Identification of Critical Habitat Corridor Patches by Cut Node Ranking

Gianni Fenu and Pier Luigi Pau

Department of Mathematics and Computer Science,
University of Cagliari,
Via Ospedale 72,
09124 Cagliari, Italy
{fenu,pierluigipau}@unica.it

First online: November 27, 2017*

Abstract

Graph-theoretic approaches are increasingly popular in the representation and analysis of ecological networks, acting as a complement to traditional techniques, particularly to perform connectivity analysis. Local network indices are useful to identify portions of an ecological network that are worth inspecting, such as bottlenecks. However, the simple identification of cut nodes may turn out to be insufficient. By providing criteria to rank identified cut nodes, land managers are able to assess priorities for further local analysis.

1 Introduction

Current policies for the protection of the environment throughout the world generally involve the creation of ecological networks in order to merge the genetic pool of populations of endangered species, and to better protect biodiversity. In the European Union, the Natura 2000 project involves the creation of an ecological network and the coordination of member states in its maintenance and administration. At a local level, land managers are to take into account the impact of decisions on large-scale conservation goals. To improve their understanding of properties of an ecological network, or the portion under their jurisdiction, graph models are used to complement traditional analysis brought about using Geographic Information System (GIS) tools. In this paper, using a case study with a focus on the area of the Metropolitan City of Cagliari in Sardinia, it is shown how graph connectivity analysis can be used to rank critical patches in proposed ecological corridors, giving land managers a way to assess priorities in the process of testing the modification of land use.

The paper is organized as follows: in Section 2, ecological networks and ecological corridors are introduced, and their identification and representation using different tools is discussed. In Section 3, a set of proposed corridors in Sardinia is presented as a useful case to apply further analysis. In Section 4, the identification of cut nodes in a graph representation of ecological corridors is discussed, and the method of analysis is applied in the previously described context. Lastly, in Section 5, conclusions are drawn and an outline for future work is given.

2 Analysis of Ecological Networks

Preservation of the environment is an increasingly important factor in determining decisions in land use management, which are taken within the administration of municipalities, as well as in the definition of policies at national and international levels, with the coordination of administrative bodies up to a continental scale. The continued existence of endangered habitats and species is held as a priority; for this reason, nature protection areas are established and maintained, and steps are taken to avoid or reduce the fragmentation of habitats. Current policies involve the creation and management of ecological networks [7], in which nature protection areas are designed to contribute to large-scale preservation goals in order to protect biodiversity [15]. Network behavior emerges as it can be inferred that the population of a species has the ability to migrate or disperse between sites, so that the populations of that species located in multiple sites can be considered to share their

*This is a pre-copyedited version of a contribution published in Complex Networks & Their Applications VI: Proceedings of Complex Networks 2017 (The 6th International Conference on Complex Networks and Their Applications), Editors: Prof. Chantal Bonner Cherifi, Prof. Hocine Cherifi, Prof. Márton Karsai, Prof. Mirco Musolesi, part of the Studies in Computational Intelligence book series (SCI, volume 689), published by Springer Nature. The definitive authenticated version is available online via https://doi.org/10.1007/978-3-319-72150-7_50

genetic pool, and be part of the same metapopulation. Migration between sites happens through so-called 'ecological corridors' (or 'habitat corridors'), which can be identified in the territory; corridors can be either contiguous or made up of sets of disconnected patches (stepping stones). In some cases, ecological corridors are to be maintained or even created by human intervention.

The current project for an ecological network in the European Union, denominated "Natura 2000", provides proper definitions and establishes rigorous policies concerning nature protection areas, referred to as 'sites', and designated as Special Protection Areas (SPA) for birds, or as Special Areas of Conservation (SAC) and Sites of Community Interest (SCI) for all the other species. On the contrary, linking elements such as ecological corridors are not as thoroughly regulated.

Land management requires the balancing of several aspects and the pursuit of diverse goals, sometimes in apparent conflict with one another: for example, the state of green infrastructure may have to be improved while taking into account the possible impact on the existing road layout, or the effects deriving from its modification [8]. Land managers are required to deal with a wide range of technical and political aspects, and the representation of ecological networks in a model that allows to perform quantitative analysis is an invaluable help in performing their tasks. Nonetheless, a useful model is not always available, due to shortcomings in data collection, while in other cases, there may be available data for multiple kinds of models, making the choice of a viable model a problem in itself.

