cities are arranged. Perhaps due to a lack of suitable geographical data, though, little work has sought to compare crime levels with features of the street network. Early work took a coarse-grained approach, comparing crime levels at the area level with street network density and finding a positive relationship with burglary risk (Bevis and Nutter 1977). Moving to a more local level, later research used the street segment as the unit of analysis, classifying each according to the number of roads which connected to it directly and using the road type as a proxy for its 'flow' (Beavon et al. 1994). These factors both showed a positive relationship with crime, supporting the theory that more permeable roads were at greater risk (Newman 1972).

The distribution of crime has also been considered in the context of the more general concept of 'space syntax'; an approach which has been applied in a variety of urban contexts and one which has been influential in arguing for the importance of incorporating street network constraints in urban-level social models. The approach itself is characterised by the principle that individual street segments cannot properly be understood in isolation, but must be considered in the context of the rest of the network and their position within it. In addition, the role of sight-lines in governing connectivity is emphasised, and indeed these are used in the definition of street segments themselves. Several metrics have been developed using these ideas, such as *integration*, which measures how close a given location is to all others in terms of paths through the network. Applying these methods to crime, and analysing data from urban areas in the UK and Australia, Hillier (2008) found that crime was positively related to connectivity but negatively related to integration. The conclusions drawn from this contradict the previous work, suggesting that permeable designs are favourable, but that where redundant connectivity is present (that which does not increase integration) the effect can be reversed, perhaps because of the provision of extra entry or escape routes.

Hillier's work also includes observations at the level of individual properties, in particular relating to modes of access (*e.g.* proximity to alleys). These are also considered by Armitage (2007), whose work involved detailed assessment of the physical features of individual houses. These observations included the type of road on which the property was situated and a subjective estimate of its usage, both of which were then individually compared with crime levels. In this case it was found that increased activity and permeability was associated with higher risk,

with the difference between isolated cul-de-sacs and those serviced by pedestrian alleys being a notable example.

Most recently, a more nuanced approach, employing a multi-level statistical approach to account for spatial nesting in the data, again found a positive relationship between permeability and burglary risk (Johnson and Bowers 2010). In that study, the permeability of roads was evaluated either by the Ordnance Survey classification (a categorical factor, *e.g.* 'Minor Road', 'B Road') or by the number of roads to which they connected, and cul-de-sacs were treated as a special case and identified manually. The study found a steady increase in effect when moving through the hierarchy from minor roads to major, and also a notable finding that cul-de-sacs in particular were found to be at significantly lower risk.

As well as simply explaining some variation in crime rates, it has also been suggested by Johnson and Bowers (2007) that street networks may play a role in the spread of crime risk. Particularly considering the movements of offenders, and the hypothesised role of awareness spaces in repeat victimisation, it is suggested that networks are a natural substrate for such diffusion processes.

Recent advances in geographical information systems have, one one hand, facilitated modelling approaches which explicitly incorporate network data (Groff 2007b), but have also been accompanied by the development of the field of network science, which uses ideas from graph theory to facilitate sophisticated analysis of real-world networks (Newman 2010). The study of spatial networks, defined as those whose features are embedded in space in some sense, is a particularly active sub-field (Barthélemy 2011) and includes the study of various properties of street networks (Porta et al. 2006a; 2006b). In order to discuss relevant results of this research, we first introduce some graph-theoretical terminology.

The mathematics of networks

Networks, to the extent that they will be used in this paper, are relatively simple mathematical objects and many of the concepts used in their analysis are formalisations of intuitive ideas. In the most basic terms, they are collections of points and lines, where each line connects a pair of points, and typical analysis is no more complicated than considering how the network can be traversed by travelling along lines. To facilitate mathematical analysis, though, it is helpful to define these concepts symbolically.

A *network* (or *graph*) $G = (V, E)$ is a collection of *nodes*, V, and *links*, E (also referred to as *edges* within graph theory). The set of nodes, $V = \{v\}$, is a non-empty set of N elements, and $E = \{e\}$ is a set of M elements, each of which is an unordered pair of nodes. The nodes are labelled using the integers $1, \ldots, N$, where the order is unimportant as long as the labelling is consistent and unique, and each node is then referred to by its label. Where a link exists

between two nodes i and j, the nodes are said to be *adjacent* and the link is therefore represented by the unordered pair of nodes (i, j). The information necessary to describe all such links in a network can be encoded as an *adjacency matrix* **A**, an $N \times N$ matrix such that

$$a_{ij} = \begin{cases} 1 \text{ if } (i,j) \in E \text{ (\textit{i.e.} there is an edge connecting } i \text{ and } j) \\ 0 \text{ otherwise,} \end{cases}$$

(1)

where, for clarity, the symbol $\in$ denotes 'is a member of'. Various simple quantities can be defined for a node i, such as the *degree* k_i, which is the number of other nodes to which it is adjacent (*i.e.* the number of links connected to it); it is straightforward to see that this is equal to $\sum_j a_{ij}$.

A *path* in a network is any ordered sequence of nodes such that every consecutive pair of nodes is connected by a link (that is, a sequence of nodes which can be traversed by following links). The length l of such a path can be defined as the number of links which feature in it (which is 1 fewer than the number of nodes in the path). For any pair of nodes $i, j \in V$ it can be determined whether a path between the two exists, and indeed there may be more than one such path. A *shortest path* between i and j is one of these such paths of minimal length (though, again, there may be more than one such if several alternatives are equally short).

Paths can be used to calculate measures of the 'centrality' of elements of the network, in various senses, and we introduce one such here. *Betweenness centrality* is a measure which seeks to quantify how regularly individual links are used during journeys through the network. Although we will formally define it below, its meaning can be most intuitively understood by describing how it is calculated. The main steps involved are:

1) initialise all links with a betweenness centrality of 0;
2) consider all pairs of nodes i and j;
3) for each pair i and j, find the shortest path(s) between them;
4) for every link that appears in the shortest path(s), increment its betweenness centrality by $\frac{1}{w}$, where w is the number of shortest paths between i and j (so if there is only one shortest path between i and j, add 1 to the centrality of each segment in it).

Effectively, then, betweenness counts the number of times that each link is traversed, assuming that one shortest-path journey occurs between all possible pairs of nodes on the network. This can be regarded as a well-motivated proxy for the likely level of usage of each link when journeys are occurring on a network.

