

# Creating Optimised Employee Travel Plans

Neil Urquhart and Emma Hart

## 1 Introduction and Motivation

Many organisations are currently faced with an increasing requirement to reduce their environmental footprint, either due to statutory requirements or to meet aspirational goals set by their employees or customers. For organisations with a mobile workforce this requirement can possibly be met, in part through the increased use of public transport links by the mobile workforce. Public transport, which can encompass travel modes such as bus, tram and rail and may also include short amounts of walking, has the advantages of low cost and low environmental footprint when compared to car based travel. However it suffers from the disadvantage of potentially being slower for many journeys. Even within developed cities, public transport will have limited coverage with some addresses being too far from the nearest access point or too many individual journey legs being required. In addition, a typical journey by public transport comprises a walking element to the nearest access point, one or more legs by public transport, then an additional walk to the final destination. There may also be short walks involved between individual legs of the journey (e.g. from bus to tram) with the result that some journeys may be considered infeasible by public transport. This paper builds upon the brief introduction made by the authors in [1] and investigates whether an Evolutionary Algorithm can incorporate a mixture of travel modes in order to increase the range of solutions created. Our aim is to explore how public transport links can be introduced within such a mobile workforce scenario and to assess the impact on the objectives of minimising the employee time required to make the visits and the environmental impact of

---

Neil Urquhart  
Edinburgh Napier University, 10 Colinton Road Edinburgh, Eh10 5DT, e-mail:  
n.urquhart@napier.ac.uk

Emma Hart  
Edinburgh Napier University, 10 Colinton Road Edinburgh, Eh10 5DT, e-mail:  
e.hart@napier.ac.uk

the solution. For the employer, there are two major considerations when examining modal choice: environmental impact ( $CO_2$  emissions) and travel times. Within the context of this paper we will examine the effect of making modal choice a decision variable within a mobile workforce problem to provide schedules that meet time window constraints but offer multiple solutions that trade-off time against carbon emissions.

The remainder of this paper is organised as follows, section 2 presents a review of previous relevant work, section 3 describes the problem instances being investigated, the methodology used (an Evolutionary Algorithm) is described in section 4, results are given in section 5 with conclusions in section 6.1

## 2 Previous Work

This paper brings together two strands of work — that of multi-modal Workforce Routing and Scheduling Problems (WRSP) and vehicle routing to optimise emissions. It extends previous work in bringing together the two issues, and treating the problem as multi-objective in terms of meeting constraints, minimising time, and minimising emissions. In addition, it addresses some issues prevalent in the literature with respect to exploiting realistic data by utilising a government sponsored journey planning service in conjunction with an emissions model.

The WRSP has been investigated by a number of researchers, for a full survey of work in this area the reader is referred to [2]. Typically the problem is formulated around mobile health workers, the aim being to find the optimum allocation of workers to jobs each which have a location and a time window associated with them. The resulting schedule has to respect constraints such as time windows, working hours and qualifications. Approaches include Linear programming, constraint programming and meta-heuristics [3], Markov chains [4] and clustering [5].

An investigation into the WRSP incorporating modal choice for transport was carried out in [6]. A two stage approach is used, the first stage uses constraint programming to produce an initial solution. The second stage attempts to improve the solutions by iteratively apply four meta heuristics - neighborhood search, memetic algorithm, scatter search and simulated annealing. The modal choices are between cars and public transportation, it is not clear how the public transport data is gathered. The problem is formulated as a single objective problem, with no focus on carbon emissions.

The issues of taking into account environmental factors within vehicle routing is examined by the authors of [7] who examine the routing of light delivery vehicles within an urban setting. The authors consider a mixed fleet of EFV (Environmentally friendly vehicle) and EUF (Environmentally Unfriendly Vehicles). The novel approach taken utilises a neural network to calculate the likely environmental benefits (in terms of air and noise pollution) for each street section and subsequently the Clarke Wright savings algorithm [8] is utilised to find a set of routes that most

effectively deploys the EFVs within the fleet to gain the maximum environmental benefits.

