

Using an evolutionary algorithm to discover low CO₂ tours within a Travelling Salesman Problem

Neil Urquhart, Cathy Scott and Emma Hart

School Of Computing, Edinburgh Napier University, Edinburgh, UK
n.urquart@napier.ac.uk

Abstract. This paper examines the issues surrounding the effects of using vehicle emissions as the fitness criteria when solving routing problems using evolutionary techniques. The case-study examined is that of the Travelling Salesman Problem (TSP) based upon the road network within the City of Edinburgh, Scotland. A low cost path finding algorithm (A*) is used to build paths through the street network between delivery points. The EA is used to discover tours that utilise paths with low emissions characteristics. Two methods of estimating CO₂ emissions are examined; one that utilises a fuel consumption model and applies it to an estimated drive cycle and one that applies a simplistic CO₂ calculation model that focuses on average speeds over street sections. The results of these two metrics are compared with each other and with results obtained using a traditional distance metric.

1 Introduction and motivation

Vehicle routing may be measured in terms of emissions produced as well as the more traditional raw distance metric. Vehicle emissions are dependant on the physical characteristics of the vehicle such as engine type and overall mass as well as factors related to the current driving activities such as acceleration and speed. Previous research in the area of optimisation has examined routing problems such as Travelling Salesman on the basis of a simplistic underlying geographical model, with a fitness function being based on the distance taken by a vehicle within a given solution. This paper will examine the effects of using emissions rather than distance covered as evaluation criterion. A more realistic underlying geographical model is required, rather than one that assumes a constant Euclidean distance between points.

In this paper we examine the Travelling Salesman Problem (TSP) based on real-world street data using CO₂ emissions as our evaluation criterion. The EA employed produces permutations of delivery points, the path taken through the street network between each point determined by the A* algorithm. It is feasible to use a CO₂ costing metric as the cost function of a search technique such as Dijkstra's algorithm, but the time taken to construct paths can be prohibitive. In this work we are using the EA to find paths created on the basis of distance that will contribute to a low CO₂ TSP tour.

2 Previous Work

The TSP is a well known mathematical problem that involves construction a Hamiltonian Path within a graph. It is traditionally presented as finding the shortest route that allows a travelling salesman to visit each location. The TSP has been extensively investigated since the 1930s (for a brief history of TSP investigation see [1]). Because of its simple construction the TSP has been used for testing newly developed heuristics and algorithms.

The use of heuristic approaches to solve the TSP, such as Lin-Kernighan [2] are discussed in [3] [4] [5] [6]. The authors of [4] [5] [6] have carried out extensive investigations into the application of the Lin-Kernighan heuristic to very large instances of the TSP. The work of these authors has included the implementation of the Lin-Kernighan heuristic within a TSP solver entitled CONCORDE (Combinatorial Optimisation and Networked Combinatorial Optimisation Research and Development Environment). A set of benchmark TSP instances is maintained on the internet at the TSPLIB site [7]. Considerable research into solving the TSP using EAs has been undertaken [8] [9] [10] some investigation into the use of EAs to solve the Vehicle Routing Problem (VRP) has also been undertaken [11][12][13][10].

3 Problem Description

3.1 The Geographical Data Source

In recent years sources of accurate geographical data have been made available online by vendors such as Google Maps, ViaMichelin and TeleAtlas. Such encapsulates the topology of streets and to a certain extent the layout and characteristics of road junctions. Within this paper the authors have used data from the Open Street Map (OSM) project (available from www.osm.org under the terms of the Creative Commons Attribution-Share Alike 2.0 licence). Data representing the street graph for the City of Edinburgh, Scotland has been downloaded and stored within a local MySQL database. For each road section the length and road class are stored. The authors have applied empirical knowledge and data presented in [14] to allocate a realistic average speed to each class of road (see table 1). At junctions the appropriate intersecting street details are stored along with attributes such as the existence of traffic signals or a roundabout. The model in use within this paper does not take into account waiting time at junctions or gradients.

The A* algorithm may be used to construct a path between any two locations within the OSM dataset, on the basis of shortest distance. A metric may subsequently be applied to calculate the cost associated with the route. In previous research this metric has often been based upon distance travelled. By examining the attributes of streets and junctions traversed within the route it is feasible to apply a CO₂ emissions metric to the route.