More formally, if we define σ_{ij} as the total number of shortest paths between i and j, and then, more particularly, $\sigma_{ij}(e)$ as the total number of shortest paths between i and j

which pass through the link $e \in E$, we can define the betweenness centrality C_e^b of a given link e as:

$$C_e^b = \sum_{i,j \in V, i \sim j} \frac{\sigma_{ij}(e)}{\sigma_{ij}} \qquad (2)$$

where $\sim$ here represents the relation 'there exists a path between i and j'. The value can then be normalised by dividing by $\frac{N(N-1)}{2}$ (the maximum possible value) in order to allow comparison between networks. Figure 1 gives a stylised example of a network for which betweenness can be used to discriminate between the roles of different links; it can be seen how the value changes between 'central' links and those on the periphery.

In order to use such ideas on street networks, we must first establish how they can be represented as graphs. There are several methods for this, with varying theoretical and mathematical implications, and the choice is dependent on the task in hand. One well-developed method, for example, represents whole roads (sets of street links which have been associated on the basis of street name or geometry) as nodes, and places links between any two which share a junction (Jiang and Claramunt 2004; Porta et al. 2006a). While the process of associating streets is useful and realistic, it can place unduly high emphasis on the importance of street name, and also leads to large variation in the physical length of streets. Perhaps a more natural choice, and the one we adopt here, is commonly known as the 'primal' representation (Porta et al. 2006b). In this, nodes are taken to represent the intersections between streets, and links represent the sections of streets which connect the intersections (each of these sections is referred to as a *street segment*, so that two nodes are connected if there is a street segment between them). An example of this can be seen in Figure 2.

Using this representation, it is possible to calculate betweenness values for each segment in a road network, and these values can also be seen in Figure 2f. The highest values are seen to occur on main thoroughfares, which would be expected to see the greatest use. As has been

pointed out elsewhere (Porta et al. 2006a), it is worth noting at this stage that betweenness is susceptible to 'edge effects': that is, the measured betweenness for segments towards the periphery of the study area is artificially low, since the starting points of many paths which would use such a segment are not included. This is unavoidable since the network considered must be geographically limited, but the problem can be ameliorated by establishing a 'buffer zone' at the extremes of the network for the purpose of betweenness calculation, which is then discounted in any following analysis.

This technical perspective on the study of street networks has developed rapidly in recent years (Crucitti et al. 2006), and the values of various measures of centrality (including betweenness) observed for many real-world networks have been well-studied. The concept of centrality can be defined in several well-motivated ways (concerning, for example, either the accessibility of places or their closeness to others) and particular metrics have been proposed to emphasise different perspectives. Building on these, the technique of 'Multiple Centrality Assessment' has been developed by Porta et al. (2006a) as a method of combining several complementary metrics in order to give an overall measure of the centrality of parts of a street network. Taking this to the natural next step, and of particular relevance for this work, the same authors then sought to establish the relationship between these measurements and urban activities, looking specifically at the densities of retail and service premises in Bologna, Italy (Porta et al. 2009) and Barcelona, Spain (Porta et al. 2012). In both cases, a positive relationship was found between economic activity and street centrality, and such results suggest that the use of metrics such as these to predict urban activity is likely to be fruitful.

The purpose of this paper is therefore to propose an explicitly mathematical model (that is, based on differential equations) for burglary, situated on a network. After describing this model, we provide examples of its behaviour on realistic networks under various scenarios, corresponding to potential real-world uses or policy interventions. With a model such as this, interest is both
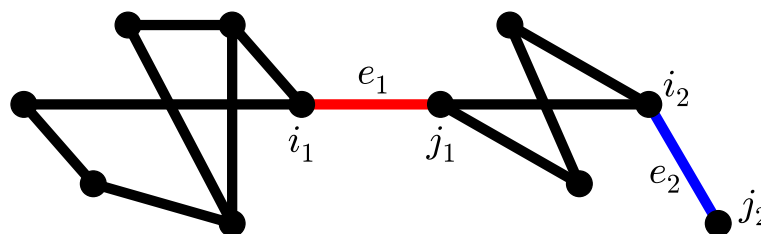

Figure 1 Simple illustration of betweenness. Link $e_1 = (i_1, j_1)$ (shown red) features in any path between a node on the 'left' of the network (there are 6) and a node on the 'right' (of which there are 5) and therefore has a relatively high betweenness value of 60. Link $e_2 = (i_2, j_2)$ (blue), on the other hand, is only traversed by paths starting or ending at j_2; there are 20 such paths and it therefore has a relatively low betweenness value of 20.
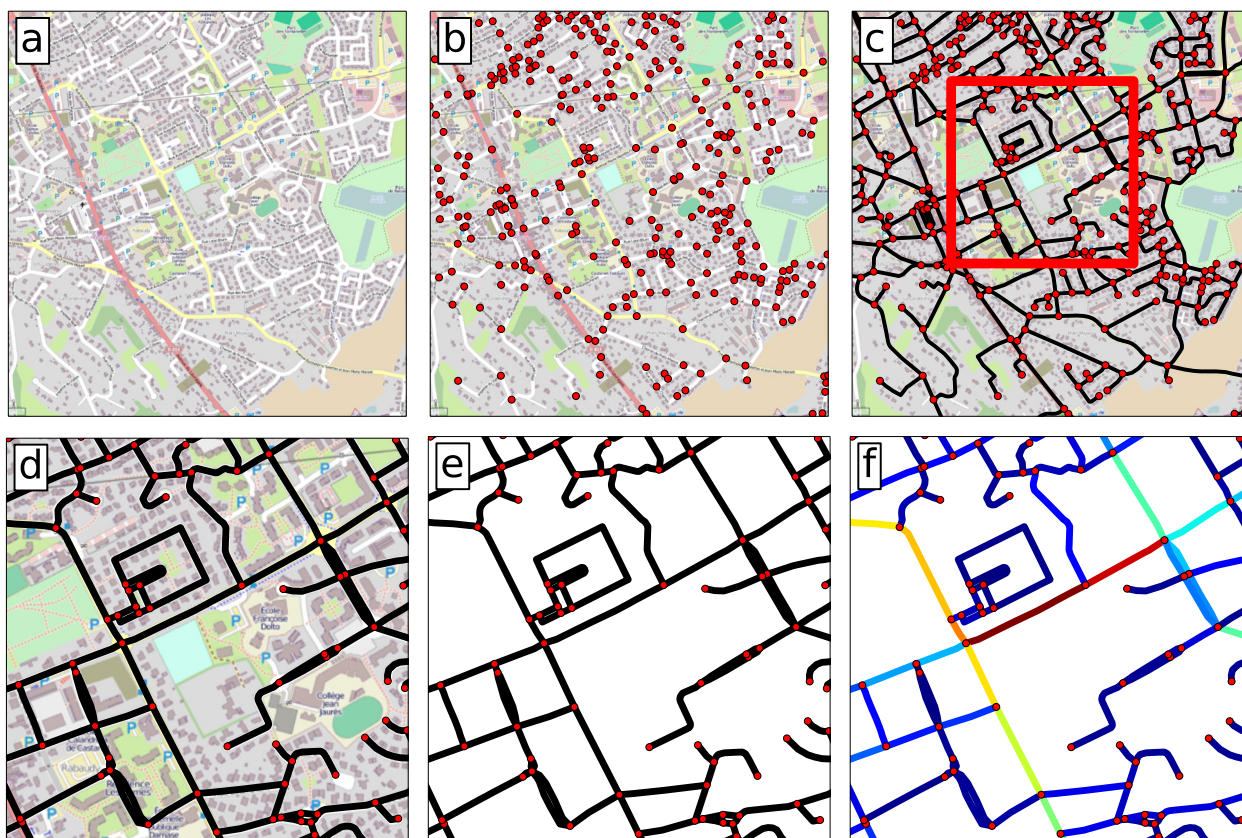
Figure 2 The construction and analysis of the 'primal' representation of a street network. This example shows a section of the network of Toulouse, France: **a)** the original street network map, as obtained from OpenStreetMap.org, **b)** nodes placed at every intersection between streets, **c)** links added between any pair of intersections which are connected by a street segment, **d)** map zoomed to the section highlighted in red in (c), **e)** background map image removed to isolate network structure, and **f)** links coloured according to betweenness (where blue is low and red is high).