The problem under consideration in this paper, may be considered to be a multiplex network [9] in that there is an interaction between the related street network representing car journeys and the public transport network. Within such a multiplex network one network may influence the other, which in turn affects processes utilising the network. As the public transport network can include tram and rail links it is not simply a subgraph of the street graph. The author of [10] presents a useful study looking at the relation between the street graphs in greater London area and New York, to the graphs represented by their respective underground rail networks. They investigate the affect of the underground networks on commuting abilities and travel connectivity across the cities, noting that the underground networks mostly influence journeys to and from the city centers and within the suburbs. Within the problem under discussion, the public transport and car networks are both related, a switch in journey mode for an individual within the WRSP is in practice switching them between travel networks.

### 3 Problem Instances

Within this paper two sets of problems are examined, the first one is based upon the City of Edinburgh and the second on the City of London. To schedule any dataset requires extensive data covering the car and public transport networks, including timetabling data to provide journey times and therefore establish feasibility, this case data from transport providers has been utilised. The problems are generic instances of routing problems with time windows, each problem comprises a number of visits which must be made within the working day (defined as 09:00 to 17:00), each visit lasts 30 minutes. All employees start and finish their working day from a city center based office location. Each visit must commence within a specified time window which will always start and end between 09:00 and 17:00. Four different datasets were created in which the number of time-windows and the length of the time-window varied ( see table 1). Note that in set 1, there is effectively no time-window as visits can be placed anywhere in the 8 hour day, whilst in set 4 all visits are allocated a 1 hour time window within which the visit must commence.

As the aim of this paper is to investigate the incorporation of multiple travel modes into travel plans these generic instances allow any member of staff to make any visit, also the number of staff available is not limited, but in many real-life problems the use of agency staff allows extra staff be deployed at short notice.

Set	Visits	Window length	No windows
Edinburgh 1	96	8	1
Edinburgh 2	96	4	2
Edinburgh 3	96	2	4
Edinburgh 4	96	1	8
London 1	61	8	1
London 2	61	4	2
London 3	61	2	4
London 4	61	1	8

**Table 1** Description of the problem instances used

### 3.1 The Edinburgh Datasets

The problem instances investigated are based upon a set of 98 random addresses within the City of Edinburgh, UK and surrounding district, distributed randomly, each problem uses the same addresses, but with differing time windows. Public transport provision within the area covered by the problem is mostly provided by buses, but with a limited tram and rail service. The journey data for public transport is supplied by the Transport Direct bulk journey planning software [11], which was downloaded and stored in a local database the data and problem instances may be downloaded [12]. Values for emissions for journeys by public transport are supplied by Transport Direct as part of their journey planning data. Car journeys were modeled using street network data obtained from Open StreetMap [13], and the GraphHopper [14] library to compute journey durations. Estimated emissions were obtained by applying emissions factors obtained from the National Atmospheric Emissions Inventory (NAEI) [15] for a medium sized petrol engined motor car.

### 3.2 The London Datasets

The London datasets are based upon 61 randomly selected addresses in central London. As with the Edinburgh datasets car journey data is derived from GraphHopper and Open StreetMap with car emissions factors derived using NAEI. Time windows are allocated in the same manner as the Edinburgh datasets.

Transport for London (TfL) is the statutory corporation responsible for transportation within the greater London area, TfL coordinate bus services as well as operating the Underground and Overground rail networks. Data is obtained via an API [16] provided by TfL, emissions values are taken from statistics provided by TfL [17].