OSM Category	Speed (kph)
default	32
unclassified	32
secondary	36
residential	36
primary	38
tertiary	38
trunk	64
motorway-link	80
motorway	112

Table 1. Average speeds allocated to OSM link classes. For any link class not in the list, the default value (32kph) is used.

3.2 Estimating Vehicle Emissions

The emissions characteristics of a specific vehicle will depend on a range of factors, such as engine size, fuel type and vehicle mass. The actual driving activity and style also influences emissions through variables such as speed, acceleration and gradient. The calculation of estimated vehicle emissions is non-trivial exercise for which a number of approaches have been proposed. In this paper we examine two estimating techniques one based on [15] which attempts to calculate fuel consumption over each second of the journey and relate emissions produced to fuel used. The other (see section 3.4) is based on the work of [15] and assumes and average emissions rate over a road based on the probable speed of vehicles on that category of road.

3.3 Emissions Calculations using a fuel consumption model

A power based instantaneous fuel consumption model for road vehicles was proposed in [15]. The model may be described as follows:

$$dF = \alpha dt + \beta_1 R_T dx + [\beta_2 a R_1]_{a>0} \text{ for } R_T > 0 \quad (1)$$

$$= \alpha dt \quad \text{for } R_T \leq 0 \quad (2)$$

where

dF = fuel (mL) consumed over distance dx (metres) during time dt (s)

α = idle fuel rate (mL/s)

β_1 = fuel consumption per unit of energy

β_2 = fuel consumption during positive acceleration

a = acceleration (m/s), negative when slowing down

R_t = total force required to drive the vehicle (kN) expressed as follows:

$$\begin{aligned}
R_t &= R_D + R_l + R_G, \\
R_D &= b_1 + b_2 v^2, \\
R_l &= Ma/1000, \\
R_G &= 9.81M(G/100)/1000,
\end{aligned}
\tag{3}$$

where

v = speed (dx/dt) m/s

G = gradient (%) +ve or -ve

M = vehicle mass (kg)

b_1, b_2 = drag force function parameters

Appropriate values are provided in [15] to allow the model to be calibrated with respect to a Ford saloon car, these values were derived from observations made using an instrumented vehicle. This model is used by the authors of this paper to estimate fuel consumed within a given solution to the TSP. The litres fuel consumed are converted to Kg of co2 by multiplying by a conversion factor of 2.317 as specified in [16]. To estimate emissions from a route the model has to be utilised in conjunction with a drive cycle. A drive cycle being a series of data points representing the vehicles speed at given intervals (typically one second in many applications). The authors convert the OSM route into a series of interconnected drive cycles. Data for likely acceleration/deceleration curves is taken from [14]. The drive cycle is built as follows:

- for each street section establish a likely average speed based upon the OSM category using the values in table 1.
- create data points for that street within the drive cycle, speeds are not constant but deviate slightly in cycles derived from TRL data [14]
- evaluate the action required for each junction (change in speed, stop/restart etc) and plot appropriate data points (using acceleration/declaration curves from the TRL data [14])

The estimated changes in speed required at a junction are based upon the features present at the junction (traffic signals, roundabout etc) and the classification of the incoming and outgoing roads. It is assumed that at any junction having the attributes of a roundabout or traffic signals the vehicle will stop and restart.

3.4 Emissions Calculations using a simpler model

The model described in 3.3 requires the construction of a drive cycle, which facilitates an attempt to model junctions, but is computationally expensive. As a contrast a simpler model based on the work of [17] is also tested. This is available as spreadsheet based model from the UK National Atmospheric Emissions Inventory [18].

The model proposed by [17] is an average speed model that is applied to each street section, with junctions being explicitly modelled. This negates the requirement to construct a drive cycle. The model may be described as follows:

$$em = (a + b.v + c.v^2 + d.v^e + f.ln(v) + g.v^3 + h/v + i/v^2 + j/v^3)$$

Where

em = the emissions produced as grams of CO2 per km

v = speed (kph)

a,b,c,d,e,f,g,h,i,j = coefficients that define the specific characteristics of the vehicle under consideration.

