Measuring connectivity patterns in a macro-corridor on the south coast of Western Australia

By Adam G. Dunn and Jonathan D. Majer

Adam Dunn was a Research Fellow with the Alcoa Research Centre for Stronger Communities and hosted by Environmental Biology, Curtin University of Technology at the time this research was undertaken. He is now a Research Fellow at the Centre for Health Informatics at the University of New South Wales (Sydney, NSW 2052, Australia; Tel: +61 (0)29385 9033; Email: a.dunn@unsw.edu.au).

Jonathan Majer is Professor of Invertebrate Conservation and Head of the Department of Environmental Biology at Curtin University of Technology (PO Box U1897, Perth, WA 6845, Australia). The project this study contributes to is part of the Alcoa Foundation’s International Conservation and Sustainability Program, of which Curtin University of Technology is one of five partners on five continents.

Introduction and context

It is understood that the heterogeneity of a landscape has significant, non-linear effects on the dynamics of an ecosystem (Wiens 1976; Turner 1989; Pickett & Cadenasso 1995), including both flora (With 2002) and fauna (Lindenmayer et al. 2005). The structure of edges (Ims 1995; Haddad 1999) and corridors within a heterogeneous landscape (specifically, a heterogeneous habitat) has potential effects that extend beyond their immediate locations (Levin 1992). Corridors are therefore valuable components of fragmented landscapes (Beier & Noss 1998).

Quantifying connectivity is regarded as a difficult problem, despite the concept of corridors being easy to understand (Li & Wu 2007). Here we present a spatially explicit and quantitative analysis of connectivity using the betweenness centrality metric (Bodin & Norberg 2007), a method traditionally used as a metric of social networks to find the power that individuals have over communication and information flow in a social network. The analysis demonstrated in this paper is carried out on the macro-corridor (Wilkins et al. 2006) on the south coast of Western Australia (currently the focus of the Gondwana Link restoration project; Watson et al. in press), represented by a vegetation survey (taken from Connell & ATA Environmental 2002). The analysis provides a map of the pathways through the corridor that are predicted to be significant to dispersal, fauna movement and gene flow.

Betweenness centrality (see Boccaletti et al. 2006), first introduced only relatively recently to landscape ecology by Bodin & Norberg (2007), belongs to a group of graph-theoretic methods (see Calabrese & Fagan 2004) for analysing corridors. Other metrics in the group include multiple least-cost paths (Pinto & Keitt 2008) and minimum-spanning trees with edge removals (Urban & Keitt 2001). Centrality (as distinct to betweenness centrality) has also previously been used as a metric for landscape connectivity (Jordán et al. 2003).

The aim here is to provide a quantitative analysis of the macro-corridor (and its component corridors) between the Stirling Ranges and the Fitzgerald River National Park (Wilkins et al. 2006), identifying the significant pathways that compose the macro-corridor and the individual corridors within the region. The presentation is devised such that the method may be replicated for other regions, for individual species within the region, and for updated vegetation surveys. A further aim is to make a recommendation for the targeted restoration and revegetation in the region, based on the pathways identified in the analysis.

Method

In essence, the manner in which important pathways are uncovered is by aggregating the results of repeated calculations of the betweenness centrality using different input parameters. The results provide a quantitative and spatially explicit view of the corridor, which may be used to support decisions regarding revegetation and restoration. A user specifies two main inputs: a map of the habitat (typically as a raster image), and a series of distances (from 100 m up to several kilometres) related to the dispersal ability of plant and animal species that are potentially likely to use the defined habitat. Using a network representation of the habitat, and for each of the distances, the betweenness centrality algorithm is applied and the aggregated output values are used to rank the importance of pathways to the connectivity through the corridor.
**Research Report**

Figure 1. The Stirling–Fitzgerald macro-corridor with three labelled corridors (see text). Shaded grey areas are habitat. The segmentation (black-edge polygons) represents the structure used in the creation of the network for the coarse-scale analysis.

