

A Note on Using the Resistance-Distance Matrix to solve Hamiltonian Cycle Problem

V. Ejov, J.A. Filar, M. Haythorpe*, J.F. Roddick, and S. Rossomakhine

the date of receipt and acceptance should be inserted later

Abstract An instance of Hamiltonian cycle problem can be solved by converting it to an instance of Travelling salesman problem, assigning any choice of weights to edges of the underlying graph. In this note we demonstrate that, for difficult instances, choosing the edge weights to be the resistance distance between its two incident vertices is often a good choice. We also demonstrate that arguably stronger performance arises from using the inverse of the resistance distance. Examples are provided demonstrating benefits gained from these choices.

Acknowledgements

This work was supported by Australian Research Council grants LP110100166 and DP150100618.

V. Ejov
Flinders University
1284 South Road, Clovelly Park SA 5042
E-mail: vladimir.ejov@flinders.edu.au

J.A. Filar
Flinders University
1284 South Road, Clovelly Park SA 5042
E-mail: jerzy.filar@flinders.edu.au

M. Haythorpe - *Corresponding Author*
Flinders University
1284 South Road, Clovelly Park SA 5042
E-mail: michael.haythorpe@flinders.edu.au

J.F. Roddick
Flinders University
1284 South Road, Clovelly Park SA 5042
E-mail: john.roddick@flinders.edu.au

S. Rossomakhine
Flinders University
1284 South Road, Clovelly Park SA 5042
E-mail: sergei.rossomakhine@flinders.edu.au

1 Introduction

The famous *travelling salesman problem* (TSP) can be summarised as follows: given a graph G containing a vertex set V and an edge set $E : V \rightarrow V$, and a weighting function f that assigns a weight to each edge in the graph, find a simple cycle that includes every vertex in V such that the sum of weights of the edges used is minimised. TSP is known to be NP-hard, and the decision variant, in which the question is whether any tour with length less than a given value k exists, is NP-complete. In the language of TSP, a simple cycle visiting every vertex is called a *tour*, however it is also known as a *Hamiltonian cycle*. A closely related problem is the *Hamiltonian cycle problem* (HCP) which simply asks if at least one Hamiltonian cycle exists in a given graph. Graphs with at least one Hamiltonian cycle are called *Hamiltonian* and graphs with none are called *non-Hamiltonian*.

Although specialised algorithms and heuristics for solving HCP do exist (e.g. see Baniasadi et al. (2014); Chalaturnyk (2008); Eppstein (2003)), a common and typically quite successful method for solving HCP is to first cast it as a TSP problem, and then use one of the suite of highly-developed TSP heuristics (e.g. see Applegate et al. (2006); Helsgaun (2000)) to solve it. It is natural to convert HCP to TSP, since in the latter a tour of minimal length is desired, and any tour suffices to give an affirmative answer to an instance of HCP. Specifically, HCP is converted to an optimisation (typically, integer programming) problem of the generic form:

$$\min_{\mathbf{x}} \sum_i \sum_j w_{ij} x_{ij}, \quad \text{s.t.} \quad \mathbf{x} = \{x_{ij} \in \{0, 1\} \mid (i, j) \in E\} \quad (1)$$

where the positive solution values corresponds to the edges in a tour, and w_{ij} is the weight or cost associated with edge (i, j) . However, there is the question of how best to assign the weights.

One approach which, perhaps, seems logical is to assign every edge in the graph an equal weight, say, 1. This avoids unnaturally biasing the heuristic used by ensuring that all tours have equal weight, equal to $|V|$. However, this naive approach can cause the algorithm to run inefficiently. Since all edges are weighted equally, it is difficult for any TSP heuristic to make good choices, and often random guesses are required to find a solution. Using randomly chosen weights for the edges can help drive the heuristic in some direction, and if the graph has many tours this may be sufficient, but if there are relatively few tours this often amounts to little better than a prohibitively exhaustive search.

