Wildlife Corridor Design: connections to Computer Science

Bistra Dilkina & Carla P. Gomes
Cornell University, Ithaca NY 14850
bistra, gomes@cs.cornell.edu

The reserve design problem is an important problem in the sustainability of biodiversity. The general problem consists of selecting a set of land parcels for conservation to ensure species viability. However, biologists have highlighted the importance of addressing the negative ecological impacts of habitat fragmentation when selecting parcels for conservation. To this effect, ways to increase the spatial coherence among the set of parcels selected for conservation have been investigated (see [10] for a review , [9, 11, 1]). We look at the problem of designing wildlife corridors, which consists of selecting land parcels for conservation in order to connect areas of biological significance (e.g. established reserves). Wildlife corridors are an important conservation method in that they increase the genetic diversity and allow for greater mobility (and hence better response to predation and stochastic events such as fire, as well as long term climate change).

Specifically, in the wildlife corridor design problem, we are given a set of reserves (biologically significant areas), a set of land parcels connecting the reserves, and the cost (e.g. land value) and utility (e.g. habitat suitability) of each parcel. The goal is to select a subset of the parcels that forms a **connected network** including all reserves. More formally, we are given a (planar) connected graph $G = (V, E)$ where the graph nodes correspond to land parcels, a set of reserves $T \subset V$, a cost function $c : V \to \mathcal{R}$, and a utility function $u : V \to \mathcal{R}$. There are three natural optimization variants. The budget problem requires that we maximize total utility of the protected parcels while the total cost is less than a specified budget. The quota problem requires minimizing cost while the total utility of the selected parcels is greater than a specified quota. Finally, the prize-collecting problem calls for minimizing the cost of the set plus the utility of the unprotected parcels (here interpreted as a penalty for losing these parcels).

In this work, we investigate different mathematical formulations of the budget-constrained variant of the Corridor problem. Corridor Design is a computationally challenging combinatorial problem with connections to graph problems studied in computer science. The existence of the budget constraint relates the Corridor problem to the class of combinatorial problems involving packing constraints. Removing the connectivity constraint, we have a 0-1 knapsack problem. On the other hand, the connectivity constraint relates the Corridor Design Problem to other important classes of well-studied problems in Computer Science such as the Traveling Salesman Problem and the Steiner Tree problem. In particular, there are two lines of work in Computer Science that are closely related to the Corridor Problem.

The first one is the work of Moss & Rabani [8] on the so-called *Constrained Node Weighted Steiner Tree Problem*. The classical Steiner Tree Problem involves a graph $G = (V, E)$, a set of terminal vertices $T \subset V$, and costs associated with edges, while the goal is to select a subgraph $G' = (V' \subseteq V, E' \subseteq E)$ that is a tree and contains all terminals ($T \subseteq V'$). Moss et al., however, are concerned with a graph with costs and profits associated with nodes (not edges). The problem concerns choosing a subset of vertices that forms a connected subgraph (or equivalently a tree). The name Constrained Node Weighted Steiner Tree Problem can be quite ambiguous. In the context of all the Steiner tree variants, we would rather refer to this problem as the *Steiner Tree Problem with Node Profits and Node Costs.* It is easy to see that the Corridor problem falls in this class. Moss et al. provide results for the special case where there is either no terminals or only one terminal - a specified root node. For all three optimization variants - the budget, quota and prize-collecting, Moss & Rabani [8] provide an approximation guarantee of $O(\log n)$, where n is the number of nodes in the graph. However, for the budget variant, the result is a bicriteria approximation, i.e. the cost of the selected parcels can exceed the budget by some fraction. For really large Corridor instances, finding optimal solutions would be very difficult. In such cases, one is interested in having approximation algorithms with good guarantees. We are interested in extending the results of Moss et al. to include multiple terminals and to guarantee better approximation ratios by exploiting some of the problem-specific structure. In particular, the graph in the Corridor problem captures a set of parcels on land and hence is a planar graph. Demaine, Hajiaghayi, & Klein [6] have recently shown that one can improve the $O(\log n)$ approximation guarantee to a constant factor guarantee when restricting the class of graphs to planar. However, their results only concern the minimum cost Steiner Tree Problem with costs on nodes (but no profits).