**Study area**

The Gondwana Link project (Watson et al. in press) designates the region between the Stirling Ranges National Park (on the west) and the Fitzgerald River National Park (on the east) as the Fitz–Stirling Operational Area (hereafter, the Fitz–Stirling macro-corridor). Within the Fitz–Stirling, there are three corridors that have been identified to be of interest by Wilkins et al. (2006, p. 47), who list these three corridors as follows:
1. The Stirling Ranges to Corackerup/Peniup (hereafter, the Stirling–Peniup corridor; see Fig. 1, label A);
2. Corackerup/Peniup to Fitzgerald River (hereafter, the Peniup–Fitz corridor; see Fig. 1, label B); and
3. Fitzgerald to Lake Magenta (hereafter, the Fitzgerald corridor; see Fig. 1, label C).

**Betweenness centrality**

Betweenness centrality is a measure of information exchange that applies to both the edges and the nodes in a graph (as in graph theory, see Boccaletti et al. 2006). In the landscape connectivity context, nodes may be thought of as distinct patches (or cells that compose patches) and edges as the representation of foraging, dispersal, other movement or gene flow between two nodes. An edge with a high betweenness centrality value relative to other edges in the graph may be said to be important to the information exchange within the graph. This is because the high-value edge is a component of a greater number of shortest paths between cells across the entire graph. Given that least cost paths are an appropriate way to represent dispersal, movement and gene flow, the betweenness centrality metric is an appropriate metric for the quantification of connectivity through a landscape. In particular, betweenness centrality may be used to identify specific pathways that are bottlenecks within a corridor that may contain multiple redundant paths between patches.

Betweenness centrality (see Girvan & Newman 2002; Boccaletti et al. 2006; Mann et al. 2008) can be used as a metric of spatial networks to find the power that certain sites may have over species flow, just as it was used traditionally to analyse the influence of individuals in a social network. Newman (2001) implemented the betweenness centrality algorithm in O(nn) time (the number of constant time operations is proportional to the number of edges times the number of nodes), which means that it is possible to run the algorithm for reasonably large and dense networks, as presented below. Newman’s description of the algorithm provides enough information to implement it in any number of programming environments. For the analyses presented here, Matlab (The MathWorks Inc. 1992) is used because it implements optimized matrix operations and has many built-in visualization tools for presenting results.

Graph-theoretic techniques for modelling spatial processes are well-established in landscape ecology (Cantwell & Forman 1999; Urban & Keitt 2001) and have been used in the analysis of connectivity and dispersal. One method that is comparable to the betweenness centrality method is the multiple shortest path tool (Pinto & Keitt 2008), which is a method for finding the set of redundant paths that form a corridor through a landscape using weighted edges between pixels. Another comparable method is the circuit theory model (McRae et al. 2008), which is a discussion of circuit theory and random walks in their application to identifying important habitat patches and corridors. This method has been used to describe both corridor redundancy and dispersal bottlenecks, which are analogous to the concept of funnelling in general network theory (see Newman 2001).

**Analysis method**

Once the GIS mapping and species habitat definitions for the corridors are gained, each analysis is completed in five steps. First, a structure is created to represent the habitat of the landscape as a series of cells, which provides a more efficient and accurate structure than is produced by using pixels as cells (see below). Second, the network is defined using a maximum distance threshold, v, which links all cells (via edges) to their close neighbours. Third, the betweenness centrality algorithm takes the network as input to produce a set of betweenness values for each edge. Fourth, the second and third steps are repeated for different values of v. Finally, the results of all of the individual betweenness centrality instantiations are aggregated to produce a ranking of important edges through the corridor, for the given underlying structure.
The GIS information used in these analyses (from Connell & ATA Environmental 2002), is exported from its original vector (polygonal) format into three raster images. The first raster image is a coarser image of the macro-corridor that includes portions of the Stirling Ranges and the Fitzgerald River biosphere. The second and third raster images are ‘zoomed-in’ sections of the Stirling–Peniup and Peniup–Fitz corridors, with higher resolutions that include roadside and degraded riparian vegetation. Each of the raster images is a binary image. The significant assumption is that the raster images are a reasonable approximation of what constitutes potential habitat for a range of important plants and animals from the region.