## 4 Methodology

### 4.1 The Evolutionary Algorithm

The Evolutionary Algorithm used is the Non-dominated Sorting Genetic Algorithm II (NSGA-II) algorithm [18, 19], with the twin objectives of reducing the predicted emissions generated by the solution and reducing the time required to make the visits. NSGA-II produces a set of non-dominated solutions to the problem under consideration allowing the final choice of solution to be left to the end user, who may examine solutions and choose the one which best fits their current requirements. Two objectives are defined as follows:

- *Minimise total time*  $T = \sum_{i=1}^n t_i$ , where  $n$  is the number of employees,  $t_i$  is the total journey time for employee  $i$ , calculated as the time between leaving and returning to the office, and including travel time, time spent on visits and any waiting time incurred due to arriving at a visit prior to the time window.
- *Minimise total carbon emissions*  $E = \sum_{i=1}^n e_i$ , where  $e_i$  is the emissions associated with the journeys made by employee  $i$  and calculated according to the emissions values stored in the database as described in section 3.

Each problem assumes an unlimited number of employees are available for allocating work too, therefore all solutions are *valid* in that they guarantee that all work items are scheduled.

The genotype representation used within the EA is that of a grand tour [20] which is a permutation of visits, each visit comprises the visit identifier (which is linked to the location and time window data) and the travel mode identifier. The travel mode identifier assigns a travel mode for that visit, a value of 1 for public transport, a value of 0 for car transport. Hence, the EA must evolve must optimise the ordering of customers fed to the decoder, and the mode of travel that should be assigned for each customer.

With the EA a fixed sized population of 300 solutions is used, the size being determined by empirical experimentation. Each new population is the same size as the previous population, the parents being selected by binary tournament from the previous population. Children are created using a permutation crossover operator that copies a complete route, selected at random, from one parent, remaining routes, less any duplicate genes, are copied from the other parent in the order they appear and are appended to the child. Each child has one of two mutation operators, selected at random, applied to it. The first operator selects an entry from the grand tour at random and moves it to a new randomly selected position within the grand tour. The second mutation operator selects a visit at random from the the grand tour and 'flips' the travel mode.

## 4.2 Decoding Solutions

A complete solution consists of a number of individual employee tours, each visit being included in a tour. When constructing a solution, the decoder considers each travel mode in turn, firstly public transport and secondly car based visits in the order that they appear within the grand tour. Visits are initially added to the first employee, as each visit is added the arrival time is calculated based upon the journey times (see section 3) for that mode of transport. If an arrival time is earlier than the commencement of the time window then the employee waits until the start of the time window in order to make the visit. If the arrival time is after the end of the time window then a new employee tour is added to the solution and the visit allocated to the new tour. This decoding process continues until of the visits for the current travel mode have been added. The decoder then passes through the grand tour considering visits with next travel mode. In this way each pass through the grand tour creates a set of employee routes for that travel mode. Note that when considering public transport, not all journeys are feasible, if a visit has been allocated to public transport, but travel from the previous visit is not possible, then the travel mode is changed to car and the decoder continues. This repair mechanism ensures that a feasible solution is always constructed, by using car transport as the default choice. A pseudo code implementation of the decoder is presented in algorithm 1. After applying the decoder, the carbon emissions  $e_i$  and time  $t_i$  associated with each individual tour can be evaluated and therefore the values of the two objectives.

## 4.3 Experiments

For each dataset described in table 1 two initial experiments were performed:

- Using only car based transport - labeled in the results as *car*
- Using a combination of car and public transport - labeled in the results as *pTrans*

The EA was limited to 1,000,000 evaluations on each run as empirical experimentation demonstrated that beyond that value improvements in the quality of solution were seldom noted. Each run of the solver was repeated 10 times in order to obtain a set of results. The Pareto-fronts obtained from the 10 runs were combined and plotted as described in the next section.

## 5 Results

Figure 2 shows the Pareto fronts obtained when combining the output of 10 runs, using only car transport (in black) and using a combination of car and public transport. A summary of all results may be found in table 2. Considering the car-only