Since randomly generated graphs typically have exponentially many tours (and in fact, can be solved in general in only slightly longer than linear time (Frieze 2015)), graphs with relatively few tours usually contain significant structure. In this note, we propose to use information about the structure present in the underlying HCP instance to provide a better choice of weighting function. Specifically, we use information obtained from the *resistance-distance matrix* that contains all of the resistance distances between pairs vertices in a graph. Using the state-of-the-art TSP solver *Concorde* (Applegate et al. 2006), we attempt to solve several difficult families of HCP instances, and demonstrate that these instances can be much more easily solved using resistance distances. Interestingly, we show that for many families of HCP instances, the best choice is to use the inverse of resistance distances.

2 Resistance-Distance Matrix

Suppose that a graph G corresponds to an electrical network, such that the vertices on the graph are electrical components, and each edge (i, j) corresponds to a 1 ohm resistor being placed between components i and j . The resistance distance Ω_{ij} is then the effective resistance between components i and j , and the matrix of all resistance distances is called the resistance-distance matrix.

Although there are several ways to formulate resistance distance mathematically, perhaps the simplest arises from the inverse of the *Laplacian matrix*. Suppose that a graph G of order n has a corresponding adjacency matrix A , and further define a diagonal matrix D where each entry d_{ii} is equal to the degree of vertex i in G . Then the Laplacian matrix $L := D - A$. Define a new matrix $\Gamma := L + \frac{1}{n}J$, where J is the $n \times n$ matrix with every entry equal to 1. Then the resistance distance between vertices i and j was shown in (Babic et al. 2002) to be equal to

$$\Omega_{ij} = (\Gamma)_{ii}^{-1} + (\Gamma)_{jj}^{-1} - 2(\Gamma)_{ij}^{-1}.$$

Since an inverse needs to be found, it takes $O(n^3)$ time to compute the full set of resistance distances for every pair of vertices in the graph. However, for difficult instances of HCP this is a relatively small expense compared to the time taken to solve the instance. Note that the above definition is only well-defined if G is connected. If G is disconnected then the resistance distance can be found between vertices in connected components by treating those connected components as individual graphs, and the resistance distance between vertices in different connected components is defined to be infinity (see p.132 in Bapat (2010)).

Research into the value of the resistance-distance matrix is still relatively young, arguably initiated in 1993 by Klein and Randić (1993), although the underlying theory of resistive electrical networks has been studied for far longer (e.g. see Doyle and Snell (1984)). The resistance-distance matrix has been used extensively in mathematical chemistry, where it is also known as the commute-time distance. For example, it is used as a tool in studying cyclicity and stability of molecules (e.g. see Babić et al. (2002); Fowler (2002)) as well as signal transduction in proteins (e.g. see Chennubhotla and Bahar (2007)). Arguably the most famous result arising from this line of research is the Kirchhoff index (Lukovits et al. (1999)), defined as one half of the sum of all entries of the resistance-distance matrix. Palacios (2001) computed closed-form formulae for the Kirchhoff index of various classes of graphs. In the years since it was introduced, the resistance-distance matrix has been recognised as a useful tool in analysing graphs and networks of various kinds, including communications networks (Tizghadam et al. 2010), social networks (Alguliev et al. 2011), ecological connectivity models (McRae et al. 2007), and graphs arising from natural language processing (Rao et al. 2008), while Spielman and Srivastava (2011) use a nearly linear-time approximation of the resistance-distance matrix to develop a fast graph sparsification algorithm.

We now extend the use of the resistance distance matrix to aid in solving a classical graph theory problem, namely, HCP. In particular, in the formulation (1), we consider $w_{ij} = \Omega_{ij}$ and $w_{ij} = \Omega_{ij}^{-1}$, as well as randomly generated w_{ij} 's. In the following section, we demonstrate that difficult classes of HCP instances can be much more effectively solved by using resistance distances to weight the edges. This finding is, perhaps, unsurprising since the resistance distance is in some sense a measure of how many paths exist between two vertices. Intuitively then, it would seem to be linked to the problem of trying to find a longest cycle. Of course, such a link need not be direct, and so we will consider both the resistance distance, and the inverse of the resistance distance, and will show that in various cases, one or both are more effective than random weights.