The second related line of work is that of Costa, Cordeau, & Laporte [3, 4] on the so-called *Steiner Tree Problem with Node Revenues and Budgets*. They study the variant of the Steiner tree problem where in addition to costs associated with edges, there are also revenues associated with nodes. The goal is to select a Steiner tree with total edge cost satisfying a budget constraint while maximizing the total node revenue of the selected tree. The case where we have edge costs can easily be reduced to the case of node costs by replacing each edge with a new node with the corresponding cost and edges to the endpoints of the original edge. Hence, the Steiner Tree Problem with Node Revenues and Budgets is easier than the budget-constrained Steiner Tree Problem with Node Profits and Node Costs. Nevertheless, the work of Costa et al. on the budget-constrained Steiner Tree Problem with Node Revenues can provide useful insights into this more general problem. In particular, their results on formulations of the connectivity constraints are of relevance.

The Wildlife Corridor Design Problem was studied recently in [2, 7]. The authors designate one of the reserves (terminals) as a root node and encode the connectivity constraints as a single commodity flow from the root to the selected parcels. This encoding is small and easy to enforce. As opposed to the Steiner tree formulation, it does not impose the constraint that the set of used edges should form a tree.

Alternatively, one can use the equivalence with the Steiner Tree Problem with Node Profits and Node Costs. Encodings of the connectivity requirement successfully applied to the Steiner Tree problem involve exponential number of constraints. In particular for the Steiner Tree Problem with Node Revenues and Budgets, Costa, Cordeau, & Laporte [4] suggest using the directed Dantzig-Fulkerson-Johnson formulation with subtour elimination constraints enforcing the tree structure of the selected subgraph [5]. Such exponential encodings cannot be represented explicitly for real life sized instances. Instead, the problem is solved using Branch-and-Cut, where a relaxed master problem omitting the exponential subtour elimination constraints is solved and connectivity constraints are added while solving using max-flow/min-cut algorithms to quickly discover connectivity violations.

In our work, we compare the effectiveness of three different encodings of the connectivity constraints in the budget-constrained Corridor Problem: 1) the single-commodity flow encoding [2]; 2) a multi-commodity flow encoding; 3) a modified directed Dantzig-Fulkerson-Johnson formulation using node costs. In multi-commodity flow encoding, the connectivity of each selected parcel to the root node is established by a separate commodity flow. It provides a middle ground between the single-commodity encoding and the exponential encoding. Although the multi-commodity flow encoding of the connectivity requirement is much larger that the single commodity encoding (yet still polynomial size), it can result in a stronger LP relaxation of the problem which results in tighter bounds on the objective function.

Acknowledgments

This research was supported by NSF Expeditions in Computing award for Computational Sustainability (Grant 0832782). The first author was partially supported by an NSERC PGS Scholarship.

References

[1] Cerdeira, J. O.; Gaston, K. J.; and Pinto, L. S. 2005. Connectivity in priority area selection for conservation. *Environmental Modeling and Assessment* 10(3):183–192.

[2] Conrad, J.; Gomes, C. P.; van Hoeve, W.-J.; Sabharwal, A.; and Suter, J. 2007. Connections in networks: Hardness of feasibility versus optimality. In *CPAIOR*, volume 4510, 16–28.

[3] Costa, A. M.; Cordeau, J.-F.; and Laporte, G. 2006. Steiner tree problems with profits. *INFOR: Information Systems and Operational Research* 4(2):99–115.

[4] Costa, A. M.; Cordeau, J.-F.; and Laporte, G. 2009. Models and branch-and-cut algorithms for the steiner tree problem with revenues, budget and hop constraints. *Netw.* 53(2):141–159.

[5] Dantzig, G.; Fulkerson, R.; and Johnson, S. 1954. Solution of a Large-Scale Traveling-Salesman Problem. *Operations Research* 2(4):393–410.

[6] Demaine, E. D.; Hajiaghayi, M.; and Klein, P. 2009. Node-weighted steiner tree and group steiner tree in planar graphs. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming.*

[7] Gomes, C. P.; van Hoeve, W.-J.; and Sabharwal, A. 2008. Connections in networks: A hybrid approach. In *CPAIOR*, volume 5015, 303–307.

[8] Moss, A., and Rabani, Y. 2007. Approximation algorithms for constrained node weighted steiner tree problems. *SIAM J. Comput.* 37(2):460–481.

[9] Sessions, J. 1992. Solving for habitat connections as a Steiner network problem. *Forest Science* 38(1):203–207.

[10] Williams, J. C., and Snyder, S. A. 2005. Restoring habitat corridors in fragmented landscapes using optimization and percolation models. *Environmental Modeling and Assessment* 10(3):239–250.

[11] Williams, J. C. 2002. A zero-one programming model for contiguous land acquisition. *Geographical Analysis* 34(4):330–349.