---

**Algorithm 1** The decoder used to transform the grand tour genotype into the phenotype.

---

```

1: procedure DECODE(chromosome)
2:   solution = new Phenotype      ▷ The chromosome will be expanded into a set of routes
   within the phenotype
3:   for travel-mode = 0 to MAX-MODE do      ▷ Cycle through each of the allowable travel
   modes
4:     current-route = new Route(travel-mode)      ▷ Create a new route and add it to the
   solution
5:     solution = current-route
6:     previous = OFFICE-BASE
7:     for each gene in chromosome do
8:       if gene.getMode() == travel-mode then
9:         if feasibleTrip(previous,gene,current-mode) then      ▷ True if the travel mode
   will allow the trip
10:        if (current-route.feasibleVisit(gene) then      ▷ True if the visit can be made
   within the time window
11:          current-route.add(gene)
12:          previous = gene
13:        else                                          ▷ Add a new route to the solution
14:          current-route = new Route
15:          current-route.setMode(travel-mode)
16:          previous = OFFICE-BASE
17:        end if
18:      else
19:        gene.setMode(CAR)      ▷ If trip not feasible by current mode, default to car
20:      end if
21:    end if
22:  end for
23:  return solution
24: end procedure

```

---

solutions as we would expect the output forms a curve ranging from high  $CO_2$  solutions taking less time than the lower  $CO_2$  solutions at the other end of the curve. Shown alongside this is the curve generated by combining the output from 10 runs where both public transport and car travel is used. As the *pt* runs control the decision variable we would expect them to be able to create car-only runs which would cover the same solution space as the *car* runs resulting in a pareto fron that overlaps that produced from the *car* runs. Viewing figure 2 shows that this is not the case, none of the graphs show the expected overlap. We note that in no case were the public transport based solutions able to find a solution with the same as or less time than the car-only solution, although in some cases (graphs a, c and d) high  $CO_2$  solutions were found, but none of them exhibited low time values. Comparing between the two cities we note that for the London based problems (graphs e to h) the solver rarely finds high  $CO_2$  solutions unlike the Edinburgh based datasets.

When considering the results to be returned to the user, it is important that the Pareto set, as it's limits encompasses extreme solutions, for instance the all car high

$CI_2$  low time solution and the all public transport low  $CO_2$  high time value solution, clearly the results presented in figure 2 do not encompass those solutions. We may surmise that the fitness landscape is not conducive to NSGA-II exploring car-only solutions. We investigate two possible approaches to resolving this issue the results of which are shown in figure 3. The initial approach taken was to combine the outputs of the *car* and *pt* runs to produce a combined set of non-dominated solutions, these are labeled *comb*. Within figure 3 we can see that combining the outputs produces a Pareto set that encompasses the desired range of solutions. In some cases (noticeably the Edinburgh datasets) we can clearly see a gap in the pareto front which would appear to delineate the *car* and *pt* solution spaces.

Whilst combining outputs produces a useful set of solutions, it does have the disadvantage that it requires the algorithm to be executed twice, thus potentially increasing the runtime by a factor of 2. A second approach towards finding low time, high  $CO_2$  solutions utilises a second mutation operator. The new mutation operator sets all of the travel mode identifiers to 0 (car travel). Whilst this may seem extreme, experiments with only setting part of the chromosome to car did not produce any solutions within the desired area. The rate for this chromosome had to be set to 0.60 (ie applied to 60% of all child solutions) in order to produce solutions in the desired space. Figure 3 show the mutation operator achieving a range of solutions comparable to those achieved by combining the separate results. Figure 3 graphs a-d show that when applied to the Edinburgh dataset the mutation operator results in the production of low time, high  $CO_2$  solutions, but has a detrimental effect on the production of low  $CO_2$  solutions.

