myjournal manuscript No.
(will be inserted by the editor)

Workforce Scheduling and Routing Problems

Literature Survey and Computational Study

J. Arturo Castillo-Salazar[1]**,**
Dario Landa-Silva, Rong Qu

Received: date / Accepted: date

Abstract In the context of workforce scheduling, there are many scenarios in which personnel must carry out tasks at different locations hence requiring some form of transportation. Examples of these type of scenarios include nurses visiting patients at home, technicians carrying out repairs at customers' locations and security guards performing rounds at different premises, etc. We refer to these scenarios as Workforce Scheduling and Routing Problems (WSRP) as they usually involve the scheduling of personnel combined with some form of routing in order to ensure that employees arrive on time at the locations where tasks need to be performed. The first part of this paper presents a survey which attempts to identify the common features of WSRP scenarios and the solution methods applied when tackling these problems. The second part of the paper presents a study on the computational difficulty of solving these type of problems. For this, five data sets are gathered from the literature and some adaptations are made in order to incorporate the key features that our survey identifies as commonly arising in WSRP scenarios. The computational study provides an insight into the structure of the adapted test instances, an insight into the effect that problem features have when solving the instances using mathematical programming, and some benchmark computation times using the Gurobi solver running on a standard personal computer.

Keywords workforce scheduling, employee rostering, routing problems, mobile workforce, mathematical programming, benchmark instances

1 Introduction

In recent times, employees often need to be more flexible regarding the type of jobs they perform and similarly, employers need to make compromises in order to

[1] The author acknowledges CONACYT for its financial support

J. Arturo Castillo-Salazar · Dario Landa-Silva · Rong Qu
E-mail: {psxjaca, dario.landasilva, rong.qu}@nottingham.ac.uk
ASAP Research Group, School of Computer Science, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, United Kingdom, NG8 1BB

retain their best employees (Eaton, 2003; Martínez-Sánchez et al, 2007). Moreover, in some cases workforce should perform tasks at different locations, e.g. nurses visiting patients at their home, and technicians carrying out repairs at different companies, etc. Therefore, the scheduling of workforce with 'flexible' arrangements and 'mobility' is of great importance in many scenarios. Many types of personnel scheduling problems have been tackled in the literature (Baker, 1976; Miller, 1976; Golembiewski and Proehl Jr, 1978; Cheang et al, 2003; Ernst et al, 2004; Alfares, 2004). We are interested in those workforce scheduling problems in which personnel is considered flexible (in terms of tasks and working times) and mobile (travelling is required in order to do the job). By mobility we refer specifically to those cases in which moving from one location to another takes significant time and therefore reducing the travel time could potentially increase productivity. To some extent, this problem combines features from the general employee scheduling problem and also from vehicle routing problems. The survey and computational study presented here represent a step towards our longer term aim of formulating and tackling the problem of scheduling flexible and mobile workforce. In the rest of this paper, we refer to this as the workforce scheduling and routing problem (WSRP).

In Section 2 we describe the WSRP and identify some of the main characteristics of this type of workforce scheduling problems. Section 3 is dedicated to the vehicle routing problem with time windows (VRPTW) because it represents the basic routing component of many of the problems discussed in this survey. Section 4 outlines some workforce scheduling scenarios that have been investigated in the literature and that in our view present a case of WSRP. Examples include home care, scheduling of technicians, manpower allocation, etc. In each subsection of Section 4 we also review the different solution techniques (optimisation, heuristics and hybrid approaches) that have been used to tackle these problems. Section 5 contains the computational study. The experiments are performed using the Gurobi solver, IP and MIP models. A set of problem instances are also presented in this section. Finally, Section 6 summarises our findings and outlines the next steps in our research into workforce scheduling and routing.

2 Workforce Scheduling and Routing Problems

2.1 Description of the Problem

Workforce Scheduling and Routing Problem (WSRP) refers to those scenarios involving the mobilisation of personnel in order to perform work related activities at different locations. In such scenarios, employees use diverse means of transportation, e.g. walking, car, public transport, bicycle, etc. Also, in these scenarios there are more than one activity to be performed in a day, e.g. nurses visiting patients at their homes to administer medication or provide treatment (Cheng and Rich, 1998), care workers aiding members of the community to perform difficult tasks (Eveborn et al, 2006), technicians carrying out repairs and installations (Cordeau et al, 2010), and security guards performing night rounds on several premises (Misir et al, 2011). The number of activities across the different locations is usually larger than the number of employees available, hence employees should travel between locations to perform the work. This results into combination of employee scheduling and vehicle routing problems. The number of activities varies depend-

ing on the duration of the working shift, but assuming that each activity needs to be performed at a different location, a routing problem also arises. A route is a sequence of locations that need to be visited (Raff, 1983) but we exclude problems in which workers need to move across work stations within the same factory for example. Work activities which need to be performed within a given time require scheduling in addition to routing. Tackling WSRP scenarios could potentially involve many objectives like: reducing employees travel time, guaranteeing tasks to be performed by qualified people only, reducing the cost of hiring casual staff, ensuring contracted employees are used efficiently, etc.

We assume employees should rather spend more time doing work than travelling, particularly in settings in which travelling time is counted as paid working time, hence reducing travel time is valuable (Fosgerau and Engelson, 2011; Jara-Díaz, 2000). Like in many workforce scheduling problems, the set of skills that an employee has for performing a task is of great importance in WSRP scenarios (Cordeau et al, 2010) so that employees perform activities at customer premises more efficiently. Many papers in the literature assume that the workforce is homogeneous regarding skills but in many scenarios, a diverse set of skills is the predominant environment. We should note that scenarios like the pick-up and delivery of goods (parcels) is not considered here as a WSRP because no significant 'work' (in terms of time) is carried out at customers' premises. Although, one could argue that the action of delivering a parcel is a task, it does not take a significant amount of time once the worker gets to the destination. This type of pick-up and delivery problems are usually defined as routing problems and are not covered in our study of workforce scheduling and routing problems.