theoretical and practical, necessitating different types of analysis. On the theoretical side, we aim to generate known generic patterns of crime, but the rare nature of burglary events (and multitude of factors not considered) also means that a comprehensive model may not be a sensible goal. From a practical perspective, however, the question of the risk diffusion patterns associated with individual shocks (*i.e.* burglary events) is likely to vary according to network structure and is of clear practical interest. By considering the qualitative behaviour of the model under both these motivating cases, we demonstrate the importance of considering the street network explicitly in crime modelling, and the significant effect which such consideration has on results.

Methods

Street network data
Street network data has been obtained from OpenStreetMap.org, a collaborative open-source mapping project. The data provide information relating to all roadways in a given area (according to a broad definition, which includes footpaths and private access roads, for example) and various levels of information about each, such as the road's name, type and how it is used. An exported file is processed in order to arrive at the primal graph representation, and cleaned in order to remove features, such as roundabouts, which might distort analysis.

Model specification
In previous attempts to model burglary mathematically, the fundamental spatial unit has been taken notionally to be the individual property, the implication therefore being that risk diffuses from house to house via intermediate properties. Here we use a coarser scale and take the street segment as the basic unit, considering therefore only segment-to-segment diffusion, and this is done for several reasons. Firstly, this is a scale at which real-world burglary data is widely available, and is also the scale at which potential police interventions (*e.g.* patrolling) are likely to be implemented. In addition, the determination of insurance premiums - which this model might also help to inform - is often based on postal codes, which typically

correspond to street segments; this is the case in the UK, for example. From a more practical mathematical perspective, the street segment is a suitable unit since it is at that scale that the majority of variation in centrality indices, such as betweenness, takes place. If the individual property was used instead, all properties on a given segment would have exactly the same betweenness value: values are based on journeys, and for two adjacent properties on a given segment, the journeys which pass one property are exactly those which pass its neighbour. Individual properties are therefore indistinguishable on that basis, and so their inclusion represents unnecessary redundancy in the present setting.

The model is therefore defined on a network $G = (V, E)$, constructed as described in the previous section, so that each link represents a street segment. For convenience, we introduce an indexing $e_1, \ldots, e_M$ of the links and also define the matrix $\mathbf{A}' = (a'_{ij})$, where

$$a'_{ij} = \begin{cases} 1 \text{ if } e_i, e_j \in E \text{ share a common node } v \in V \\ 0 \text{ otherwise.} \end{cases}$$

(3)

This matrix therefore encodes the structure of the network in terms of which pairs of links are coincident; in fact, it is just the adjacency matrix of the 'line graph' of G, in which the roles of nodes and edges are inverted (as described by Diestel 2010). As such, it does not introduce any further statistical properties beyond those of G itself, but simply allows for brevity in mathematical representations.

We model burglary risk $R_i(t)$ at time t for each link e_i of G, and this quantity represents our main dependent variable. Specifically, it is the rate of burglary at the given point in space and time. Each link e_i can also be assigned *a priori* two values: a fundamental *basic attractiveness* B_i (summarising static features likely to have an influence on burglary, such as security and affluence) and a *centrality* measure C_i (such as betweenness) determined by the network structure. The notion of centrality is intended to be a broad one, and refers to the extent to which a given street features in activity patterns. Many such measures could be used within this framework; for the sake of concreteness when considering the model, it can be thought of as betweenness.

Inspired by Short et al. (2008), we take the burglary rate on a link to be composed of two parts: a static component S_i and a dynamic component $Q_i(t)$:

$$R_i(t) = S_i + Q_i(t)$$

(4)

Regarding the static rate, for the purposes of this model we regard this as being dominated by opportunistic burglary and therefore driven by a combination of activity patterns and target attractiveness. At the time scale we

consider - that over which repeat victimisation effects are seen - both activity and attractiveness can be taken to be constant (though spatially varying).

To calculate the static rate S_i, we require an estimate of the level of movement activity W_i on a link e_i, and the probability p_i that an opportunist offender would victimise the link in a given time unit, were they to be present. Criminological theory implies that the means by which the network will influence crime patterns is by shaping the activity patterns of individuals. In contrast to the random walk assumption used in previous work, this assumes that people are taking specific trips, between home and work for example, and that patterns are driven by the accumulation of these trips. One approach to modelling this might therefore be to expect that such patterns at segment level are revealed as emergent properties of offender behaviour, but including this is clearly very challenging in a simple mathematical framework (as opposed to within, say, an agent-based model).

The notion of betweenness, however, whilst typically thought of purely as defining a graph metric, could in fact be regarded as a model of sorts for the throughput of a segment when traffic is flowing on the equivalent street network. In this light, its direct inclusion in the model is not necessarily artificial, but rather represents a submodel of typical pedestrian behaviour. For our measure of activity, therefore, we use some function f of centrality (the value itself could be used, but W_i may need to be a modified/smoothed quantity due to the inherent non-linearity of betweenness); thus

$$W_i = f(C_i).$$

(5)

Turning to the probability of offending, we model this as a Poisson process (a series of events occurring probabilistically in continuous time), the rate of which is determined by the basic attractiveness B_i defined previously, implying a series of independent offender decisions based on the underlying attractiveness. Accordingly, the probability within a given time period $[t, t + \delta t)$ is

$$p_i = 1 - e^{-B_i \delta t}.$$

(6)

Leaving aside the precise form of this relationship, the key point is that an increase in the attractiveness B_i will raise the probability of an offence: the exponential term $e^{-B_i \delta t}$ decreases, and p_i becomes closer to 1 (which corresponds to the offence occurring with certainty).