Values for a-j are provided within [17] for a range of vehicles. For the purposes of linking this model to the EA under consideration the spreadsheet model has been re-implemented as a set of Java classes. The model is applied to each street within the route, the length of street and class of street being available from the underlying OSM database. The average speed for that class of road is derived from observations within [14]. This is a computationally simpler model, but lacks the detail of that in section 3.3.

4 Experimental method and Results

4.1 Problem Instances

Six problem instances were utilised, each requiring a visit to between 10 and 30 delivery points (addresses within the City of Edinburgh data set, see section 3.1) starting and ending at specific start point. Table 2 shows the respective distances between delivery points. The two data sets known as DS1 and DS2 were utilised their delivery points were chosen at random, two data sets referred to as Ring Road and City Centre were also utilised and two data sets big-20 and big-30 are used to explore longer runs. Ring Road uses delivery points located adjacent to the city’s outer ring-road and City Centre used 15 delivery points located within the city centre area.

	DS1	DS2	Ring Road	City Centre	Big-20	Big-30
Delivery points	10	10	10	15	20	30
Average dist (km)	8.29	12.9	18.52	1.45	11.65	9.181

Table 2. A summary of the data sets used

4.2 The evolutionary algorithm employed

The evolutionary algorithm uses a steady state population of 10 individuals. Each individual is a permutation of the points to be visited within the TSP instance. Within each generational cycle a parent is selected using tournament selection of size 2. A child is cloned from the parent and a mutation applied. The child is copied back into the population, by conducting another tournament and having the child replace the looser.

The fitness function builds a route through the street data visiting the points in the order they appear in the chromosome. To determine the path to be taken between each set of points the A* algorithm is utilised as described in section 3.1. The metric used to cost the route will be as set in section 4.3.

4.3 Experimental Method

The EA was executed on each of the four data sets with each of the three routing metrics were used to cost the TSP tours. The metrics used were raw distance and those outlined in sections 3.3 and 3.4. Within the results tables these metrics are referred to as dist, em and sEm respectively. Because of the non-deterministic nature of EAs all experiments were repeated 10 times and the results averaged.

Where results produced using dist and sEm as the evaluation criterion are being evaluated the em metric is applied to the final solution to produce a comparable emissions value.

4.4 Results

The results obtained by the methods outlined in section 4.3 are illustrated in table 5. The distances of tours do not increase significantly when the emissions metrics are utilised. Note that in two cases (City Centre and DS2) a slight decrease in the average distance is noted when optimising for emissions. The biggest increase in distance being the 10% increase noted on the Ring Road dataset. Table 3 shows the average emissions (kg/CO₂) produced by solutions when optimising for distance and for emissions. When comparing the emissions produced by tours evaluated using the dist and em metrics all of the data sets show an improvement. Less of an improvement is noted when sEm is utilised, in the example of DS1, no improvement is noted. The improvements produced when using the larger datasets (Big-20 and Big-30) are less than the improvement with the smaller data sets.

Table 4 compares the emissions values produced by the individual results for each dataset and metric using t-tests. The T-test is used to establish where the difference in results is significant underlying changes. It suggests that the improvements noted in emissions between TSP tours evolved with the dist and em metrics are most significant in the smaller data sets. When optimising for low emissions some increase in distance may be expected as evidenced in table 5. The reduction in emissions is shown in table 3.

Table 6 shows how the road categories used in solutions based on raw distance differs from those based on the emissions model. There do not appear to be any general trends, which suggest that a specific class of road is generally utilised more or less when costing the tours using a specific metric. But it should be noted that the Ring Road data set shows a significant switch to roads of the trunk class when optimising for emissions and that all data sets show a slight reduction in their use of unclassified roads when optimising for emissions



Fig. 1. Two typical tours produced with the DS1 problem instance, showing each delivery point (B-K) and the start/end point (A and L). The upper tour is optimised for minimum distance, the lower for minimum emissions. (Map ©2009 CloudMade - Map data CCBYSA 2009 OpenStreetMap.org contributors. [http : \\cloudmade.com/terms_conditions](http://cloudmade.com/terms_conditions))