2.2 Main Characteristics

We outline here the main characteristics of WSRP. Some of these characteristics are 'obvious' since they are in the nature of the problem while others were identified during our survey. We include the characteristics that appear the most in the literature.

Time Windows for performing a task (duty, job) at a customer premises. It is assumed that employees can start the work as soon as they arrive at the location. Time windows can be flexible or tight and in accordance to contractual arrangements. In some cases, no time window is defined as employees work based on annualised hours. Also, in some cases employees can benefit from over-time payment, making compliance with the time window a soft constraint.

Transportation Modality refers to employees using different means like car, bicycle, walking or public transport. We assume that time and cost of transportation is not the same for each employee.

Start and End Locations can be from one location, where all employees start at the main office (Eveborn et al, 2006), up to many locations (perhaps as many as the number of employees) assuming each employee may start from their home. In some cases company's policy enforces employees starting their working time at the main office but returning home directly after the last job performed.

Skills and Qualifications act as filters on who can perform a task and there are two main cases. 1) All employees have the same skills and qualifications so anyone

can perform every task. This tends to be expensive for the organisation. 2) Employees with diverse levels of abilities. This is common in industries such as consulting and healthcare. Matching employees' skills to the tasks assigned has been tackled for complex organisations (Cordeau et al, 2010).

Service Time corresponds to the duration of the task and it varies depending on the employee who performs it and the type of task. Most models in the literature assume a fixed duration. If service times are long enough so that they restrict each worker to perform only one job, then the problem reduces to task allocation since every route would consider only one job per employee.

Connected Activities refer to dependencies among two or more activities. Sequential dependency occurs when one activity must be performed before or after another. Temporal dependencies are as defined by Rasmussen et al (2012). Synchronisation is when two or more activities need to start at the same time. Overlap occurs when at any point in time, activities happen simultaneously. In minimum difference dependency, the second activity starts after some given time has passed since the end of the first activity. A maximum difference dependency establishes a limit for the start on the second activity from the end of the first activity. A min+max difference dependency is a combination of the previous two by creating an additional time window for the start of the second activity based on the end of the first activity.

Teaming is necessary sometimes due to the nature of the work to be carried out (Li et al, 2005). If team members remain unchanged then the team can be treated as a single entity and we can assume that all start and end the joint activity at the same time. Nevertheless, if teams change according to the task, then synchronising the arrival of employees to the location of the activity is required. Within teaming, synchronisation refers to employees and not to activities like *Connected Activities*. Also, when teams change frequently then skill matching must be ensured every time teams change to perform a job requiring multiple skills not present in a single employee (Cordeau et al, 2010).

Clusterisation may be necessary for several reasons. One is that employees may prefer not to travel more than a number of miles. Another reason is when companies assign employees to perform work only in certain geographical areas. Clusters may also be created just to reduce the size of the problem by solving a number of clustered sub-problems.

3 Vehicle Routing Problem with Time Windows

The routing part in many of the problems considered here as examples of WSRP is based on the vehicle routing problem with time windows (VRPTW). In this problem the main objective is to minimise the total distance travelled by a set of vehicles serving customers spread across different locations. Every customer must be visited once by one vehicle. Each customer specifies a time window when the visit should take place. The delivery vehicle must arrive at the location within that specified time window. If the vehicle arrives before the time window, it must wait until the time window opens to perform the delivery (Desrochers et al, 1992; Kallehauge et al, 2005). Extensions of the VRPTW include other features such as multiple trips, multiple depots and synchronisation of vehicles.

In the VRPTW extension that covers multiple depots, the fleet of vehicles is distributed across multiple depots allowing vehicles to return to the closest depot once all the deliveries by that vehicle have been completed. This VRPTW variant (Desaulniers et al, 1998) is relevant to our study because its formulation is applicable to workforce scheduling and routing. Many papers in the literature dealing with WSRP scenarios use this VRPTW variant and associate every employee's starting and ending point to a depot. It is also possible for every employee to start at the same location (depot) but to end their working day at a different location (home).

Another extension of the VRPTW allows multiple trips (Brandão and Mercer, 1998). This applies to the case when employees could perform more than one trip on a day to visit the same location. A trip in this context involves a series of tasks before going back to the depot. In WSRP scenarios, an employee is assumed to have a means of transportation. Sometimes the employee might need to go back to the main office (depot) to replenish resources. The type of vehicles that can be used to access a particular customer's location might also be restricted as pointed out by Brandão and Mercer (1997). Vehicles have different capacities which can be associated to model an heterogeneous workforce. Vehicles can also be hired for some time which is associated to hiring casual staff.

Finally, another extention of VRPTW which is important to WSRP is the synchronisation of vehicles. Two or more workers executing a task can be modelled in the same way as when two or more vehicles need to arrive simultaneously at the same customer location (Bredström and Rönnqvist, 2007). Precedence constraints in WSRP are also related to synchronisation of vehicles (Bredström and Rönnqvist, 2008). Assuming a client should be visited more than once per day, the order of visits might matter. For example, before technicians install a satellite TV, the antenna might need calibration and then a demodulator is set. These activities could be performed by different people at different times but the order matters and must be respected.