To account for the fact that not all patterns of offending (and activity) will be accounted for by street network properties, we also include a constant *background rate* of activity, D_i, to encapsulate other constant factors not directly related to the network. We therefore have:

$$S_i = (D_i + W_i)p_i,$$

(7)

Turning now to the dynamic component of risk $Q_i(t)$, this term is intended to encapsulate 'boost' effects; that is, those acting locally and at short time scales in response to previous victimisations. This value, for a given link, will again be composed of several parts: that which is due to 'new' offending taking place; remaining boosts from earlier time; and boost effects acquired from neighbouring locations. For the first term - new offending - the boost will simply be proportional to the offending rate $R_i(t)$ on the link, according to some boost parameter Γ which simply determines the magnitude of the effect. The latter two of the three terms have previously been modelled as a diffusion process with decay (Short et al. 2008), and are done so again here. These can be put together to form a differential equation - an equation for the rate of change of Q_i with time, denoted $\frac{dQ_i}{dt}$ - thus:

$$\frac{dQ_i}{dt} = \Gamma R_i - \omega Q_i + \eta \sum_j a'_{ij}(Q_j - Q_i) \qquad (8)$$

where ω is a decay parameter, η a coefficient of diffusion and the a'_{ij} in the summation are elements of $\mathbf{A}'$. On the right hand side, the first term ΓR_i represents the increase in proportion to new offending, and $-\omega Q_i$ determines that Q_i decays in proportion to its present value. The latter term represents diffusion from link to link whenever both meet at a common node (recalling that $a'_{ij} = 1$ if and only if i and j are coincident). Specifically, it encodes the idea that risk will flow from higher- to lower-risk segments: if segment i has a neighbour j which is at higher risk (*i.e.* $Q_j > Q_i$), the term $a'_{ij}(Q_j - Q_i)$ will be positive and drive an increase in Q_i. The summation simply averages this effect over all neighbours.

Naturally, R_i can be re-written as $(S_i + Q_i)$, and S_i in turn as $(D_i + W_i)p_i$, so that we have a single dynamical equation for Q_i:

$$\frac{dQ_i}{dt} = \Gamma((D_i + W_i)p_i + Q_i) - \omega Q_i + \eta \sum_j a'_{ij}(Q_j - Q_i). \qquad (9)$$

Such a differential equation can be used to carry out numerical simulations, and the results of several of these will be shown in the following sections. To perform these, the differential equation is repeatedly applied at regular discrete time-steps, with parameters as specified in figure captions in each case. Unless otherwise specified, street network data from Toulouse, France, is used as the spatial setting, as depicted in Figure 2. For each link, the equation for the rate of change, when coupled with the initial condition, is sufficient to determine the state of the system at all subsequent times.

Results and discussion

Response to burglary events

In the above formulation, risk is modelled as increasing according to current levels of offending via the 'boost' term with parameter Γ. The purpose of this is to represent an evolving system with ongoing offending and ultimately to find persistent patterns of crime; however, also of interest in this scenario is the question of how the risk associated with known, exogenous offending will evolve over a short timescale. One way of examining this is to consider the response to a hypothetical burglary event occurring at a particular location, in the absence of further offending. Removing the boost term from the model (or, indeed, setting $\Gamma = 0$) and setting initial conditions to represent some state of the system, we can see how the patterns of risk would evolve. The diffusion-decay model in this case is simply

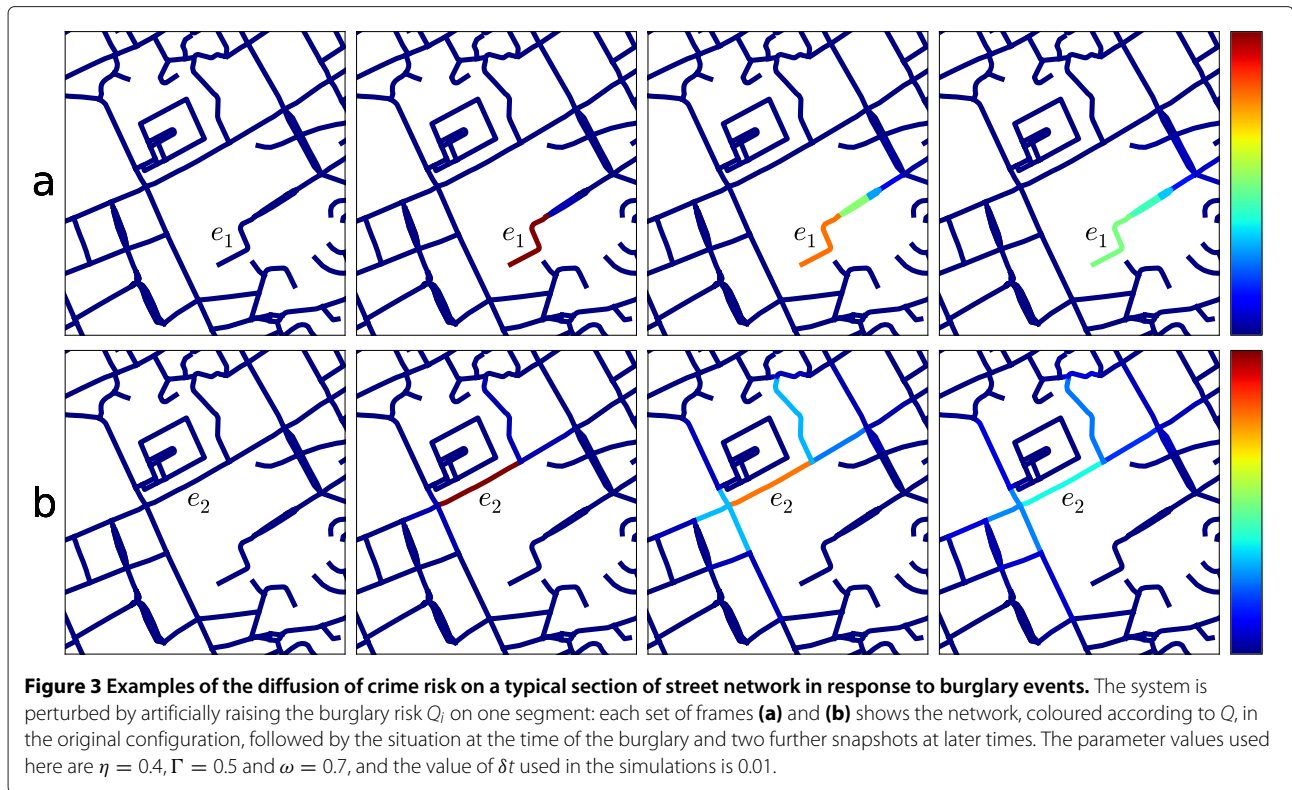$$\frac{dQ_i}{dt} = -\omega Q_i + \eta \sum_j a'_{ij}(Q_j - Q_i). \qquad (10)$$

Figure 3 shows the diffusion of risk after the system has been perturbed in a single location (representing a burglary event). Although this has little to offer in terms of providing an explanation for hotspot formation, it demonstrates the potential use of a model such as this as a predictive tool: given a recent pattern of offending, a quantitative indication of the local change in risk owing to that offending can be found. In this case, we can see the risk ensuing from a single victimisation.