Geographic Information System (GIS) software represents the main tool available to land managers, as it allows the representation and processing of spatial data as well as attribute data, while supporting proper geographical referencing. This includes the translation between different map projections, on which datasets of varying origins may be based upon. Geographic features may be represented in the form of raster data, based on the tessellation of regions into a regular grid with a given cell size; or as vector data, with the representations of lines and polygon boundaries as vectors. The latter enables better map precision, completely bypassing the problem of choosing a cell size and those deriving from the inability to represent boundaries of arbitrary shapes; however, raster data is more suitable for most kinds of analysis, as well as for the storage of certain continuous attributes, such as altitude.

Graph models are often employed to complement these data types for the purpose of connectivity analysis. In the context of ecological networks, it is usually meant for this concept to refer to the possibility of migration of a certain target species, or a set thereof. In a graph model, each node may represent up to an entire nature protection areas for large-scale studies [9], or down to a single habitat patch for small-scale studies [12]. The two approaches can be thought to complement one another, as large-scale studies can be performed to understand the global impact of local decisions, or to identify an area which requires local intervention in accordance with large-scale goals and policies [5]. Particularly in the former case, linking elements such as ecological corridors may be represented as edges; this is not necessary if the choice of scale and the granularity of data allows for single patches that make up corridors to be represented as nodes, or if the study is focused on birds: in this case, it is considered viable to simply link nodes within a set geographical distance, corresponding to the dispersal distance of the target species. In general, edges are drawn to connect nodes if migration has been detected (e.g. by tracking the movements of tagged individuals) or is otherwise thought to be possible between a pair of nodes. Raster data is particularly useful in the latter case, as the estimation of a resistance value for each cell can be used in an attempt to predict migration paths [10]. Other techniques involve the comparison of genetic pools of samples of distant populations [3]; this method can also be applied to infer the possibility of plant dispersal.

A meaningful graph representation of landscape enables the application of complex network analysis, with the extrapolation of global network indices, as well as local properties at the node level. This can be very helpful in improving the understanding and the representation of properties of the underlying ecological network [1]. For example, the degree of redundancy of links can be assessed using the clustering coefficient, which expresses the frequency of occurrence of full triangles in network topology, and a method to identify bottlenecks in a network is given by ranking nodes according to their betweenness centrality index, which expresses the proportion of shortest paths going through it. The identification of bottlenecks is a common step in vulnerability assessments, performed in various contexts of application, which include ecological networks [9] as well as several kinds of infrastructure networks, such as power grids [4].

However, particularly if a graph model is built in a way that does not accurately represent topological properties of the landscape, the exclusive use of a graph-based approach is prone to providing misleading results [11]. Building a meaningful graph model is a challenge in itself, and generally involves the identification of a set of target species and the study of functional connectivity [14].

3 Case Study

To provide a case study, the portion of the Natura 2000 ecological network within the Metropolitan City of Cagliari is considered. Initial data is based on the identification of ecological corridors performed by assessing the negative contribution to habitat fragmentation of land patches and their suitability as green infrastructure [2].

Relevant information on Natura 2000 sites is extracted directly from public databases made available within the project; however, it is necessary to complement this dataset with additional sources that provide data pertaining to land areas outside of recognized Natura 2000 sites. A land cover map based on CORINE land use codes, which was made available by the Region of Sardinia in year 2008, was used for this purpose. Land use codes can be used to make distinctions of the kind

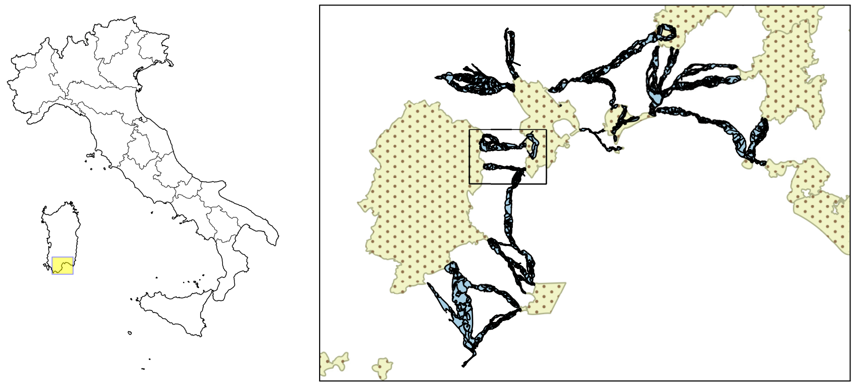

Figure 1: Left: a map of Italy. The highlighted box marks the area represented on the right. Right: map of Natura 2000 sites (dotted areas) and proposed corridors in the Metropolitan City of Cagliari, Sardinia. The superimposed box marks the area represented in Figures 4, 5, and 6.