There are many solution methods proposed to tackle the VRPTW. When using exact approaches, researchers tend to model the problem as multi-commodity network flow problems (Desaulniers et al, 1998; Salani and Vaca, 2011) or following a set partitioning/covering formulation (Bredström and Rönnqvist, 2007). Such models have been tackled using constraint programming, branch and bound, and branch and price (column generation) (Barnhart et al, 1998; Desrosiers and Lübbecke, 2005). Other researchers use hybrid methods that employ heuristics for the generation of columns within a column generation setting (Bredström and Rönnqvist, 2008) or use heuristics to improve an initial solution found with mathematical programming (Fischetti et al, 2004). Alternative approaches include dividing the problem (clustering) into smaller sub-problems and then obtaining a global solution. This approach does not guarantee finding the overall global optimal solution but it is sufficient if the objective is to quickly find feasible solutions (Landa-Silva et al, 2011).

There are a few benchmark data sets for the VRPTW, here we refer to the ones by Solomon (1987) and Castro-Gutierrez et al (2011) which we adapt for our computational experiments.

4 Workforce Scheduling and Routing Problems in the Literature

In this section we review some of the problems tackled in the literature that can be considered as a type of workforce scheduling and routing problem (WSRP). The intention is to illustrate the variety and importance of WSRP scenarios in the real-world. Each subsection focuses on a problem domain and the solution methods that have been used in the literature to tackle it. We distinguish between exact, heuristics and hybrid methods in our review.

4.1 Home Health Care

Bertels and Fahle (2006) describe home health care (HHC) as visiting and nursing patients at their home. Patients preferences regarding the time of visit are respected as much as possible, as they should not be kept waiting. Additionally, nurses have time limitations on the number of working hours per day or the starting and ending time. In HHC, transportation modality is present when nurses travel (car, public transport or walking) to visit more than one patient. The start and end location can vary. Nurses can depart from their homes or from a central health care office, and end their day once they return home or in some cases at the last visited patient's location. A diverse set of skills and qualifications exists usually among nurses. Health care organisations often cannot afford to have nurses trained in all procedures. Then, the use of highly qualified nurses should be restricted to tasks that demand those skills. Nursing tasks can vary in duration (service time), e.g. from a 10 minutes injection to a 45 minutes physical therapy. Connected nursing activities can arise when administering medication, e.g. the first dose is applied during the morning followed by another dose 3 hours later. Some activities require more than one nurse at the same time, e.g. handling a person with epilepsy. In such cases, nurses can be synchronised to arrive at the location at the same time. Clusterisation is used by the organisation providing health care, to avoid nurses having to travel too long distances.

Other characteristics of HHC which are not considered in this WSRP study include nurses preferences, shift types and other legal requirements. Also, it is desirable not to change much the nurses who visit particular patients. This is because patients and nurses develop a bond that is usually good to maintain. Cheng and Rich (1998) explore the use of casual nurses. Their work does not consider different nurses' skills and qualifications but instead, proposes a matching method in which a pairing patient-nurse is either feasible or not for some reason. The objective in their work is to reduce the amount of overtime and part-time work employed.

Home health care has been tackled mainly with hybrid approaches. For example combining mixed integer programming with heuristics for either the routing or the scheduling component (Begur et al, 1997). Another example of combining two approaches is when using constraint programming to obtain a good feasible solution and in a second stage applying a series of meta-heuristics including simulated annealing and tabu search (among others) to improve the quality of the solution (Bertels and Fahle, 2006).

Exact methods have also been used, particularly branch and price, using a set partitioning formulation for the master problem. The model includes real variables

for the scheduling of the activities, and binary variables for deciding whether an activity is performed by a specific employee or not. The pricing problem is an elementary shortest path (Barnhart et al, 1998; Bredström and Rönnqvist, 2007). An extension of such models includes the addition of side constraints in the master problem (Dohn et al, 2008). Not all models for the set partitioning part have both real and binary variables, for example pure integer models are also used by Kergosien et al (2009). The addition of cuts on the time windows improves the branch and price approach and turns it into a branch cut and price which leads to good results.

4.2 Home Care

The home care problem also called domiciliary care, refers to the provision of community care service by local authorities to their constituents (Akjiratikarl et al, 2007). The aim is to schedule care workers across a region in order to provide care tasks within a time window while reducing travel time. This problem is related to the home health care problem described earlier (Bertels and Fahle, 2006; Cheng and Rich, 1998). The difference is that home health care involves helping people for a relatively short period of time to recover after hospitalisation. Home care however usually refers to helping elderly and/or disabled people to perform their daily activities such as shopping, bathing, cleaning, and cooking, etc. (Eveborn et al, 2009). Once a person starts receiving home care support it is likely that he remains receiving such care for a long time.

Home carers usually start travelling from home to deliver support at their predefined destinations using their own transport arrangements (mixed transportation modality) and return home at the end of the day. In some cases reported in the literature, care workers do not start from their home but from a home care office as last minute changes to their schedules are possible and need to be agreed before starting the working day (Eveborn et al, 2009). In some cases, travel time is considered as work hours and hence the objective is to reduce the time used not providing care. Some assumptions are made such as given travel speed for a carer and travel distances to be euclidean. In other cases like the work by Dohn et al (2008), the objective is to maximise the quality level of care service provided. Reducing cost, although important, is not usually the main objective. Dohn et al (2008) study the problem as a variant of the VRP with time windows. Although not as much as in HHC, there are some skills and qualification required in home care when caring for others, e.g. health and safety, handling people with dislexia, etc. Service time is standarised and it only varies due to the experience of the carer or difficulties with the person receiving care. Connected activities also exist in home care, e.g. taking a shower before grocery shopping. Teaming is usually not present as carers tend to be syncronised to perform difficult tasks, e.g. assisting a heavy person. Clusterisation is based on municipalities borders to clearly define which authority (e.g. council, district, etc.) is responsible for each area.