It is notable that when a relatively non-central segment is victimised in Figure 3a, the spread is localised and one-directional, whereas risk spreads in a much more diverse manner when a highly central segment is victimised in Figure 3b. It can also be seen that, when segments are far apart in terms of network distance, there is negligible spread of risk from one to the other, even if they are close in purely spatial terms; this is contrary to what is predicted by models which do not incorporate the street network, and arises simply because risk is constrained to spread only along links. This demonstrates a fundamental sense in which this model differs from those proposed previously.

Modelling dynamic burglary events

We now move on to investigate numerically the behaviour of the system when levels of offending feed back into the system; that is, when crime is not taken as an exogenous initial condition, but the ongoing offending originates from within the model itself. We therefore consider the case where $\Gamma \neq 0$.
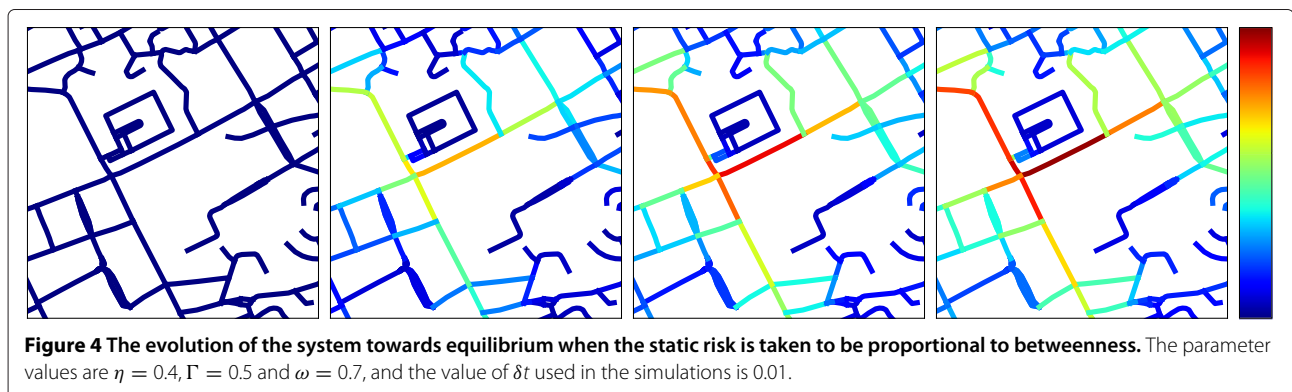
The long-term behaviour of the model under a relatively simple formulation is shown in Figure 4. In this simulation, the static attractiveness S of each street segment (*i.e.*

Figure 3 Examples of the diffusion of crime risk on a typical section of street network in response to burglary events. The system is perturbed by artificially raising the burglary risk Q_i on one segment: each set of frames **(a)** and **(b)** shows the network, coloured according to Q, in the original configuration, followed by the situation at the time of the burglary and two further snapshots at later times. The parameter values used here are $\eta = 0.4$, $\Gamma = 0.5$ and $\omega = 0.7$, and the value of δt used in the simulations is 0.01.

each link in the network) is taken to be directly proportional to its betweenness centrality C^b, in line with the reasoning outlined previously. It is also assumed that the basic attractiveness of street segments, B, is uniform but non-negative across the system, and the background rate D is taken to be 0 everywhere, in order to concentrate on repeat-victimisation effects. In addition, the system is initialised with a uniform initial condition of 0 for Q. With initial values chosen in this way, and for suitable values of the other parameters, it is seen numerically that the system tends towards an equilibrium (in that the value for each link reaches a steady state), with the final state shown in the final frame of Figure 4. The fact that equilibrium is reached implies that, for the parameters chosen, the final

configuration shown is a sustainable pattern for the long-term distribution of crime; the value for each link is its long-term rate of crime.

Although comparison with Figure 2f shows that the links with highest value of Q are, as expected, those with highest betweenness centrality C^b, it is also seen that it is not the case that there is a direct relationship between the two. This can be seen more explicitly in Figure 5, in which betweenness values are plotted against the equilibrium values of risk. Behaviour of this form suggests that the diffusive pattern adds structure to the crime patterns, even in steady state, over and above that which would be predicted purely on the basis of the betweenness-based static risk (in which case a straight line would be expected
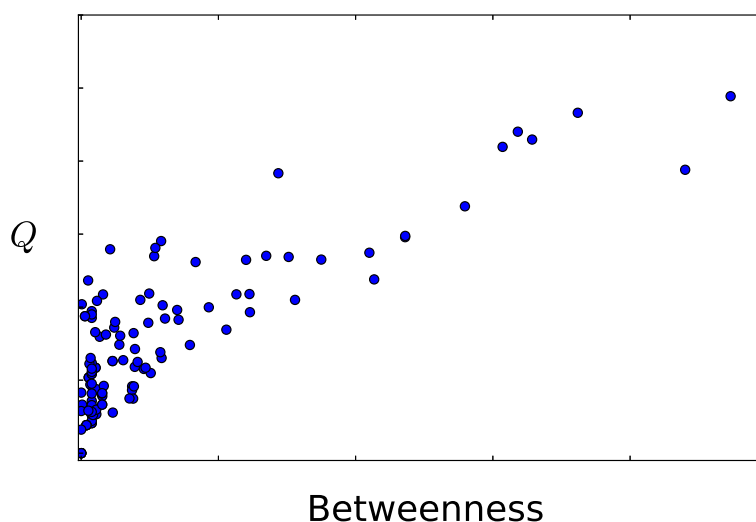

Figure 4 The evolution of the system towards equilibrium when the static risk is taken to be proportional to betweenness. The parameter values are $\eta = 0.4$, $\Gamma = 0.5$ and $\omega = 0.7$, and the value of δt used in the simulations is 0.01.