Instance	Criterion for best	Car only		Public Transport		Modified	
		Time	Emissions	Time	Emissions	Time	Emissions
Edin 1 Window	Time	3369	119995.34	3474.00	118917.03	3358	119620.88
	Emissions	3432	94471.37	4778.01	73986.22	4725	79761.21
Edin 2 Windows	Time	3576	138192.45	3835	114682.47	3585	130331.66
	Emissions	4091	104052.55	5627	74084.64	5347	76455.2
Edin 4 Windows	Time	3779	123826.16	4154	161021.05	3768	140161.99
	Emissions	4532	108137.63	5936.01	85692.63	6752.01	92530.02
Edin 8 Windows	Time	3932	181014.68	4277.01	195616.28	3998.01	179210.1
	Emissions	6618.01	114441.52	6719.01	82449.23	7062.01	94264.53
London 1 Window	Time	2176	43456	2258	34931.14	2194	38685
	Emissions	2286	36252	3429	9292.98	3523	9104.5
London 2 Windows	Time	2321	45322	2550	35318.07	2381	48022
	Emissions	3537	40226	4185	9621.33	4638	10071.96
London 4 Windows	Time	2411	51454	2813	41631.31	2423	56755
	Emissions	3080	42195	4886	11296.9	4576	11966.33
London 8 Windows	Time	2544	66842	2796	58451.19	2542	72612
	Emissions	3503	50985	4868	14019.83	5266	14148.74

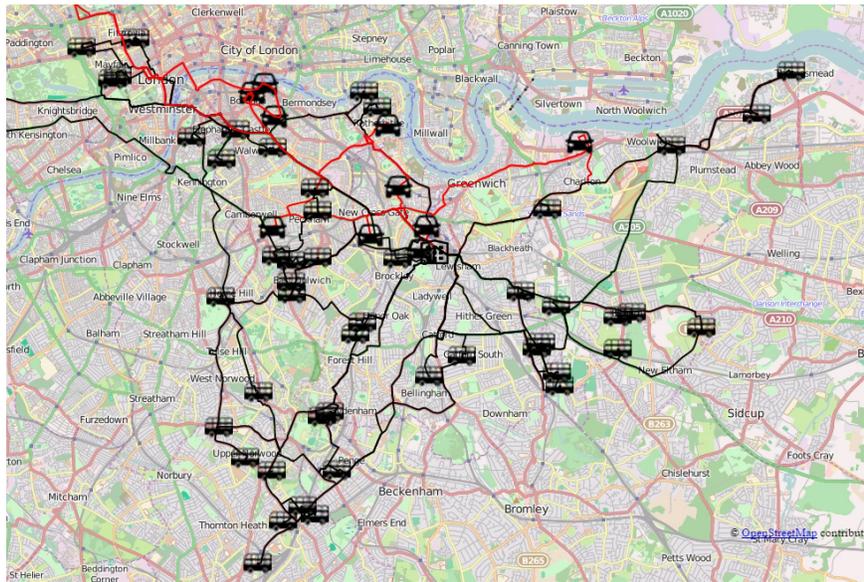
**Table 2** Key: Time - Minutes, Emissions -  $CO_2$  g. A summary of the best solutions found with and without using public transport links. The figures in brackets within the time columns represent the number of employees required within the solution.