of vegetation in natural and semi-natural areas (e.g. between broad-leaved, coniferous and mixed forests), as well as to identify agricultural and urban areas. Additional datasets provided by the Region of Sardinia provide values of 'habitat suitability' of each land code for a set of species of interest, which is also defined according to Natura 2000 guidelines. The aggregation of suitability values with other ecological values is the basis for the creation of a resistance map, where land cells are associated with a 'cost-weighted distance' (CWD) value. Land patches within the first decile of CWD values are selected as suitable patches for ecological corridors. Least-cost paths for the migration of species can also be identified on the map; the combination of patch selection and identification of paths between Natura 2000 sites is used to identify potential habitat corridors in the area.

To land managers, the corridors thus identified represent a scope of analysis; their state on the real network should be assessed, in order to determine whether any migration actually takes place on them, and if so, whether there is any threat to their continued existence. In the case where no migration is taking place, reasons should be inferred and it can be meaningful to assess whether it is indeed possible to take action to enable corridors within the identified land patches.

Even though the area being analyzed covers a relatively small portion of the entire Natura 2000 network (for reference, see Figure 1), the number of land features that make up the proposed set of corridors is such that the instantiation of software tools to facilitate their analysis is largely justified. A way to assist land managers in the interpretation of such a large map is to provide a way to rank patches according to their importance with respect to a certain goal. The continued existence of a corridor is clearly a meaningful goal in this context.

4 Identification and Classification of Cut Nodes

To build a representation of proposed ecological corridors, previously identified land cells were intersected with the original land patches from the land cover map, obtaining a vector map of land patches, which are homogeneous with respect to land use codes. The resulting set of land patches can be used as a node set; edges are added to connect pairs of nodes corresponding to adjacent patches, as detected on the map. The graph resulting from this process has a total of 6995 nodes and 17464 edges, and is made up of 18 connected components, each corresponding either to a candidate corridor, or to a connected set of candidate corridors to form bridges between Natura 2000 sites.

In graph theory, a connected component is a maximal subgraph such that a path exists between any pair of its nodes; if the removal of a single node in a connected component causes a loss of this property, that node is referred to as a cut node. If a proper graph model of a corridor can be created, with nodes corresponding to habitat patches, then the identification of cut nodes is a first step in ranking critical patches, which should be given priority for further inspection. Concerning this study, since nodes represent homogeneous areas from the point of view of land use, results of the analysis are to be interpreted accordingly; that is, the identification of a critical patch is to be treated as a sign that the protection of that area's land use type is to be considered important.

While the identification of cut nodes seems straightforward enough as a process, the number of cut nodes in a graph built with the method described earlier turns out to be too large to be of any practical use: as this graph model represents a set of corridors, cut nodes are abundant by construction. Out of 6995 nodes in the graph model being analyzed, 569 are cut nodes (about 8.13%), and since these are distributed over the whole area, the identification of critical habitat patches is not made easier by a simple process of marking cut nodes. A significant factor in determining this property is that the original land cover map presents a large number of enclaves, i.e. patches that are topologically surrounded by a single other patch, which translate to leaf nodes in the graph model; every node connected to a leaf node is a cut node, even those that do not

Table 1: Occurrences of component count resulting from the removal of cut nodes.

Component count	Occurrences
2	372
3	108
4	45
5	18
6	10
7	6
8	2
9	2
10	2
12	1
13	1
18	1
36	1

cause an interruption of an ecological corridor.

It is therefore necessary to rank cut nodes by priority according to some criterion, in order for their list to be useful to land managers. For that purpose, the removal of each cut node is simulated, and some relevant measures are recorded for each node removal: specifically, the number of separate connected components that are obtained, and the size of each component. The size of the second largest component resulting from the split can be considered of special interest, since it will be 1 for those cut nodes, the removal of which only disconnects a leaf node, thus including the cases where an enclave is disconnected; this makes it easier for land managers to filter out cases that may be of lesser interest. It is also useful to visualize results on a map, using gradients to represent a higher or lower component count, and a larger or smaller size of second-largest component. While neither measure acts as a clear indicator of how critical a patch is, a higher value for either attribute (or both) can be used to establish a priority for further examination; the map visualization is useful for comparison with results from other elaborations made on GIS software suites.