Additional features of home care include prioritising visits. Usually, there is not enough personnel to perform all the visits in a single day. Therefore, visits are rescheduled or even canceled in the worst case. Deciding which visit is not carried out is part of the problem. For example, it is more important to assist someone with his diabetes medication than to help another person in grocery shopping.

The shift patterns are either given by contracts or expressed as preferences by carers. Many organisations emphasise respecting carers' preferences to increase staff retention. Also, tolerance on time windows to perform care tasks can vary widely, e.g. 5 minutes tolerance for critical medical activities, 15 minutes to 2 hours tolerance for support activities, etc.

Home care problems have been solved using all three exact, heuristics and hybrid approaches. Among the exact methods we find linear programming (De Angelis, 1998). Mixed integer linear programming is also used on assignment and scheduling models of home care problems. The assignment model is used when new visits are introduced and the scheduling model is used to generate weekly visits (Borsani et al, 2006). Heuristic methods include local search based on simple heuristics, meta-heuristics like tabu search (Blais et al, 2003), evolutionary approaches such as particle swarm optimisation (Akjiratikarl et al, 2006, 2007) and agent-based modelling (Itabashi et al, 2006). Other methods include hyper-heuristics Misir et al (2010). Among all heuristic methods the solving strategy seems to be similar, to generate a good initial solution followed by local improvement procedures. Common neighbourhood moves include insertion, removal and swaps to interchange both activities among workers and activities in an employee's route. Finally, the combination of a set partitioning model and a repeated matching algorithm, to find suitable pairs of employees and routes in a hybrid approach has also been used to tackle Home Care (Eveborn et al, 2006, 2009).

4.3 Scheduling Technicians

Some telecommunication companies require scheduling employees to perform a series of installation and maintenance jobs, e.g. Cordeau et al (2010). In the literature, this problem is referred to as technician and task scheduling problem (TTSP). In this sector, commitments on time to perform the jobs are enforced, resulting into strict time windows. Technicians need to carry equipment, then it is common to use company vehicles to travel from one customer location to the next one. Technicians start and end at the company premises, although in some cases they are allowed to take home the company car depending on the location of the first job the following day. Technicians, are often highly skilled and this can be related to their experience and training. As a result companies define levels of seniority (e.g. junior technician, supervisor, etc.) among their workforce. Those seniority levels to some extent help to estimate the service time required to complete the job. Tasks tend to be independent from each other within the same day, but in a wider time frame there are some connections between them. In this scenario, teams are often formed with the aim of having a balanced set of personnel with as many skills as possible. Teaming also helps technicians to learn from each other, hence improving their performance. Companies with many branches across different regions use clusterisation to assign jobs to each branch when the scheduling is done centrally for all branches.

The scheduling of technicians has been solved using heuristics approaches, particularly a fast constructive heuristic to reach feasible solutions. Then, local heuristics based on destroy and repair moves are used to improve the solutions (Ropke and Pisinger, 2006). Different heuristics are used depending on the stage of the problem that is is being tackled: activities allocation to employees, skill

matching, and routing (Cordeau et al, 2010). Moreover, evolutionary approaches like particle swarm optimisation have been reported to find good enough solutions for instances of 300 employees (Günther and Nissen, 2012).

4.4 Security Personnel Routing and Rostering

In this problem, round of visits are performed by security personnel to several customer premises in different locations over a 24 hours period (Misir et al, 2011). Many organisations outsource security guards duties only when premises are closed while in other cases, security is outsourced at all times. Round visits must be performed at the contracted time often given as a time window. Security personnel often use vehicles to go from one location to the next one and then walk once they get to the facility but require to check several buildings. Security guards often have their own home as the start and end of the working shift. In Misir et al (2011) the authors mention 16 types of skills that the company records for its workforce and some visits require enforcing those skills. The duration (service time) of each visit can vary but it should be within a time window in which the visit must finish. Visits are independent from each other. Customers are divided into regions (clusterisation), so that security guards living nearby are assigned to each region to reduce travelling time. In this industry, contract terms vary considerably and this leads to many different additional constraints in the problem. Although not mentioned in the scenario, it is not unreasonable to assume that teams of two or more guards are used.

A mathematical programming approach was used by Chuin Lau and Gunawan (2012) when solving a similar problem that involved security teams to patrol different underground stations within the network. Hyper-heuristics is another method that has been applied to this problem, by using two different heuristics selection methods, simple random and adaptive dynamic, followed by an improvement heuristic (Misir et al, 2011).

4.5 Manpower Allocation

The manpower allocation problem (Lim et al, 2004) refers to assigning servicemen to a set of customer locations to perform a diverse range of activities. The objectives are to minimise the number of servicemen used, minimise the total travel distance, minimise the waiting time at service points, and maximise the number of tasks assigned. The manpower allocation problem therefore can be seen as another example of WSRP. Manpower allocation with time windows is particularly relevant since customers explicitly define when the workforce is required. There is no mention of transportation modality so we assume all servicemen use the same type of transport. Every serviceman starts and finishes his working day at the control centre. Skills among the workforce are assumed to be the same, making no difference on who performs the service.

There are restrictions on the number of hours each employee can work. Waiting time, the time that servicemen have to wait at a customer location before the start of the time window, is included within the service time making it vary accordingly. Li et al (2005) add job teaming constraints, where a team is assembled at every

location and work cannot start unless all members of the team have arrived. More recently, a variation of the manpower allocation problem was tackled in the context of scheduling teams to do ground handling tasks in major airports (Dohn et al, 2009). In the work by Li et al (2005) teams are set at the beginning and do not change over the working day. Additional characteristics include teams having mandatory breaks within certain time windows, hence breaks are treated as just another visit.