The first step is completed by first distributing a set of pseudo-random points over the landscape at a specific density. We use a Halton point distribution (Halton & Smith 1964) because it produces an irregular distribution that has the appearance of randomness, but is deterministic and repeatable. We use an irregular distribution to avoid the inherent directional bias that is created when using regular grids (O’Regan et al. 2007; Schönfisch 1997; Holland et al. 2007). Second, to describe the cells and modify the granularity of homogeneous regions, we use Voronoi decomposition (Okabe et al. 2000). This process is described in more detail in an implementation of hierarchical patch dynamics (Wu & Loucks 1995), in which multiple scales of habitat are described (Dunn & Majer 2007). In the final part of the first step, cells whose originating points do not coincide with the habitat as defined by the image of the study area are removed (as illustrated in Fig. 1).

The second step is to define the network based on the locations of each of the cells. The distance from every cell to every other cell is compared to the maximum distance threshold, \( v \), and any cells whose distances are smaller than the value of \( v \) are recorded. The format of the recorded values is an adjacency matrix, which is simply a two-dimensional matrix of size \( n \times n \), where \( n \) is the number of cells in the landscape.

In the third step, the betweenness centrality algorithm is instantiated using the adjacency matrix created in the previous step. The algorithm performs a search through the network defined by the adjacency matrix and aggregates the number of times each edge features as a component of the shortest path between two cells.

In the fourth step, \( v \) is modified to produce a new network and the betweenness algorithm is instantiated with the new adjacency matrix. Increasing the value of \( v \) results in a network of greater density; any network with a higher value for \( v \) will contain more (or at least as many) links as in the network with a smaller value for \( v \). The returned betweenness values are recorded for each edge and each cell.

In the final step, the returned betweenness values from each of the repetitions of steps two and three are aggregated. Visualizing the set of all betweenness values does not always produce intuitively useful results, so an additional step is performed to find the highest ranked edges among all graphs, producing the final view as shown in the results section below.

**Parameter sensitivity and modelling issues**

For the underlying habitat patterns presented here, the betweenness value is maximized when increasing the value of \( v \) first causes the largest patches in the landscape to become connected (see Figs 2a, 2b, and 3). This corresponds to the geometric phase transition in the graph (Keitt et al. 1997). Since the aggregation is the summation of betweenness values across a range of minimum distances, the first connected pathway is often heavily represented in the final aggregation. As the minimum distance increases, other (potentially redundant) pathways begin to feature in the results (see Fig. 2c). A characteristic curve for the average and maximum betweenness may be plotted for each analysis (see Fig. 3) and this curve demonstrates the minimum distance required to traverse the corridor for a specific functional granularity. The characteristic curve additionally provides information about which values for \( v \) are most appropriate to the landscape pattern and the current functional granularity represented in the structure (capturing the geometric phase transition is important).

The method is not particularly sensitive to the choice of boundaries for these habitat profiles, within reason. Since the total habitat area inside the larger protected regions are orders of magnitude greater than the total area of habitat between them (in all three examples), small changes in boundary choices have an insignificant effect on the top-ranked edges. Boundary choice is potentially more significant in cases where the difference between corridor and patch is not clear (imagine an environment where fragmentation and habitat loss is relatively consistent throughout the region chosen). The choices of boundaries used in the experiments presented below are consistent with those used by Pinto & Keitt (2008) in a connectivity study for an environment with a similar level of overall fragmentation.

The analyses are sensitive to the granularity with which the structure is created – namely the structure may not perceive smaller patches within corridors as habitat at coarse scales. The problem corresponds to the issue of functional granularity, which is central to landscape ecology in general (Levin 1992). This is particularly important in examples where small patches of vegetation, such as roadside vegetation, play significant roles in the connectivity through a corridor. This occurs when the patches are not sampled during the Halton point allocation step (see the previous section). To ensure that all patches identified in the underlying data are captured in the structure requires sampling to be carried out at a sub-pixel (or sub-vector) level. This is computationally infeasible for the coarse scale, but is possible if sections of the landscape are analysed individually.