The entire process can be summarized as follows:

1. create a graph model by instantiating a node for each homogeneous patch, and linking nodes that correspond to adjacent patches on the map;
2. identify a set of connected components in the graph model;
3. for each connected component, identify its set of cut nodes;
4. simulate the removal of each cut node, noting the number of split components (*component count*) and their size (particularly the second-largest size, or *second-size*) as attributes of the cut node;
5. rank patches according to their *component count* or *second-size* attributes (a value of zero can be assumed if a node is not a cut node) and visualize the result.

Custom scripts for the QGIS software suite [13] were developed to perform steps 1 and 5, while the remaining steps were implemented in custom Java software, using the open source library JGraphT [6] for graph representation and analysis. The connectivity inspector included within JGraphT was used for the identification of connected components (steps 2 and 4), while the biconnectivity inspector was used to identify cut nodes (step 3). QGIS was also used to export maps as images.

Table 1 shows the number of occurrences of each component count attribute, resulting from removal simulations for the 569 cut nodes detected in the Metropolitan City of Cagliari. A majority of cut nodes (372, i.e. over 65%) generate the smallest possible component count of 2, with the number of occurrences quickly dropping for higher component counts; the distribution is visualized in Figure 2. A similar phenomenon is observed for the second-size attribute. In this case, 338 cut nodes (over 59%) are shown to disconnect single leaf nodes; the number of occurrences of larger values becomes progressively smaller, and detected values become sparser: there is at least a cut node with a second-size value from 1 to 19, with 20 being the first value with no occurrences. The largest value with at least two occurrences is 57, and the largest detected value overall is 810. An excerpt comprising values up to 12 is shown in Table 2; the full distribution is shown in Figure 3.

The number of cut nodes for which both attributes take the smallest possible value (2 resulting components, with a second-size of 1) is 265, corresponding to 46.57% of cut nodes. A larger value for either attribute acts as an indicator of

Table 2: Occurrences of size of second-largest component resulting from the removal of cut nodes (excerpt).

Size of component	Occurrences
1	338
2	50
3	17
4	15
5	9
6	10
7	10
8	5
9	5
10	6
11	3
12	3
...	...

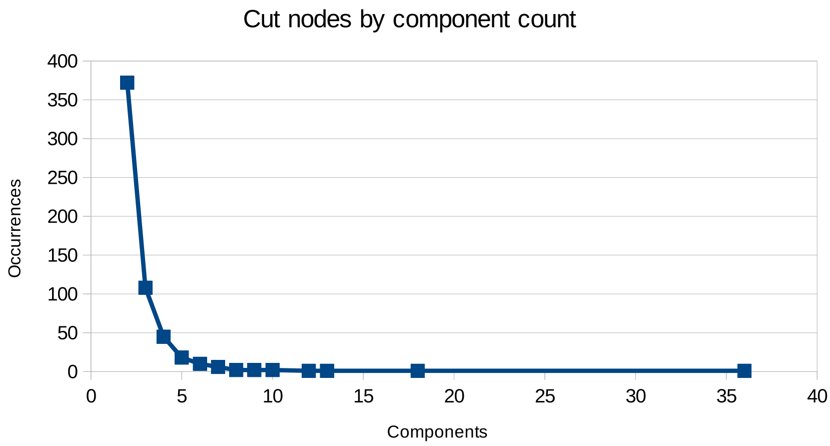

Figure 2: Statistical distribution of component count resulting from the removal of each cut node.

higher potential interest for land managers in a particular cut node; instances exist where the smallest component count is associated to a large value in the second-size attribute, and this may indicate a cut node close to the middle of an ecological corridor, which should not be ignored. A cut node with a small value for both attributes, while not necessarily uninteresting, has lower chances to be as critical as a node associated with larger values. This supports the notion that these attributes are useful in the assessment of priorities.