In manpower power allocation papers, we identified the use of the three types of solution methods. An exact method uses integer programming, based on a set covering formulation which is solved with branch and price (Dohn et al, 2009). Meta-heuristics including tabu search, simulated annealing and squeaky wheel optimisation have also been applied (Lim et al, 2004). Finally, Li et al (2005) relaxed an integer programming formulation of a network flow model to obtain lower bounds. The upper bounds were obtained using constructive heuristics and simulated annealing was the main component of their solution framework.

5 Computational Study

5.1 Data Sets For WSRP

As the above survey reveals, workforce scheduling and routing problems (WSRP) arise in a variety of scenarios. However, there seems to be no established set of benchmark problem instances and results for this type of problems. Hence, from the literature we identified five data sets that can be used as instances of WSRP, either in their original form or after some adaptations.

We consider here the following data sets from the literature: (1) VRPTW instances by Solomon (1987), (2) multi-objective VRPTW instances by Castro-Gutierrez et al (2011), (3) home health care instances by Rasmussen et al (2012), (4) security personnel scheduling instances by Misir et al (2011) and (5) field technicians scheduling instances by Günther and Nissen (2012). Not all these data sets include the features that we have identified in the survey of this paper as characteristics of WSRP. The following paragraphs outline the way in which each of these data sets can be seen as instances of WSRP. The specific adaptations that we made to each data set are detailed later when discussing results of our computational study.

1. Solomon (1987) VRPTW data set has been used broadly in the literature. These VRPTW instances can be transformed into WSRP instances by adding additional constraints. Such constraints are: fixing the number of employees, introducing skills to employees, adding activities which require more than one employee for its completion and establishing connected activities constraint where required. Vehicles visiting customers are considered employees visiting client locations. Since only one vehicle (with no additional features) is involved in a customer visit in these VRPTW instances, this can be seen as single skill activity.

2. Castro-Gutierrez et al (2011) VRPTW data set is a more recent addition to the literature. This data set includes realistic distances and times for travelling between customer locations. This is the main difference with Solomon's data set but their interpretation as WSRP is the same as described above.

3. Rasmussen et al (2012) home health care (section 4.1) data set is actually a WSRP in which carers with different skills provide assistance to elderly, disabled and ill members of the community by visiting them at their homes. Visits are often restricted to a time window. Sometimes visits involve activities requiring more than one carer. This data set has been used in a few publications (Dohn et al, 2008; Misir et al, 2010; Rasmussen et al, 2012).

4. Misir et al (2011) security personnel scheduling data set is also a WSRP in which security guards with different skills visit locations. Visits have associated starting time, due time and duration which then create a time window feature.

5. Günther and Nissen (2012) field technicians scheduling data set is another WSRP in which technicians travel to different locations in a day in order to perform certain jobs. This is often referred in the literature as the technician and task scheduling problem (TTSP) (Cordeau et al, 2010) or as field force optimisation.

Overall, a total of 375 test instances are obtained across the 5 data set gathered and adapted from the literature.

5.2 Initial Experiments With an Integer Programming Model

Here we present our initial experiments on a subset of instances using an integer programming (IP) model. The purpose was to identify benchmark computing times and capabilities of the IP approach. This IP model is based on the work of Bredström and Rönnqvist (2008) as they identified two potential real applications, home care and forest operations, for their combined vehicle routing and scheduling model. Since their model is for VRPTW, we decided to use the instances that we adapted from VRPTW as well. Such instances are the ones by Solomon (1987) and Castro-Gutierrez et al (2011), which represent a total of 183 instances for these initial experiments.

Typically in VRPTW instances, all customers must be visited subject to an upper limit on the number of vehicles used. Also, all vehicles start and end their trip at a common location or depot. In the WSRP context, the above means that all activities must be performed in order to have a feasible solution, and that employees start and end their routes of visits at the same location. One aim of these experiments is to have an insight into the computational effort required to find solutions (first feasible and optimal ones) when performing all activities.

The IP model is as follows:

$$min \qquad \alpha_p \sum_{k \in K} \sum_{(i,j) \in A} c_{ik} x_{ijk} + \alpha_T \sum_{k \in K} \sum_{(i,j) \in A} T_{ij} x_{ijk} \qquad (1)$$

$$s.t. \qquad \sum_{k\in K}\sum_{j:(i,j)\in A} x_{ijk} = 1 \quad \forall\, i \in N, \tag{2}$$

$$\sum_{j:(o,j)\in A} x_{ojk} = \sum_{j:(j,d)\in A} x_{jdk} = 1 \quad \forall\, k \in K, \tag{3}$$

$$\sum_{j:(i,j)\in A} x_{ijk} - \sum_{j:(j,i)\in A} x_{jik} = 0 \quad \forall\, i \in N, \quad \forall\, k \in K, \tag{4}$$

$$t_{ik} + (Tij + D_i)x_{ijk} \leq t_{jk} + b_i(1 - x_{ijk}) \quad \forall\, k \in K, \quad \forall\, (i,j) \in A, \tag{5}$$

$$a_i \sum_{j:(i,j)\in A} x_{ijk} \leq t_{ik} \leq b_i \sum_{j:(i,j)\in A} x_{ijk} \quad \forall\, k \in K, \quad \forall\, i \in N, \tag{6}$$

$$a_i^k \leq t_{ik} \leq b_i^k \quad \forall\, k \in K, \quad \forall\, i \in o, d, \tag{7}$$

$$\sum_{k\in K} t_{ik} = \sum_{k\in K} t_{jk} \quad \forall\, (i,j) \in P^{sync}, \tag{8}$$

$$\sum_{k\in K} t_{ik} + p_{ij} \leq \sum_{k\in K} t_{jk} \quad \forall\, (i,j) \in P^{prec}, \tag{9}$$

$$x_{ijk} \in \{0,1\} \quad \forall\, k \in K, \quad \forall\, (i,j) \in A, \tag{10}$$

$$t_{ik} \in \mathbb{Z}_+ \quad \forall\, k \in K, \quad \forall\, i \in N. \tag{11}$$