We circumvent the above trade-off between computational feasibility and fine-scale representation by repeating the simulations at two granularities – first at a coarse scale to reveal the pathways through the entire macro-corridor for one functional granularity, and second at finer scales and finer functional granularities, which ensure that the minor patches within the corridor are captured (particularly roadside vegetation).

The choice of threshold distances as a model for the dispersal of multiple species is consistent with the multi-species analysis of Keitt et al. (1997), in which distances between 10 km and 100 km are used in a graph-theoretic model of connectedness. Threshold distances for modelling dispersal...
ability are also used in single-species analyses (Bunn et al. 2000; Bodin & Norberg 2007; Fall et al. 2007). However, Baguette & Van Dyck (2007) note that different species (and indeed different groups and individuals from the same species) perceive fragmented landscapes differently. Baguette and Van Dyck suggest precaution when attempting to generalize among species in this regard. There is clearly a trade-off between collecting detailed dispersal observations for individual species and attempting to complete a broad analysis for the purpose of ecosystem management across the gamut of plant and animal species. We suggest that the approach may be used in both situations, but follow the trend for stressing precaution in interpreting the results in either case.

Results

The results are provided for three scenarios, each with their own level of functional granularity. The first is a coarse-scale representation where smaller habitat patches are ignored and only the larger patches within the corridors contribute to the construction of pathways. The second and third scenarios are for the Stirling–Peniup corridor and the Peniup–Fitz corridor. Both of these scenarios use a much finer functional granularity, recognizing roadside vegetation and smaller tracts of riparian vegetation as habitat that may potentially contribute to pathways through the corridors.

In the coarse-scale analysis over the full range of the macro-corridor, the range of \( \nu \) is 1000 m to 8000 m. The top 100-graph edges, presented in Figure 4, show two separate complex pathways. The significant geometric phase transition occurs at around 2000 m and many redundant pathways are available beyond 3000 m. This means that edges that belong to graphs with \( \nu > 2000 \) m contribute significantly to the top-ranked pathway through the corridor at this functional granularity.

In the Stirling–Peniup corridor, the range of \( \nu \) is 300 m to 2500 m. The top 100-graph edges, presented in Figure 5, show one mostly east-west pathway that utilizes roadside vegetation in several locations. The geometric phase transition occurs at around 775 m and redundant pathways contribute significantly to the top-ranked pathway through the corridor at this functional granularity.

In the Stirling–Peniup corridor, the range of \( \nu \) is 300 m to 2500 m. The top 100-graph edges, presented in Figure 5, show one mostly east-west pathway that utilizes roadside vegetation in several locations. The geometric phase transition occurs at around 775 m and redundant pathways contribute significantly to the top-ranked pathway through the corridor at this functional granularity.
are available nearly immediately above this value. This means that edges that belong to graphs with \( \nu > 775 \) m contribute significantly to the top-ranked pathway through the corridor for this functional granularity.

In the Peniup–Fitzgerald corridor, the range of \( \nu \) is 100 m to 2200 m. The top 100-graph edges, presented in Figure 6, show one significant east–west pathway (albeit dividing into two parallel pathways at two places) through the corridor. The pathway makes use of roadside and riparian vegetation within the corridor. The significant geometric phase transition occurs between 500 m and 610 m and many redundant pathways are available above 800 m. This means that edges that belong to graphs with \( \nu > 500 \) m contribute significantly to the top-ranked pathway through the corridor.

Discussion and summary

The results show a series of specific pathways through the corridors that are ranked highly. At the coarse functional granularity of the Fitz-Stirling analysis, the highly ranked pathways through the macro-corridor are mostly along the coast near the Fitzgerald River biosphere and further north between Corackerup/Peniup and the Fitzgerald River biosphere. At the finer functional granularities in the two constituent corridors, roadside vegetation becomes significant.