To provide an example of how results can be visualized, further analysis is conducted on one of the connected components in the graph model; more specifically, the component lying closest to the North-West corner of the area represented in Figures 4 to 6. Table 3 reports the cut nodes detected in this component, each with its ID (which is simply an internal identification number and does not correspond to any meaningful attribute), the number of components that are created with its removal, and the size of these components.

In this case, one of the detected cut nodes (with ID 1949) splits the current connected component into six parts; the same cut node also outranks all others with respect to size of the second-largest split component. This example shows that ranking cut nodes according to these criteria (number of split components and second-largest component size) can be useful to make sense of results obtained for larger areas. This becomes evident by visual comparison: in Figure 4, habitat patches are represented with a dark shade of color if they have been detected as a cut node, or with a light shade otherwise; that is, being a cut node or not is represented as a binary attribute. As it turns out, a significant share of the corridor surface area should be considered critical according to this criterion. By using a gradient to represent the count of components (Figure 5) or the second-largest component size (Figure 6), it is easier to identify land patches with a score expressing higher criticalness (the patch corresponding to node ID 1949, in the eastern part of the component, has the darkest shade of color in both figures).

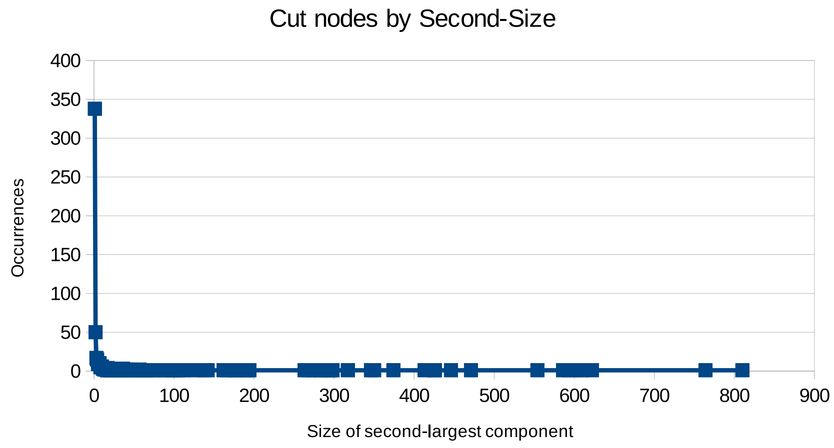

Figure 3: Statistical distribution of size of second-largest component resulting from the removal of each cut node.

Table 3: Excerpt of indicators of detected cut nodes in a connected component.

Cut node ID	Components	Largest Size	Other Sizes
1604	2	296	1
1607	2	296	1
1949	6	270	**21**, 3, 1, 1, 1
2638	2	296	1
2712	2	296	1
4383	3	295	1, 1
4579	2	296	1
5094	2	296	1
5541	2	296	1
6884	4	294	1, 1, 1
6885	2	296	1
6869	2	295	**2**
6889	2	296	1

5 Conclusions and Future Work

Together with nature protection areas, ecological corridors are one of the founding elements of ecological networks, which represent the concept upon which current policies for nature preservation are based. By their definition, contiguous ecological corridors are made up of strips of land, which are easily interrupted by natural disasters or other exogenous factors. Ecological corridors are not necessarily homogeneous from the point of view of land use; creating a graph model where each node represents a homogeneous land patch, the identification of cut nodes gives way to a defining a criterion for determining which patches and land use types are critical for the continued existence of an ecological corridor or a set thereof, possibly complementing other evaluations based on different kinds of analysis.

In this paper, a method to rank the detected cut nodes is described, including a preliminary discussion of its application in a case study based on the Natura 2000 ecological network. Ranking criteria are useful to assess priorities in presence of a large number of cut nodes, which is a common occurrence in this context of operation, due to the definition of ecological corridors. While low-rank cut nodes may still be of interest, priority can be given to cut nodes that are ranked high according to the proposed criteria.

Future work will focus on an extension of this method to take into account additional attributes, such as surface area or distance from a nature protection area, in an attempt to determine a set of attributes that is more apt for filtering out uninteresting patches. Studying the effect of the simultaneous removal of multiple cut nodes is also being considered, as is the comparison of results from the application of this method on graph models created under different preconditions.