## 6 Conclusions and Future work

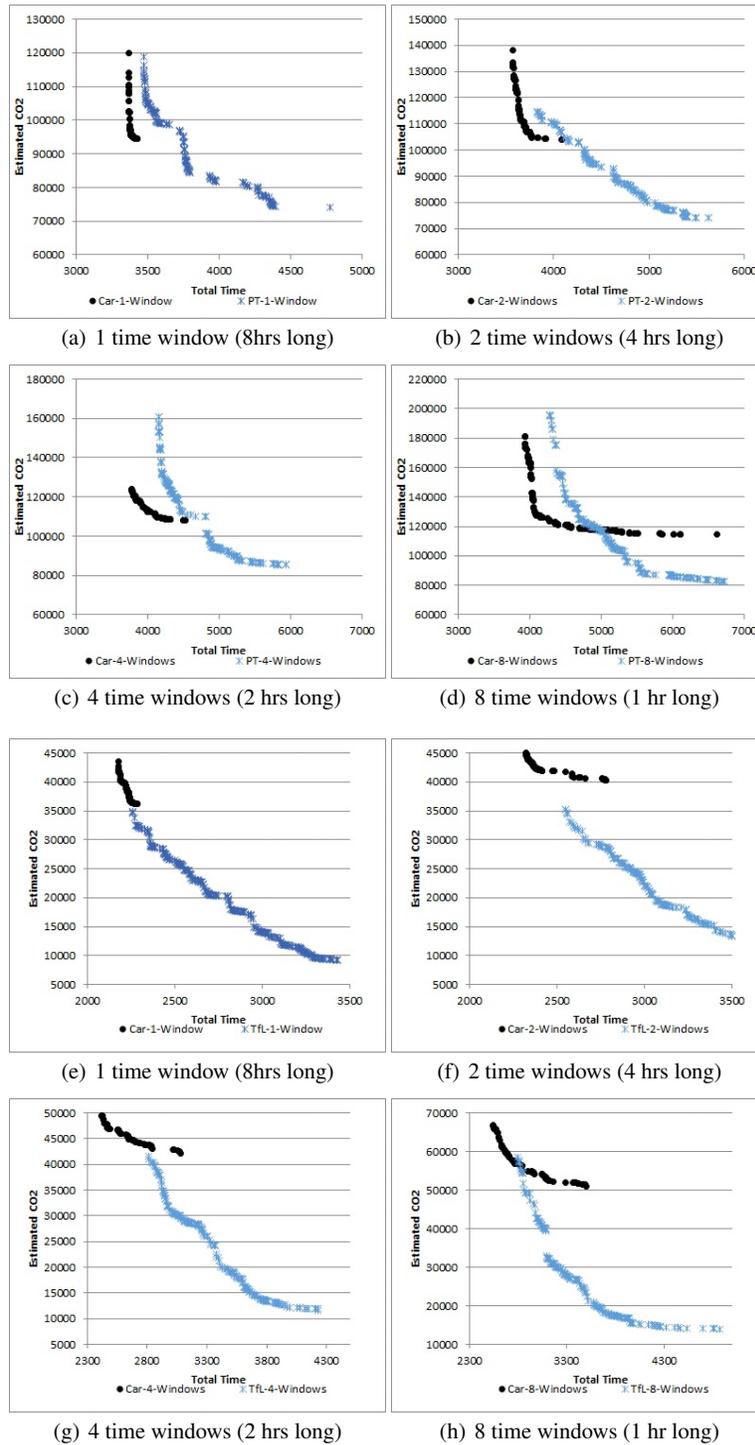
### 6.1 Conclusions

The paper set out to investigate the trade-offs that could be obtained in terms of time and emissions when scheduling a mobile workforce to complete a set of customer visits. Using data from a real public transport and road network, a multi-objective EA was used to evolve solutions under three scenarios: using a car only, using mixed modes of transport and using mixed modes with an added constraint that excluded some potential routes based on public transport due to excessive walking or change-overs. Experiments investigated the trade-offs between the two objectives in four different scenarios in which the time-window for a visit was varied between 1 and 8 hours.