3 Examples

We now provide examples of families of HCP instances which, in our experience, are difficult to solve, primarily due to the relative rarity of Hamiltonian cycles.

- **Aldred-Thomassen graphs** - Aldred and Thomassen (1999) produced a family of 3-connected graphs which has only a single Hamiltonian cycle, and for which all vertices are degree 3 except two vertices of degree 4. One example of such a graph is displayed in Aldred and Holton (1999) and can be generalised easily. The graphs have order $16 + 4k$ for integer $k \geq 1$.
- **Fleischner-2 and Fleischner-3 graphs** - Fleischner (2014) provides two infinite families of graphs which all have minimum degree 4, and only a single Hamiltonian cycle. The first family contains 2-connected graphs of order $169k$ for integer $k \geq 2$. The second family contains 3-connected graphs of order $85 + 323k$ for integer $k \geq 2$. We will henceforth refer to the two families as Fleischner-2 and Fleischner-3, respectively.
- **Modified Flower Snarks** - Isaacs (1975) constructed an infinite family of Snarks, that is, non-Hamiltonian 3-regular graphs with chromatic index 4, called the Flower snarks. The family contains graphs of order $8 + 4k$ for any odd integer $k \geq 1$, which are constructed by taking k copies of the Star graph and joining them together in a prescribed way. The graphs are all *maximally non-Hamiltonian*, that is, the addition of any edge renders them Hamiltonian. We modify the Flower snarks by adding a single edge. Specifically, we add an edge between any two non-adjacent vertices in one of the star graphs (any such choice is equivalent).
- **Minimally-Hamiltonian Regular Graphs** - Haythorpe (2016) proposed a family of regular graphs which are conjectured to contain the minimal number of Hamiltonian cycles over all k -regular graphs (for $k \geq 5$) of equivalent order. The family contains graphs of order $m(k + 1)$ for integer $m \geq 2$ and $k \geq 5$. The graphs contain $(k - 1)^2 [(k - 2)!]^m$ Hamiltonian cycles. In our experiments we chose $k = 10$ in all cases.

We considered several instances of graphs of the above types. In each case, we attempted to solve the graph by submitting it to *Concorde* in sparse format with edge weights chosen from the following four schemes (rounding up when necessary):

1. All edge weights are equal to 1.
2. All edge weights are randomly chosen integers between 1 and 100.
3. All edge weights are equal to the resistance distance between their two incident vertices multiplied by 100.
4. All edge weights are equal to the inverse of the resistance distance between their two incident vertices multiplied by 100.

In Table 1 we demonstrate how quickly *Concorde* was able to solve the various instances for the four schemes. In each case, we limited the size of the instance to less than 2000 vertices to ensure memory management was not an overriding factor in the experiment. We stopped *Concorde* if an individual run took longer than 24 hours. The experiment was conducted on a Linux machine with an AMD Opteron 6282 SE 2.6GHz Processor and 512GB RAM.

It is interesting to note the vast range of solving times, depending on the scheme used. For example, the Fleischner graphs with over 1000 vertices all took longer than 24 hours to solve with three of the schemes, but mere seconds with Scheme 4. In fact, Scheme 4 typically dominated the other methods, but was unable to solve the modified Flower Snarks which in turn were solved in seconds by Scheme 3. These findings highlight the importance of using appropriate weights for instances of this kind. In the case of the modified Flower Snarks, the single added edge must be used in all