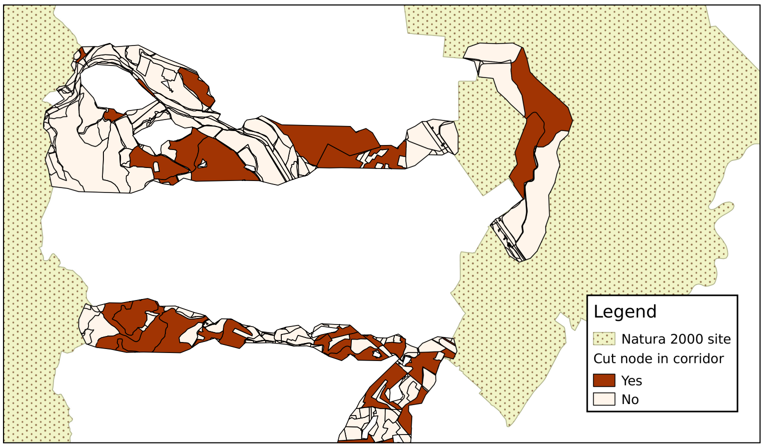

Figure 4: Excerpt of the map of proposed ecological corridors. In the depicted corridor between recognized Natura 2000 sites (dotted), patches identified as cut nodes are marked in a darker shade. However, these include patches that surround other patches completely.

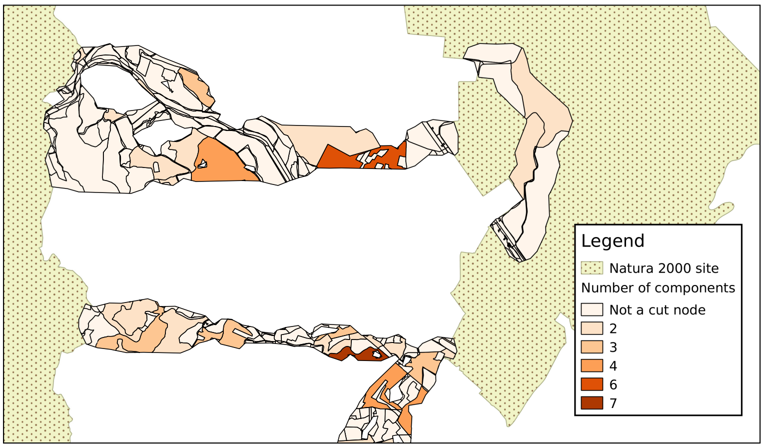

Figure 5: A darker shade represents a higher count of components resulting from the removal of a cut node.

Acknowledgments

This essay is written within the Research Program "Natura 2000: Assessment of management plans and definition of ecological corridors as a complex network", funded by the Autonomous Region of Sardinia (Legge Regionale 7/2007) for the period 2015-2018, under the provisions of the Call for the presentation of "Projects related to fundamental or basic research" in year 2013, implemented at the Department of Mathematics and Computer Science of the University of Cagliari, Italy.

References

[1] Bergsten, A., Zetterberg, A.: To model the landscape as a network: A practitioner's perspective. Landscape and Urban Planning **119**, 35–43 (2013). DOI 10.1016/j.landurbplan.2013.06.009. URL http://www.sciencedirect.com/science/article/pii/S0169204613001254

[2] Cannas, I., Zoppi, C.: Ecosystem Services and the Natura 2000 Network: A Study Concerning a Green Infrastructure Based on Ecological Corridors in the Metropolitan City of Cagliari. In: Computational Science and Its Applications – ICCSA 2017, Lecture Notes in Computer Science, pp. 379–400. Springer, Cham (2017). DOI 10.1007/978-3-319-62407-5_27. URL https://link.springer.com/chapter/10.1007/978-3-319-62407-5_27

[3] Dyer, R.J., Nason, J.D.: Population Graphs: the graph theoretic shape of genetic structure. Molecular Ecology **13**(7), 1713–1727 (2004). DOI 10.1111/j.1365-294X.2004.02177.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2004.02177.x/abstract