Figure 5 The relationship between centrality and burglary risk. This shows a scatter plot of betweenness values against equilibrium values of Q for the simulation depicted in Figure 4. Each point corresponds to a link in the network.

in Figure 5). Indeed, streets must be considered explicitly in the context of those around them and in the context of the urban space as a whole. In the configuration shown, because of heterogeneities in the way the network is connected, the underlying risk is 'smoothed' in the equilibrium.

Modelling policing interventions

A natural progression from the study of the long-term emergence of crime patterns is to consider how these react when the system is manipulated; considering this as a manifestation of a real-world intervention, this is of immediate relevance to policy. A natural intervention to study is the activity of police, via patrolling or otherwise, which is intended to have the direct effect of reducing crime in the area in which it is employed. To this end, we investigate the numerical behaviour of the model under such an intervention.

Modelling the effect of police activity in the simplest way possible - by artificially lowering the crime rate to zero on a chosen street segment - we can examine the effect on the long-term behaviour of the system. Figure 6 shows such a simulation, where the street segment with highest betweenness has been selected for intervention. From the colouring it can be seen that the risk has been reduced not only on the segment with intervention, but on other segments also. Moreover, it is apparent that the reduction on nearby segments is not simply a function of their distance from the chosen segment; again, the complexity of the network structure has caused effects which are non-linear in this sense. Figure 7 shows the ratio of post- to pre-suppression Q for segments according to their distance from the intervention, and it can be seen that there is considerable variance within groups: even among direct neighbours, the reduction is more pronounced in some than others. This again emphasises the way in which

Figure 6 The evolution of the system when crime risk is artificially suppressed on one segment. The suppression is achieved by holding the value Q at 0 on e for the duration of the simulations, with all other aspects of the model unchanged. Segments are coloured according to Q, the parameter values here are $\eta = 0.4$, $\Gamma = 0.5$ and $\omega = 0.7$, and the value of δt used in the simulations is 0.01.

Figure 7 Variation in the effect of suppression. The graph shows the magnitude of reduction of Q on nearby segments resulting from the artificial suppression of crime on the link with highest betweenness, as shown in Figure 6. Values show the ratio of Q after suppression to that in its absence, and are grouped according to the distance of the segment in question from the directly-affected segment, e.
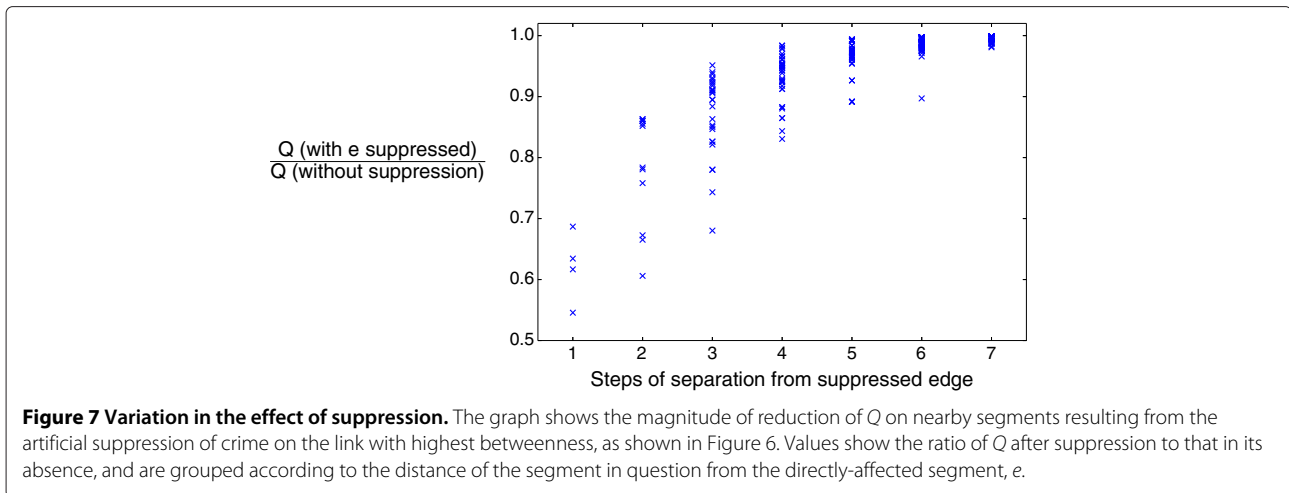
interventions must be considered in their proper spatial context, and is encouraging since it suggests that informed placement has the potential to compound the positive effect of policing.

Modification of urban form

In addition to short-term interventions such as policing, one of the potential outcomes of analysing urban crime is to inform decisions taken by planners in relation to urban design. This might take the form of developing heuristics for future planning projects, but there is also the possibility to modify existing structures in order to address a crime problem.

With this in mind, we address such a possibility; specifically, modification of the street network by removal of a certain link. This removal could represent the blocking of that segment or, equivalently, the state of the network if that segment had never originally been constructed. A natural choice for such an intervention, given the role of network centrality in the model as a driver of crime, is the segment with highest betweenness, and the removal of this segment for Toulouse is illustrated in Figure 8.

In Figure 9, we show the evolution of the system after the removal of this link from the original network. Once again, we initialise Q as zero on every link and allow the system to reach equilibrium. It can be seen that, relative to the example in Figure 4, burglary activity is displaced in an unanticipated way; the peak which was previously seen on the removed segment has not simply moved to neighbouring segments, but has appeared several segments away. The fact that such a central link has been removed means that its role in the network must be assumed by other links: in terms of betweenness, many paths passed through the deleted link and must be re-routed, with these routes possibly being quite distinct from the originals (in terms of the number of links shared by both the original and re-routed paths). This can cause a dramatic change

in the betweenness of other links, perhaps several degrees of separation away, and this effect is seen in the changed pattern of crime in Figure 9, where the main 'hot spot' of criminal activity has moved to the upper left of the map.

Variation between networks

The structure of street networks can vary widely both within urban areas and, at a larger scale, between cities and countries. Depending on its intended use, and on the dominant planning practices at the time of construction, the properties of networks can be measurably different. Given that our model is based partially on street-level metrics, such variation will clearly influence results. As well of being of theoretical interest, it also means that the model may have different implications according to local circumstances.