Although the riparian vegetation mainly creates north–south pathways in the region, the roadside vegetation supports dispersal and movement in the east–west directions. In Figures 5 and 6, it is apparent that the pathways through the corridor are utilizing roadside vegetation. The distances at which the geometric phase transitions occur are important to the interpretation of the results. Given the specific functional granularity of the Peniup–Fitz analysis (in Figs 2, 3 and 6), we find the geometric phase transition occurs at a value of \( \nu \) between 500 m and 610 m. The implication is that an animal or plant whose dispersal is limited to lesser distances is unlikely to be able to move through this corridor. These animals or plants would benefit from targeted revegetation along the important corridor identified in the analysis. Animals or plants whose vagility is close to this value are at risk of losing their ability to move through the corridor because they are relying on very few pathways through which they may move. These species would benefit from restoration of the identified pathways and revegetation along redundant pathways identified at greater values of \( \nu \). Animals and plants that have the ability to disperse over the matrix at greater distances have many redundant pathways through which they may move and face less risk as a result of further habitat loss.

The difference between the pathways of highest importance shown in Figures 4 and 6 may be accounted for by the difference in granularity between the structures for the two analyses. These differences may also be apparent when observing different species that have a different perception of patch size as habitat. Where some species recognize smaller patch areas as habitat, other species may ignore the same patch if it is below their functional grain (Baguette & Van Dyck 2007). Specific distances may be applied with more confidence for a single-species analysis, if the functional grain is known.

Although this type of analysis may provide useful results for representative groups of species and individual species, we issue a word of caution when using this type of analysis for the purpose of designing revegetation and restoration projects for individual
species. Long-distance dispersal (Nathan & Muller-Landau 2000; Jordano et al. 2007) is a phenomenon that is sometimes observed in nature, particularly in the dispersal of seeds by frugivores or wind. For example, Jordano et al. provide an interesting picture of seed dispersal for a Spanish fleshy-fruited shrub, when they measure the frequency of seeds dispersed at different distances, including distances greater than 1500 m. They produce a seed dispersal curve for small birds, medium birds and carnivorous animals showing each animal’s contribution to the dispersal of seeds at a distance. The question of linking these curves to the values of $v$ in this type of analysis does not yield a simple answer. Given that the majority of seeds in that study are dispersed at distances less than 200 m yet a small proportion of seeds are still dispersed at distances greater than 1500 m, it is not clear what values of $v$ will yield valuable results. Choosing values of $v$ around 200 m may miss the important long-distance dispersal and choosing values around 1500 m may overestimate the contribution of long-distance dispersal. We suggest that the application of this type of analysis to individual species requires context-specific decisions about the relationship between dispersal curves and the interpretation of the analysis results, especially considering the values at which the geometric phase transition occurs.

The value of the betweenness centrality approach is that it is simple to understand and perform relative to the comparative methods, the conditional minimum transit cost and multiple shortest paths tools (Pinto & Keitt 2008). Furthermore, standard and computationally feasible implementations of the betweenness centrality algorithm are consistently available for different languages and platforms. There is intuitive appeal about the way in which the analysis relates to both multi-species dispersal and multi-species functional granularity. Recommendations for future work include the detailed implementation for individual species in the region, and the comparison of analyses between updated vegetation surveys to elucidate the change in connectivity through these corridors over time.

The analyses allow for the identification of locations where targeted revegetation and restoration will have the greatest effect on strengthening the connectivity for the whole region from the Stirling Ranges to the Fitzgerald River biosphere. Additionally, these locations represent the habitat patches where further loss of vegetation will have the greatest detrimental effect on connectivity. A specific location identified for the macro-corridor is adjacent to Marnigarup East Road, east of the South Coast Highway (see Fig. 6). This location represents the highest concentration of high betweenness values when aggregated across...
all scales of analysis. Without consideration for the social, political, economic or other environmental restrictions or incentives of these and other locations, we suggest that the locations identified in this analysis be considered for their relative importance to the sustained biodiversity of the entire region from the Stirling Ranges to the Fitzgerald River biosphere.

Acknowledgements

The research fellowship of the first author was funded by the Alcoa Foundation’s International Sustainability & Conservation Program. We thank Erik Doerr and an anonymous reviewer, whose comments significantly improved the manuscript. The authors acknowledge the advice of Angela Wardell-Johnson, Simon Judd, Paula Deegan and Kerrie Wilson. The views expressed in this paper do not necessarily reflect the views of those groups involved with the Gondwana Link.

References