In this model, N represents the set of customer locations. o, d refer to the same starting and ending location: o is the starting one and d the ending one. Note that two different nodes are required for modelling purposes. The set A contains all locations N plus $o$and d. The set of employees is defined by K. Time window for activity i is given by a_i and b_i as the earliest start time and the latest start time respectively. The constant D_i indicates the duration of activity i. Travel time between the location of activities i and j is given by T_{ij}. Variable t_{ik} is a decision variable that defines the starting time of i when performed by employee k. The employees' working time is given by a_i^k and b_i^k which represent the start and the end working time for employee k. Finally, The term E_{ij} includes the travelling time from i to j plus the duration of visit at node i ($E_{ij} = T_{ij} + D_i$).

The objective function (1) is composed by the cost for assigning visits to employees c_{ik} and the travel time of all employees when performing their visits. We set the weights in the objective function to the same value $\alpha_p = \alpha_T$. Constraints are as follows. All visits must be performed (2). Employees must start and return to the same location (3). Constraint (4) maintains flow conservation among employees. Time windows of visits must be satisfied (5, 6). Visits should be performed during the employees starting and ending times (7). A Synchronisation constraint (8) is necessary for every pair of visits. Other type of temporal dependencies are covered by constraint (9). Decision variables x_{ijk} are set to 1 when employee k travels from location i to j and 0 otherwise (10). Scheduling variables are positive integers (11). In the data sets, all employees have the same starting and ending time which match the time horizon of every instance, and times are considered after the start of the time horizon.

The modifications to the Bredström and Rönnqvist (2008) model include the change of the scheduling variables t_{ik} from being real positives to being positive integers (11) so that t_{ik} represents a given minute in the scheduling period. Another modification to the original model is in the objective function (1), the balancing component w and its weight α_B (see Bredström and Rönnqvist, 2008, pg. 25) were

removed. Such balancing component could be included when factors like fairness on service time or workload for every employee are taken into account. The other weights of the objective function α_p and α_T remain the same (set to a value of 0.5).

5.3 Initial Experiment Results

For every one of the original 56 Solomon's instances with 100 activities (customers), we generated two more with 25 and 50 activities. For example, from C101 we generated C101_25, C101_50 and C101_100. From Castro-Gutierrez et al instances we selected the ones with 50, 150 and 250 activities, 5 time window types and limited vehicle capacity for all. We used a total of 183 VRPTW instances, 168 from Solomon (56 x 3 instance sizes) and 15 from Castro-Gutierrez et al (3 instance sizes x 5 window types). The integer programming model was implemented in the Gurobi solver version 5.1 and executed on a X64-based PC running Microsoft Windows 7 Enterprise with 2 Duo CPU (3.16GHz) and 4 gigabytes of RAM. There was no parameter tuning for Gurobi, it ran using the default parameter values (no computation time limit).

Table 1: Summary of the outcome infeasible, optimal, out of memory (OoM) reported by Gurobi. Computation times (Min, Mean, Max, Std Dev) are in seconds. Times in brackets show when the solver found the first feasible solution when available. [a] Solomon data set. [b] Castro-Gutierrez et al data set.

Size	Outcome	No. of Instances	Min	Mean	Max	Std Dev
25	Infeasible	30^a	0	5417.80	162340	29637.88
	Optimal	2^a	79072 (8)	160346.00 (10.5)	241620 (13)	114938.79 (3.53)
	OoM	21^a	38430 (188)	117198.48 (2329.0833)	348129 (9293)	75519.83 (2599.1286)
50	Infeasible	31^a	0	24.29	331	73.49
	Optimal	0^a	-	-	-	-
	OoM	$26^{a,b}$	1651 (63)	67294.31 (5365)	269150 (14898)	61362.68 (8273.169)
100	Infeasible	53^a	0	55.64	123	32.27
	Optimal	0^a	-	-	-	-
	OoM	0^a	-	-	-	-

Table 1 summarises the results from these initial experiments. Column Outcome gives the outcome of the experiments. Column Size shows the instances size regarding the number of activities. Column No. of Instances indicates the number of instances of that size for which the solver reports the given outcome. Columns Min, Mean and Max provide the minimum, mean and maximum computation time in seconds within that group of instances where the given outcome was obtained. Column Std Dev shows the standard deviation of the computation time. For example, we can see that out of the 61 instances of size 50 (56 from Solomon plus 5 from Castro-Gutierrez et al), the solver reported infeasibility in 31 cases, reported out of memory for 26 cases and could not find optimality in any case. We also note that in 10 instances (3 instances of size 25, 4 of size 50 and 3 of size 100) some computing error was reported by the solver when processing the instance.

Also, none of the 10 largest instances (sizes 150 and 250) from Castro-Gutierrez et al could be processed with Gurobi using our hardware setting. Adding these 20 instances to the 163 reported in Table 1 gives the total of 183.