Figure 10 shows results for a section of the street network of Santiago, Chile, which has a grid-like structure and is of clearly different character to that of Toulouse (Figure 2). This structure means that travel loads are more
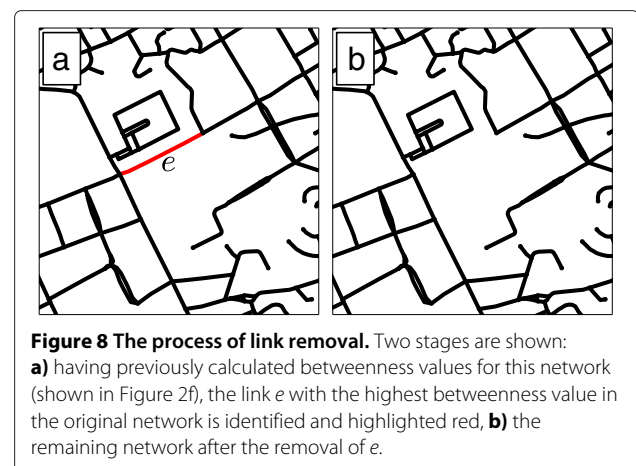

Figure 8 The process of link removal. Two stages are shown: **a)** having previously calculated betweenness values for this network (shown in Figure 2f), the link e with the highest betweenness value in the original network is identified and highlighted red, **b)** the remaining network after the removal of e.

Figure 9 The evolution of the system when one link is removed from the network. The removal is as demonstrated in Figure 8, and betweenness values are recalculated accordingly. As previously, segments have been coloured according to the value of Q. The parameter values are $\eta = 0.4$, $\Gamma = 0.5$ and $\omega = 0.7$, and the value of δt used in the simulations is 0.01.

evenly distributed - the betweenness centrality values shown in Figure 10a) are generally not as high as those for Toulouse and show little variation across the network. The effect of this for the model is that no pronounced peaks of burglary risk are predicted, as seen in Figure 10b) where the equilibrium state is shown. In this case the effect is not caused by the diffusion on the network, but rather by the lack of variation in underlying risk caused by travel patterns. In a situation such as this, patterns of pedestrian flow have less predictive value in terms of crime risk, since streets can be less clearly distinguished on that basis.

Conclusions

Theories of environmental criminology emphasise the importance of urban configuration in determining patterns of crime, since it is the primary determinant of human movement patterns. The street network is the primary structure by which this configuration is determined, and there is evidence that it shapes patterns of crime. Despite recent interest in the mathematical modelling of burglary, however, few models have sought to explicitly include network effects. We have presented a novel mathematical model for residential burglary which takes such

effects into account. The influence of the street network is manifested in two ways: by restricting the spread of crime to only occur along network connections, and by incorporating network metrics as a proxy for human activity.

We have presented several stylised examples demonstrating the qualitative behaviour of the model. These examples correspond to scenarios for which a model such as this are likely to be of use, ranging from 'real-time' predictive circumstances to the general analysis of policy interventions. In all cases, the effect of the network is evident: there is a marked difference between model results and what would have been predicted by a non-network model, and the non-linear nature of network dynamics are illustrated.

The consideration of street networks in models of crime is well-motivated, and the results shown here illustrate its importance. From a practical point of view, this is encouraging, since networks (the properties of which are quantifiable) represent another means by which crime prevention efforts can be concentrated. Indeed, the non-linear effects shown here suggest that full understanding of network effects may amplify the effect of targeted policing.

The potential use of the model presented here is, of course, dependent on the existence of a relationship between network properties and crime rates of the expected form. Although an empirical base has been established in support of a general relationship, analysis has not yet been conducted using metrics of the type considered in this paper. Future research will involve the exploration of this question via statistical analysis, as well as considering the influence of network configuration on the phenomenon of near-repeat victimisation. This would then provide a basis for the implementation of the model proposed here in a practical setting. Research addressing the question of how the output of such predictive models can best be translated into police activity is currently ongoing, and represents the crucial stage in the evolution from abstract models to practical outcomes.
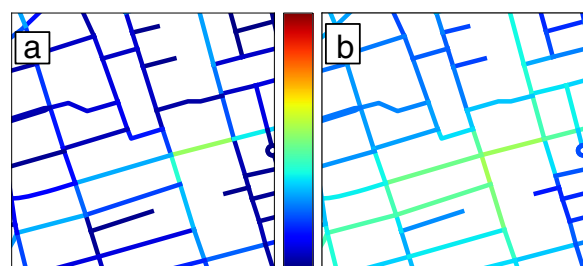

Figure 10 The example of Santiago, Chile. Figure **a)** shows betweenness centrality values, plotted on the same colour scale as used for Toulouse in Figure 2f; values in this case are generally not as high and vary little. Figure **b)** shows equilibrium values of Q for this network, for an equivalent simulation to that shown in Figure 4; again, the colour scale is the same as used in 4.

References
Armitage, R (2007). Sustainability versus safety: confusion, conflict and contradiction in designing out crime. In G Farrell, KJ Bowers, SD Johnson, M Townsley (Eds.), *Imagination for crime prevention* (pp. 81–110). Monsey: Criminal Justice Press.

Ashton, J, Brown, I, Senior, B, Pease, K (1998). Repeat victimisation: offender accounts. *International Journal of Risk Security and Crime Prevention*, *3*(4), 269–279.

Barthélemy, M (2011). Spatial networks. *Physics Reports*, *499*(1–3), 1–101.

Beavon, DJK, Brantingham, PL, Brantingham, PJ (1994). The influence of street networks on the patterning of property offences. In RV Clarke (Ed.), *Crime prevention studies, vol. 2*, (pp. 115–148). Monsey: Criminal Justice Press.

Berestycki, H, & Nadal, J-P (2010). Self-organised critical hot spots of criminal activity. *European Journal of Applied Mathematics*, *21*(4-5), 371–399.

Bernasco, W (2008). Them again? *European Journal of Criminology*, *5*(4), 411–431.

Bevis, C, & Nutter, JB (1977). *Changing street layouts to reduce residential burglary*. Paper presented to the American Society of Criminology annual meeting in Atlanta. https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=51937

Birks, D, Townsley, M, Stewart, A (2005). Generative explanations of crime: using simulation to test criminological theory. *Criminology*, *50*(1), 221–254.

Bowers, KJ, Johnson, SD, Pease, K (2005). Victimisation and revictimisation risk, housing type and area: A study of interactions. *Crime Prevention and Community Safety: An International Journal*, *7*(1), 7–17.

Brantingham, PJ, & Brantingham, PL (1981). *Environmental criminology*. Beverly Hills: Sage Publications.

Brantingham, PL, & Brantingham, PJ (1993a). Environment, routine and situation: Toward a pattern theory of crime. *Advances in Criminological Theory*, *5*, 259–294.

Brantingham, PL, & Brantingham, PJ (1993b). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, *13*(1), 3–28.

Brantingham, PL, Brantingham, PJ, Vajihollahi, M, Wuschke, K (2009). Crime analysis at multiple scales of aggregation: A topological approach. In D Weisburd, W Bernasco, GJ Bruinsma (Eds.), *Putting crime in its place* (pp. 87–107). New York: Springer.