| Family | Order | Scheme 1 | Scheme 2 | Scheme 3 | Scheme 4 |
|--------------|-------|----------|----------|-----------------|-----------------|
| Aldred-Thom. | 736 | 00:07:52 | 00:02:21 | 00:02:35 | 00:00:01 |
| Aldred-Thom. | 976 | 00:07:19 | 00:03:20 | 00:03:32 | 00:00:01 |
| Aldred-Thom. | 1216 | Failed | 00:04:11 | 00:10:19 | 00:00:02 |
| Aldred-Thom. | 1536 | Failed | 00:06:37 | 00:06:29 | 00:00:02 |
| Fleischner-2 | 338 | Timeout | 00:06:08 | 00:02:29 | 00:00:01 |
| Fleischner-2 | 507 | Timeout | 00:37:52 | 00:06:28 | 00:00:01 |
| Fleischner-2 | 676 | Timeout | Timeout | 00:21:42 | 00:00:01 |
| Fleischner-2 | 845 | Timeout | Timeout | 01:08:31 | 00:00:01 |
| Fleischner-3 | 731 | Timeout | Timeout | 00:19:08 | 00:00:01 |
| Fleischner-3 | 1054 | Timeout | Timeout | Timeout | 00:00:02 |
| Fleischner-3 | 1377 | Timeout | Timeout | Timeout | 00:00:02 |
| Fleischner-3 | 1700 | Timeout | Timeout | Timeout | 00:00:02 |
| Flower | 324 | Timeout | Timeout | 00:00:01 | Timeout |
| Flower | 684 | Timeout | Timeout | 00:00:02 | Timeout |
| Flower | 996 | Timeout | Timeout | 00:00:05 | Timeout |
| Flower | 1332 | Timeout | Timeout | 00:00:06 | Timeout |
| Minimal Reg. | 1100 | 00:08:52 | 00:00:25 | 00:06:44 | 00:00:14 |
| Minimal Reg. | 1375 | 04:34:59 | 00:00:30 | 00:10:40 | 00:00:21 |
| Minimal Reg. | 1650 | 09:07:27 | 00:00:58 | 00:13:26 | 00:00:15 |
| Minimal Reg. | 1925 | 21:09:23 | 00:02:06 | 00:27:10 | 00:00:26 |

Table 1 Times taken (hh:mm:ss) to solve each instance in Concorde using the three weighting schemes given above. The fastest runtime for each instance is given in bold. Any run that took longer than 24 hours was terminated (labelled “Timeout”). The label “Failed” means that Concorde encountered an unexpected bug and crashed during execution.

Hamiltonian cycles (since the graph was non-Hamiltonian before its addition). This edge appears to always have the lowest resistance-distance out of all edges in the graph, which perhaps explains the strong performance. It appears that “important” edges often have outstanding resistance-distance, although this can be in a minimal or maximal sense.

The *Minimal Reg* instances are a strong example of the kind of graphs which contain many Hamiltonian cycles, but contain some complex structure. Choosing to give each edge equal weight appears to have been a very poor strategy, but since each graph contains many Hamiltonian cycles, choosing random weights worked relatively well. Indeed, for these instances, random weights turned out to be superior to resistance distances, but not to the inverse of resistance distances.

Given the strong performance obtained by using the inverse of resistance-distance, and inspired by the field of electronics where the inverse of electrical resistance is called conductivity, we propose that this useful metric be given the name *conductivity distance*. Intuitively, it seems reasonable that conductivity distance would be a good choice of weights, since a large resistance distance implies many paths go between two vertices, and hence there are potentially more ways of finding a Hamiltonian cycle using the corresponding edge.

Although we have only used resistance distance and conductivity distance to solve instances of HCP in this note, the impact on solving time is so dramatic that it hints at the potential for future research into using these metrics for sparse TSP instances. Specifically, by augmenting the edge weights of a TSP instance with the resistance distance or conductivity distance (with intelligently chosen coefficients to prevent the true edge weights from being swamped), it might be possible to help coax TSP solvers such as Concorde to find good solutions more rapidly. This development could be widely useful, since even non-sparse TSP instances (such as those based on the Euclidean distance between each pair of vertices) are typically first converted to a sparse instance by a sparsification heuristic.