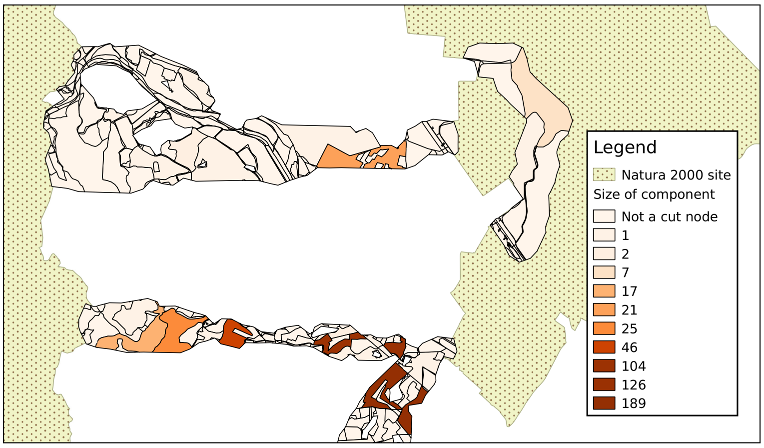

Figure 6: Cut nodes are ranked according to the size of the second-largest component resulting from their removal. A darker shade represents a larger size.

[4] Fenu, G., Pau, P.L.: Evaluating complex network indices for vulnerability analysis of a territorial power grid. Journal of Ambient Intelligence and Humanized Computing pp. 1–10 (2015). DOI 10.1007/s12652-015-0264-0. URL http://link.springer.com/article/10.1007/s12652-015-0264-0

[5] Fenu, G., Pau, P.L., Dessì, D.: Functional models and extending strategies for ecological networks. Applied Network Science **2**(1), 10 (2017). DOI 10.1007/s41109-017-0032-5. URL https://appliednetsci.springeropen.com/articles/10.1007/s41109-017-0032-5

[6] JGraphT Team: Welcome to JGraphT - a free Java Graph Library (2017). URL http://jgrapht.org/

[7] Jongman, R.H.G.: Nature conservation planning in Europe: developing ecological networks. Landscape and Urban Planning **32**(3), 169–183 (1995). DOI 10.1016/0169-2046(95)00197-O. URL http://www.sciencedirect.com/science/article/pii/016920469500197O

[8] Loro, M., Ortega, E., Arce, R.M., Geneletti, D.: Ecological connectivity analysis to reduce the barrier effect of roads. An innovative graph-theory approach to define wildlife corridors with multiple paths and without bottlenecks. Landscape and Urban Planning **139**, 149–162 (2015). DOI 10.1016/j.landurbplan.2015.03.006. URL http://www.sciencedirect.com/science/article/pii/S0169204615000602

[9] Mazaris, A.D., Papanikolaou, A.D., Barbet-Massin, M., Kallimanis, A.S., Jiguet, F., Schmeller, D.S., Pantis, J.D.: Evaluating the Connectivity of a Protected Areas' Network under the Prism of Global Change: The Efficiency of the European Natura 2000 Network for Four Birds of Prey. PLoS ONE **8**(3), e59,640 (2013). DOI 10.1371/journal.pone.0059640. URL http://dx.doi.org/10.1371/journal.pone.0059640

[10] McRae, B.H., Dickson, B.G., Keitt, T.H., Shah, V.B.: Using Circuit Theory to Model Connectivity in Ecology, Evolution, and Conservation. Ecology **89**(10), 2712–2724 (2008). DOI 10.1890/07-1861.1. URL http://onlinelibrary.wiley.com/doi/10.1890/07-1861.1/abstract

[11] Moilanen, A.: On the limitations of graph-theoretic connectivity in spatial ecology and conservation. Journal of Applied Ecology **48**(6), 1543–1547 (2011). DOI 10.1111/j.1365-2664.2011.02062.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2011.02062.x/abstract

[12] Pereira, M., Segurado, P., Neves, N.: Using spatial network structure in landscape management and planning: A case study with pond turtles. Landscape and Urban Planning **100**(1–2), 67–76 (2011). DOI 10.1016/j.landurbplan.2010.11.009. URL http://www.sciencedirect.com/science/article/pii/S0169204610002926

[13] QGIS Development Team: QGIS Geographic Information System. Open Source Geospatial Foundation (2009). URL http://qgis.osgeo.org

[14] Urban, D.L., Minor, E.S., Treml, E.A., Schick, R.S.: Graph models of habitat mosaics. Ecology Letters **12**(3), 260–273 (2009). DOI 10.1111/j.1461-0248.2008.01271.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2008.01271.x/abstract

[15] Vimal, R., Mathevet, R., Thompson, J.D.: The changing landscape of ecological networks. Journal for Nature Conservation **20**(1), 49–55 (2012). DOI 10.1016/j.jnc.2011.08.001. URL http://www.sciencedirect.com/science/article/pii/S1617138111000471