5 Conclusions and future work

The results presented suggest that within the context of the TSP there exist new challenges based upon the use of real-world data and emissions metrics. Of the

metric	DS1	DS2	Ring Road	City Centre	big-20	big 30
dist	3.56	4.59	4.35	1.02	5.61	6.46
em	3.29	4.17	3.83	0.97	5.44	6.45
sEm	4.44	4.39	4.04	1.01	5.57	6.76

Table 3. The average emissions (kg/CO₂) produced by solutions when optimising for distance and for emissions. Note that the final solution produced when using sEM or dist is recalculated using em to ensure that it is comparable with that produced using em.

metric	DS1	DS2	Ring Road	City Centre	big-20	big 30
em	0.0034	0.0001	0.0001	0.079	0.391	0.9637
sEm	0.445	0.1096	0.0079	0.7	0.1748	0.2944

Table 4. T-test results based on the raw emissions data presented in table 3 . The t-test compares the emissions values produced by the individual tours with the em and sEm metrics versus the dist metric.

metric	DS1	DS2	Ring Road	City Centre	big-20	big 30
dist	60.2	84.9	70.6	13.1	107.6	114.43
em	60.8	88.33	79.8	13.3	113.3	125.08
sEm	60.7	83.9	77.8	12.8	110.9	117.15

Table 5. The average solution distance (km) when optimising for distance and for emissions

	ds1			ds2			city centre		
	dist	em	sEm	dist	em	sEm	dist	em	sEm
primary	28.39	30.95	29.5	21	21.34	21.18	24.26	21.43	24.38
secondary	9.86	12.49	11.48	17.26	16.6	19.32	0.12	0.12	0.13
residential	29.38	29.85	26.34	27.52	24.49	27.93	16.03	16.89	15.99
unclassified	6.49	3.83	5.63	15.99	14.32	13.52	54.29	55.96	54.40
trunk	2.19	2.31	2.44	3.58	4.6	4.12	0.00	0.00	0.00
tertiary	23.69	20.57	24.62	14.65	18.67	13.92	5.30	5.59	5.10
	ring road			big-20			big-30		
	dist	em	sEm	dist	em	sEm	dist	em	sEm
primary	30.46	23.59	27.07	24.58	27.48	24.41	25.81	23.51	25.08
secondary	15.99	9.72	13.67	13.44	10.62	11.93	11.76	14.31	11.46
residential	13.73	10.72	11.12	28.74	26.51	28.29	26.81	27.26	28.47
unclassified	9.8	4.08	7.24	9.02	8.09	9.97	12.79	12.26	13.14
trunk	16.96	43.13	30.96	3.81	4.41	4.68	2.85	3.29	2.66
tertiary	13.06	8.76	9.95	20.41	22.88	20.72	19.97	19.37	19.20

Table 6. The effect of routing using the emissions metric by road class (expressed as a %).

metrics used sEm (see section 3.4) results are broadly similar to those obtained using raw distance, this might be expected given that both methods rely heavily on road weights. The em metric (see section 3.3) appears to reduce emissions further with relatively little increase in distance, the principle difference between the metrics being that em takes account of junctions. It should also be noticed that within the City-Centre data set, very little improvement in emissions is achieved. This data set has delivery points that are on average 1.45km apart, which provides few opportunities for alternative routes that provide a lower emissions factor. Overall it should be noted that the emissions reductions achieved by the EA have been based on the ordering of the points to be visited. The path taken through the streets is determined by the A* algorithm. The improvement is due to the EA evolving tours that incorporate paths with a low emissions cost. There exists further potential for emissions reduction by applying the emissions metric at the path finding stage.

Both of the emissions metrics discussed in this paper have weaknesses, neither of them take into account all of the possible variables which may effect emissions. Further research to determine those variables that most affect emissions is required. Also required is to enhance the underlying map data to include factors such as congestion patterns and gradients. The inclusion of such additional factors (especially those with a dynamic aspect such as congestion) will further increase the complexity of the fitness function. A major challenge is to integrate such techniques into existing EA and heuristics. It is desirable to apply the techniques suggested in this paper to larger and more complex routing problems such as the Vehicle Routing Problem with Time Windows.