References

1. R.E.L. Aldred and D. Holton. Planar Graphs, Regular Graphs, Bipartite Graphs and Hamiltonicity. *Australasian Journal of Combinatorics*, 20:111–131, 1999.
2. R. Alguliev, R. Aliguliyev and F. Ganjaliyev. Investigation the role of similarity measure and ranking algorithm in mining social network. *Journal of Information Science*, 37:229–234, 2011.
3. D.L. Applegate, R.B. Bixby, V. Chavátal and W.J. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, 2006.
4. D. Babić, D.J. Klein, I. Lukovits, S. Nikolić and N. Trinajstić. Resistance-Distance Matrix: A Computational Algorithm and Its Application. *International Journal of Quantum Chemistry*, 90(1):166–176, 2002.
5. P. Baniasadi, V. Ejov, J.A. Filar, M. Haythorpe and S. Rossomakhine. Deterministic “Snakes and Ladders” Heuristic for the Hamiltonian cycle problem. *Mathematical Programming Computation*, 6(1):55–75, 2014.
6. R.B. Bapat. *Graphs and Matrices*, Springer, 2010.
7. A. Chalaturnyk. A Fast Algorithm For Finding Hamilton Cycles. Masters Thesis, University of Manitoba, 2008.
8. C. Chennubhotla and I. Bahar. Signal Propagation in Proteins and relation to Equilibrium Fluctuations. *PLoS Computational Biology*, 3:1716–1726, 2007.
9. P.G. Doyle and J.L. Snell. *Random walks and electric networks*, Mathematical Association of America, 1984.
10. D. Eppstein. The Traveling Salesman Problem for Cubic Graphs. In Frank Dehne, Jörg-Rüdiger Sack, and Michiel Smid, editors, *Algorithms and Data Struct.*, volume 2748 of *Lecture Notes in Computer Science*, pages 307–318. Springer Berlin (2003).
11. H. Fleischner. Uniquely Hamiltonian Graphs of Minimum Degree 4. *Journal of Graph Theory*, 75(2):167–177, 2014.
12. P.W. Fowler. Resistance Distances in Fullerene Graphs. *Croatica Chemica Acta*, 75(2):401–408, 2002.
13. A. Frieze and S. Haber. An almost linear time algorithm for finding Hamiltonian cycles in sparse random graphs with minimum degree at least three. *Random Structures and Algorithms*, 47(1):73–98, 2015.
14. M. Haythorpe. On the Minimum Number of Hamiltonian Cycles in Regular Graphs. *Experimental Mathematics*, submitted, 2016. Available at: <http://arxiv.org/abs/1608.00713>
15. K. Helsgaun. An Effective Implementation of Lin-Kernighan Traveling Salesman Heuristic. *European Journal of Operations Research* 126:106–130, 2000.
16. R. Isaacs. Infinite families of non-trivial trivalent graphs which are not Tait colorable. *American Mathematical Monthly* 82:221–239, 1975.
17. D.J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993.
18. I. Lukovits, S. Nikolić and N. Trinajstić. Resistance distance in regular graphs. *International Journal of Quantum Chemistry*, 71(3):217–225, 1999.
19. B.H. McRae and P. Beier. Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50):19885–19890, 2007.
20. J.L. Palacios. Closed-form formulas for Kirchhoff index. *International Journal of Quantum Chemistry*, 81(2):135–140, 2001.
21. D. Rao, D. Yarowsky and C. Callison-Burch. Affinity measures based on the graph Laplacian. In: *Proceeding of the 3rd Textgraphs Workshop on Graph-Based Algorithms for Natural Language Processing*, Association for Computation Linguistics, pp. 41–48, 2008.
22. D.A. Spielman and N. Srivastava. Graph Sparsification by Effective Resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.
23. A. Tizghadam and L-G. Alberto. Betweenness centrality and resistance distance in communication networks. *IEEE Network*, 24(6):10–16, 2010.