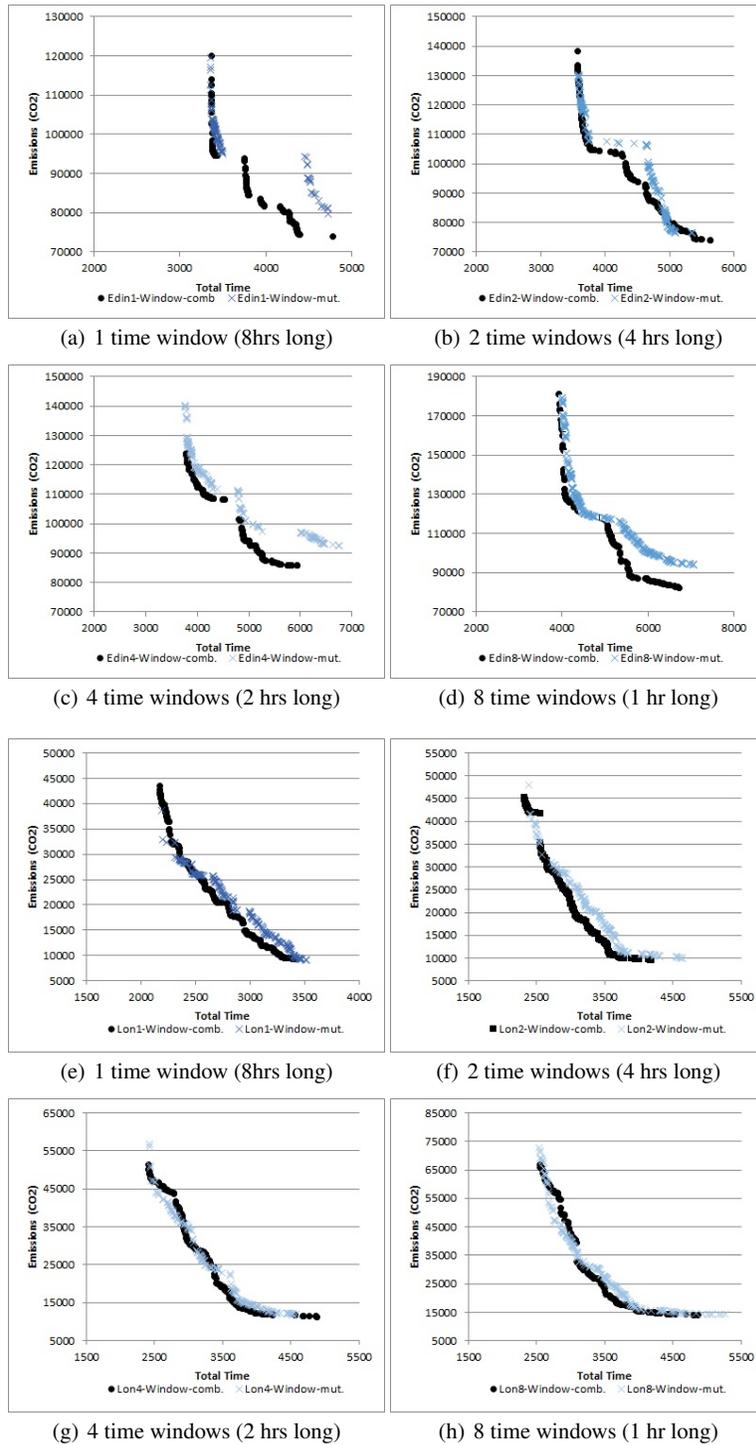
Based on the evidence presented in section 5 the NSGA-II based solver is capable of producing solutions that make use of public transport links in order to reduce  $CO_2$ . The aim of using NSGA-II is to give the user a choice of non-dominated solutions to choose from, in this case to allow the user the option of trading off  $CO_2$  against travel time by allowing public transport to be included within the solution. As initially presented (figure 2 the NSGA-II solver has difficulty in evolving a set of solutions that encompass low time values, despite the car only based solutions



**Fig. 1** An example of a solution constructed from the london data set. The start point is represented by the office building in the middle of the map, visits made by car are noted by the car icon and visits by public transport by the bus icon. Map data ©OpenStreetMap contributors using the Open Database License. <https://www.openstreetmap.org/copyright>



**Fig. 2** Total solutions produced for each problem instance using only car transport and combining car (shown as *car*) and public transport (shown as *PT*). Figures a-d represent the Edinburgh problem instances, figures e-h represent the London instances.



**Fig. 3** Total solutions produced for each problem instance by combining the outputs from figure 2 noted as *comb* and the modified mutation noted as *mut*. Figures a-d represent the Edinburgh problem instances, figures e-h represent the London instances.

proving that low time solutions are feasible. This problem is overcome by the simplistic expedient of combining the Pareto fronts produced using car only solutions and public transport with car solutions, but this has a penalty in the form of increased run times.

Finally, we note that a number of commercial geographical information systems offer public transport data, but currently the costs involved in acquiring data for a problem such as this make them uneconomic. Storing the data in a local database as described not only reduces the requirement for repeated access to on line databases but also decreases the running time of the evolutionary algorithm. However, it is recognised that a future extension of the system would include frequent updating of the database in order to take account of changing travel conditions.