As expected, the overall observation is that these VRPTW instances are computationally difficult to solve. More importantly, we observe that for a good proportion of them no feasible solution was obtained. A closer look at the results revealed that many of the infeasible instances did not have enough workers to complete all activities (hard-constraint). Also, the generated precedence constraints between activities provoked infeasibility in several cases when interacting with the time window constraints. Another useful observation from these results is that the mean computation time taken by the solver to find a first feasible solution is between 10.5 (row 5) and 5365 (row 11) seconds. Hence, we decided to set 7200 seconds as the maximum execution time in the next set of experiments.

Figure 1 gives an insight into the performance of the solver in one of instances of size 25 for which the optimal solution was found. The main solid line corresponds to the gap reduction as the computation time progresses. The inside graph shows a close-up of the first 14400 seconds. We can see that after finding a feasible solution relatively quickly (no more than 13 seconds according to row 5 of Table 1), the solver continuously improves the solution but finds the optimal only after around 67 hours of computation time. The red square shows that the greater gap reduction occurs during the first 2 hours (7200).

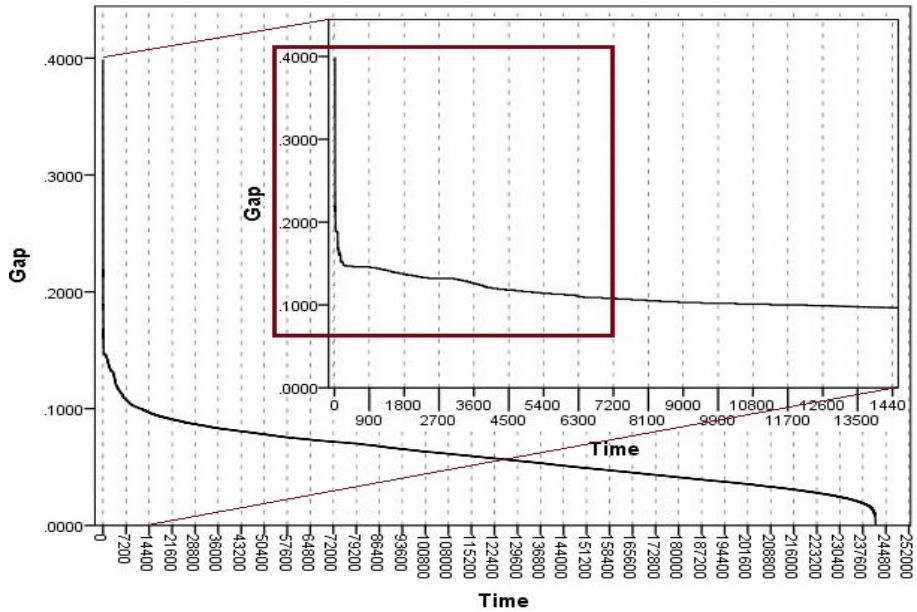

Fig. 1: Gap reduction as computation time progresses in a case in which the solver found the optimal solution. The optimal solution is reported after 241620 seconds (approx. 67 hours) but a considerable gap reduction is achieved during first two hours.

One conclusion from these experiments is that tackling the VRPTW instances as WSRP instances is not very useful unless care is taken into specifying the nature (hard or soft) of the scheduling-all-activities and activities-precedence requirements. Another conclusion is that WSRP with more than 100 activities are

still beyond the capacity of a mathematical programming solver running on a standard personal computer, but for instances of smaller size we can expect to generate good quality feasible solutions within a couple of hours which is also a reasonable time to wait for an automated system to produce a solution in real-world scenarios.

5.4 Experiments with an Mixed Integer Programming Model

It was clear from the previous experiments that if the constraint that allocates all activities is not relaxed, the solver would not be able to provide feasible results for many scenarios. Moreover, with the integer model no information is provided on the number of activities that could not be performed. In this section we used a mixed integer programming (MIP) model which we believe is more suitable for WSRP. It introduces another binary decision variable for every activity per employee. This model, although requiring more memory resources, provides feasible solutions for all instances and allows the solver to provide a lower bound and report on gap between the lower bound and the feasible solution found. In this second set of experiments the MIP model is used for all 375 instances in all 5 data sets. The only difference in the experiment settings is that for each instance the solver is now allowed to run for two hours only, unlike the unlimited execution time that was used before.

Following our observations in the previous experiments, we now relax the constraint requiring all activities to be performed (constraint (2) in the IP model). We then use the MIP model by Rasmussen et al (2012) to tackle all the adapted instances from the 5 data sets. This model introduces a binary variable y_i to indicate whether activity i is performed or not (constraint (13) in the MIP model). This MIP model requires the set of variables t_i^k, corresponding to activity i's start time by employee k, to be a real positive value (see eq. (23)). The cost C_{ij}^k in the objective function (12) is calculated as travel time plus distance from i to j. It is important to notice that in the Solomon's based instances these two values (travel time and distance) are the same. As it can be noticed C_{ij}^k differs from the way the cost c_{ik} of the previous IP Model is calculated. In the IP model c_{ik} depends on the activity i and the employee k. Whereas in the MIP model the cost C_{ij}^k varies depending the starting point i and the ending point j. The weights $(\omega_1, \omega_2, \omega_3)$ are calculated as in Rasmussen et al (2012).