Chainey, S, & Ratcliffe, J (2005). *GIS and crime mapping, vol. 6*. Chichester: John Wiley & Sons Inc.

Cohen, LE, & Felson, M (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, *44*(4), 588–608.

Cornish, D, & Clarke, R (1986). *The reasoning criminal: rational choice perspectives on offending*. New York: Springer Verlag.

Cromwell, PF, Olson, J, Avary, D (1991). *Breaking and entering: an ethnographic analysis of burglary, vol. 8*. Newbury Park: Sage.

Crucitti, P, Latora, V, Porta, S (2006). Centrality measures in spatial networks of urban streets. *Physical Review E*, *73*(3), 036125.

Davies, TP, Fry, HM, Wilson, AG, Bishop, SR (2013). A mathematical model of the London riots and their policing. *Scientific Reports*, *3*, 1303.

Diestel, R (2010). *Graph theory*. New York: Springer-Verlag.

Eck, J, & Weisburd, D (1995). Crime places in crime theory. In J Eck & D Weisburd (Eds.), *Crime and place,* (pp. 3–31). New York: Criminal Justice Press.

Farrell, G (2005). Progress and prospects in the prevention of repeat victimization. In N Tilley (Ed.), *Handbook of crime prevention and community safety,* (p. 143–170). Cullompton: Willan.

Groff, ER (2007a). Simulation for theory testing and experimentation: An example using routine activity theory and street robbery. *Journal of Quantitative Criminology*, *23*, 75–103.

Groff, ER (2007b). "Situating" simulation to model human spatio-temporal interactions: an example using crime events. *Transactions in GIS*, *11*(4), 507–530.

Grubesic, TH, & Mack, EA (2008). Spatio-temporal interaction of urban crime. *Journal of Quantitative Criminology*, *24*(3), 285–306.

Hillier, B (2008). Can streets be made safe? *Urban Design International*, *9*(1), 31–45.

Jiang, B, & Claramunt, C (2004). Topological analysis of urban street networks. *Environment and Planning B: Planning and Design*, *31*(1), 151–162.

Johnson, SD (2008). Repeat burglary victimisation: a tale of two theories. *Journal of Experimental Criminology*, *4*, 215–240.

Johnson, SD, & Bowers, KJ (2007). Burglary prediction: theory, flow and friction. In G Farrell, KJ Bowers, SD Johnson, M Townsley (Eds.), *Imagination for crime prevention* (pp. 203–223). Monsey: Criminal Justice Press.

Johnson, SD, & Bowers, KJ (2010). Permeability and burglary risk: are cul-de-sacs safer? *Journal of Quantitative Criminology*, *26*(1), 89–111.

Johnson, SD, Bowers, KJ, Hirschfield, A (1997). New insights into the spatial and temporal distribution of repeat victimization. *British Journal of Criminology*, *37*(2), 224 –241.

Johnson, SD, Bernasco, W, Bowers, KJ, Elffers, H, Ratcliffe, J, Rengert, G, Townsley, M (2007). Space–Time Patterns of risk: a cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology*, *23*(3), 201–219.

Johnson, SD, Bowers, KJ, Birks, DJ, Pease, K (2009a). Predictive mapping of crime by ProMap: accuracy, units of analysis, and the environmental backcloth. In D Weisburd, W Bernasco, GJ Bruinsma (Eds.), *Putting crime in its place,* (pp. 171–198). New York: Springer.

Johnson, SD, Summers, L, Pease, K (2009b). Offender as forager? A direct test of the boost account of victimization. *Journal of Quantitative Criminology*, *25*, 181–200.

Laycock, G (2005). Defining crime science. In MJ Smith & N Tilley (Eds.), *Crime science: new approaches to preventing and detecting crime,* (pp. 3–24). Uffculme: Willan.

Malleson, N, Evans, A, Jenkins, T (2009). An agent-based model of burglary. *Environment and Planning B: Planning and Design*, *36*(6), 1103–1123.

Mohler, GO (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, *106*(493), 100–108.

Morgan, F (2001). Repeat burglary in a Perth suburb: indicator of short-term or long-term risk? In GF Pease (Ed.), *Repeat victimisation: crime prevention studies, vol. 12*. Monsey: Criminal Justice Press.

Newman, O (1972). *Defensible space; crime prevention through urban design*. New York: Macmillan.

Newman, MEJ (2010). *Networks: an introduction*. Oxford: Oxford University Press.

Pease, K (1998). *Repeat victimisation: taking stock*: Home Office Police Research Group Crime Detection and Prevention Series, Paper 90. London, UK: Home Office.

Pitcher, AB (2010). Adding police to a mathematical model of burglary. *European Journal of Applied Mathematics*, *21*, 401–419.

Pitcher, AB, & Johnson, SD (2011). Exploring theories of victimization using a mathematical model of burglary. *Journal of Research in Crime and Delinquency*, *48*(1), 83–109.

Porta, S, Crucitti, P, Latora, V (2006a). The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications*, *369*(2), 853–866.

Porta, S, Crucitti, P, Latora, V (2006b). The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design*, *33*(5), 705–725.

Porta, S, Latora, V, Wang, F, Strano, E, Cardillo, A, Scellato, S, Iacoviello, V, Messora, R (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design*, *36*(3), 450–465.

Porta, S, Latora, V, Wang, F, Rueda, S, Strano, E, Scellato, S, Cardillo, A, Belli, E, Càrdenas, F, Cormenzana, B, Latora, L (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies*, *49*(7), 1471–1488.

Shaw, C, & McKay, H (1969). *Juvenile delinquency and urban areas: a study of rates of delinquency in relation to differential characteristics of local communities in American cities*. Chicago: University of Chicago Press.

Sherman, LW, Gartin, PR, Buerger, ME (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, *27*(1), 27–56.

Short, M, D'Orsogna, M, Pasour, V, Tita, G, Brantingham, P, Bertozzi, A, Chayes, L (2008). A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, *18*(S1), 1249–1267.

Short, M, D'Orsogna, M, Tita, G, Brantingham, P (2009). Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*, *25*(3), 325–339.

Summers, L, Johnson, SD, Rengert, GF (2010). The use of maps in offender interviewing. In W Bernasco (Ed.), *Offenders on offending: Learning about crime from criminals,* (pp. 246–272). Cullompton: Willan.

Townsley, M, Homel, R, Chaseling, J (2003). Infectious burglaries. A test of the near repeat hypothesis. *British Journal of Criminology*, *43*(3), 615–633.

Weisburd, D, Bushway, S, Lum, C, Yang, S-M (2004). Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology, 42*(2), 283–322.