## 6.2 Future work

In the immediate future further investigation of the issues involved in generating a reasonable spread of solutions are to be investigated, both of the solutions proposed in this paper have disadvantages (increased runtime and inconsistent performance).

It is hoped to undertake further studies based on real-world mobile work force problems, and to increase the constraints within the problem to match those typically found in such problems, such as some visits requiring specific workers or workers with a specific attribute or multiple workers. It is also planned to expand the problem formulation to include more transport modes, such as car share, cycling and taxis, such a formulation could include transport costs as one of its optimisation criterion. Future work will also include scaling journey times a specific times in order to reflect rush hour congestion, such scaling would affect road based transport to a far greater degree than rail transport. The scenario investigated here is simplistic, future work will examine problems where there are fixed numbers of staff and constraints on travel modes (e.g. not all staff can drive). We recognise that the problem of finding and optimising transport routes through a city is not restricted to workers, but also occurs in domestic and social situations as well, in the longer term an investigation into optimising travel activities in these areas would be appropriate.

## References

1. Urquhart, N., Hart, E., and Judson, A. Multi-modal employee routing with time windows in an urban environment. In *To appear in Proceedings of the Genetic and Evolutionary Computation Conference 2015*, (2015).
2. Castillo-Salazar, J. A., Landa-Silva, D., and Qu, R. A survey on workforce scheduling and routing problems. In *Proceedings of the 9th international conference on the practice and theory of automated timetabling*, 283–302, (2012).
3. Bertels, S. and Fafle, T. A hybrid setup for a hybrid scenario: combining heuristics for the home health care problem. *Computers & Operations Research* **33**(10) (2006).

4. Out, P. M., Bhulai, S., and Van Meersbergen, M. Optimal patient and personnel scheduling policies for care-at-home service facilities. *European Journal of Operational Research* **219**(3) (2012).
5. Rasmussen, M. S., Justesen, T., Dohn, A., and Larsen, J. The home care crew scheduling problem: Preference-based visit clustering and temporal dependencies. *European Journal of Operational Research* **219**, 598–610 (2012).
6. Hiermann, G., Prandtstetter, M., Rendl, A., Puchinger, J., and Raidl, G. Metaheuristics for solving a multimodal home-healthcare scheduling problem. *Central European Journal of Operations Research* **23**(1) (2013).
7. Dragan Pamuar Goran irovi. Green logistic vehicle routing problem: Routing light delivery vehicles in urban areas using a neuro-fuzzy model. *Expert Systems with Applications* **41**(9), 4245–4258 (2014).
8. Clarke, G. and Wright, J. W. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research* **12**, 568–581 (1964).
9. Gomez-Gardenes, J., Reinares, I., Arenas, A., and Floria, L. M. Evolution of cooperation in multiplex networks. *Sci. Rep.* **2** August (2012).
10. Dobson, S. Multiplex cities: interacting transport networks in metropolitan areas., (2015).
11. Direct, T. <http://www.transportdirect.info>, (2014).
12. Urquhart, N. Public Transport Journey Data. (2015).
13. Foundation, O. <http://www.openstreetmap.org>, (2014).
14. Karich, P. Graphhopper. <https://graphhopper.com/>, (2014).
15. Department for the Environment, F. and Affairs, R. Naei national atmospheric emissions inventory. <http://www.naei.org.uk> (2009).
16. TfL. Transport for london public api. <http://api.tfl.gov.uk/k> (2015).
17. TfL. Travel in london, key trends and developments. report number 1. <http://www.tfl.gov.uk/cdn/static/cms/documents/Travel-in-London-report-1.pdf> (2015).
18. Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A fast elitist multi-objective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation* **6**, 182–197 (2000).
19. Seshadri, A. Nsga-ii a multi-objective optimization algorithm. (2006).
20. Beasley, J. Route first-cluster second methods for vehicle routing. *Omega* **11**(4), 403–408 (1983).