The successful use of A*, rather than a heuristic specially designed to build paths that meet low emission criterion suggests the potential to integrate with existing routing services and products. It is anticipated that demand for such services will increase as legislation forces communities and organisations to account for, and reduce emissions.

References

1. Applegate, D., Cook, W.: Concorde, combinatorial optimization and networked combinatorial optimization research and development environment. <http://www.math.princeton.edu/tsp/concorde.html> (2002)
2. Lin, S., Kernighan, B.W.: An effective heuristic algorithm for the travelling salesman problem. *Operations Research* **21** (1973) 498–516
3. Baraglia, R., Hidalgo, J.I., Perego, R.: A hybrid heuristic for the travelling salesman problem. *IEEE Transactions on Evolutionary Computation* **5** (2001) 613–622
4. Cook, W., Applegate, D., Bixby, R., Chvatal, V.: Finding tours in the TSP. Technical Report 99885, Research Institute for Discrete Mathematics, University of Bonn (1999)
5. Cook, W.J., Cunningham, W.H., Pulleyblank, W.R., Schrijver, A.: The Travelling Salesman Problem. In: *Combinatorial Optimization*. John Wiley and Sons Ltd. (1998) 241–271
6. Applegate, D., Cook, W., Rohe, A.: Chained lin-kernighan for large travelling salesman problems. <http://www.citeseer.nj.nec.com/applegate99chained.html> (2000)

7. : Tsplib, a library of sample instances for the tsp (and related problems) from various sources and of various types (2002) <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>.
8. Freisleben, B., Merz, P.: New genetic local search operators for the travelling salesman problem. In Voigt, H.M., Ebeling, W., Rechenberg, I., Schwefel, H.P., eds.: Proceedings of the Fourth Conference on Parallel Problem Solving from Nature (PPSN IV). Volume 1141., Berlin, Springer (1996) 890–899
9. Tamaki, H., Kita, H., Shimizu, N., Maekawa, K., Nishikawa, Y.: A comparison study of genetic codings for the travelling salesman problem. In: Proceedings of the 1st IEEE Conference on Evolutionary Computing (ICEC '94), IEEE (1994)
10. Homaiifar, A., Guan, S., Liepins, G.E.: A new approach on the travelling salesman problem by genetic algorithms. In Forrest, S., ed.: Proceedings of the 5th International Conference on Genetic Algorithms, Morgan Kaufmann (1993) 460–466
11. Thangiah, S.R.: A hybrid genetic algorithm, simulated annealing and tabu search heuristic for vehicle routing problems with time windows. Practical Handbook of Genetic Algorithms, Complex Coding Systems **III** (1999) 347–381
12. Blanton, J.L., Wainright, R.L.: Multiple vehicle routing with time and capacity constraints using genetic algorithms. In Forrest, S., ed.: Proceedings of the 5th International Conference on Genetic Algorithms, Morgan Kaufmann (1993) 452 – 459
13. Thangiah, S.R., Vinayaamoorthy, R., Gubbi, A.V.: Vehicle routing with time deadlines using genetic and local algorithms. In Forrest, S., ed.: Proceedings of the Fifth International Conference on Genetic Algorithms, Morgan Kaufmann (1993) 506 – 515
14. Green, J.M., Barlow, T.J.: Traffic management and air quality: realistic driving cycles for traffic management schemes. Technical Report TRL Report TRL596 (2004)
15. Akcelik, R., Biggs, D.C.: A discussion on the paper on fuel consumption modeling by post et al. Transportation Research Part B: Methodological **19** (1985) 529–533
16. : Act on co2 calculator: Data, methodology and assumptions paper. Department for Environment, Food and Rural Affairs (2008)
17. Barlow, T.J., Hickman, A.J., Boulter, P.: Exhaust emission factors 2001: database and emission factors. Technical Report PR/SE/230/00, Transport and Road Research Laboratory, Crowthorne, UK (2001)
18. Department for the Environment, F., Affairs, R.: Naei, national atmospheric emissions inventory. (<http://www.naei.org.uk/>)