C represents the set of customer locations. The variables 0^k and n^k refer to the starting and ending location for each employee k. This model supports different starting and ending locations for each employee. The set of employees is defined by K. $N^k = C \cup \{0^k, n^k\}$ is the set of available locations for employee k. Time window for activity i is given by α_i for earliest start time and β_i for latest start time. The binary variables ρ_i^k indicate whether employee k can perform activity i. These binary variables are used to match skill set or any other characteristic that prevents the allocation of employees to activities. Binary variables y_i are set to 1 if activity i is left uncovered and 0 if an employee is assigned to it. Variables s_{ij}^k indicate the duration of travel time from activity i to activity j for employee k. These variables already include the duration of activity i when performed by employee k. Scheduling variables t_{ik} capture the starting time of activity i when performed by employee k. Employees' working time is given by the terms α_{n^k} and β_{n^k} which represent the start and the end working time for employee k. Connected

activities constraints are indicated by a set of pairs of activities P and for every pair of activities i and j a constant value p_{ij} is given depending on the type of connected activity constraint. Finally, binary variables x_{ij}^k are set to 1 if employee k moves to activity j after performing activity i.

$$min \qquad \omega_1 \sum_{k\in K}\sum_{i\in N^k}\sum_{j\in N^k} C_{ij}^k x_{ij}^k + \omega_2 \sum_{k\in K}\sum_{i\in C}\sum_{j\in N^k} \delta_i^k x_{ij}^k + \omega_3 \sum_{i\in C} \gamma_i y_i \quad (12)$$

$$s.t. \qquad \sum_{k\in K}\sum_{j\in N^k} x_{ij}^k + y_i = 1 \qquad \forall\, i \in C, \quad (13)$$

$$\sum_{j\in N^k} x_{ij}^k \le \rho_i^k \qquad \forall\, k \in K, \quad \forall\, i \in C, \quad (14)$$

$$\sum_{j\in N^k} x_{0^k,j}^k = 1 \qquad \forall\, k \in K, \quad (15)$$

$$\sum_{i\in N^k} x_{i,n^k}^k = 1 \qquad \forall\, k \in K, \quad (16)$$

$$\sum_{i\in N^k} x_{ih}^k - \sum_{j\in N^k} x_{hj}^k = 0 \qquad \forall\, k \in K, \quad \forall\, h \in C, \quad (17)$$

$$\alpha_i \sum_{j\in N^k} x_{ij}^k \le t_i^k \le \beta_i \sum_{j\in N^k} x_{ij}^k \qquad \forall\, k \in K, \quad \forall\, i \in C \cup \{0^k\}, \quad (18)$$

$$\alpha_{n^k} \le t_{n^k}^k \le \beta_{n^k} \qquad \forall\, k \in K, \quad (19)$$

$$t_i^k + s_{ij}^k x_{ij}^k \le t_j^k + \beta_i(1 - x_{ij}^k) \qquad \forall\, k \in K, \quad \forall\, i,j \in N^k, \quad (20)$$

$$\alpha_i y_i + \sum_{k\in K} t_i^k + p_{ij} \le \sum_{k\in K} t_j^k + \beta_j y_j \qquad \forall\, (i,j) \in P, \quad (21)$$

$$x_{ij}^k \in \{0,1\} \qquad \forall\, k \in K, \quad \forall\, i,j \in N^k, \quad (22)$$

$$t_i^k \in \mathbb{R}_+ \qquad \forall\, k \in K, \quad \forall\, i \in N^k, \quad (23)$$

$$y_i \in \{0,1\} \qquad \forall\, i \in C. \quad (24)$$

The objective function (12) is composed by the cost C_{ij}^k previously defined, the preference δ_i^k that employees have when performing activity i and the priority of the visit i given by γ_i. We set the weights in the objective function to the same values suggested by Rasmussen et al (2012). Such weights favour the allocation of all activities when possible by giving ω_3 a significant greater value than ω_2 and similarly ω_2 is greater than ω_1. Constraints are as follows. Visits are either performed or left unassigned (13). Activities can only be assigned to employees able to perform them (14). All employees must start from the initial location (15) and return to their own final location (16). Constraint (17) maintains flow conservation among employees. Time windows of visits must be satisfied (18,20). Visits should be performed during the employees starting and ending times (19). Connected activities constraints exist among related activities (21). Decision variabless x_{ij}^k are set to 1 when employee k travels from location i to j and set to 0 otherwise (22). Scheduling variables are positive integers (23). Finally, if an activity i is not performed binary decision variables y_i is set to 1 and 0 otherwise.

Some instances in the data sets were changed in order to require two employees to perform some tasks, since this type of teaming data was not present in any of the original data sets. We implement this by creating an additional virtual activity for every additional employee required. Then, we enforce a synchronisation constraint between every pair of activities, i.e. the original and the virtual activity (Bredström and Rönnqvist, 2008, p. 24). Additionally, we forbid transition between these pairs of activities since an employee performing one of the activities cannot perform the other activity too, because the aim is to have different employees. The prohibition is achieved by never creating extra binary variables x_{ij}^k between those activities. This approach is illustrated in Figure 2.

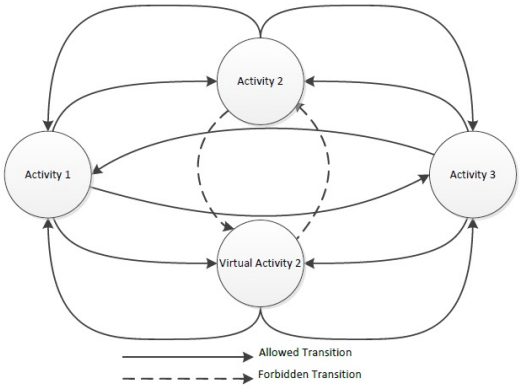

Fig. 2: Modelling a virtual activity by creating a node (virtual activity 2) for the case in which Activity 2 requires two employees. The dotted lines correspond to decision variables that are not created for every employee, ensuring two employees are assigned to Activity 2 while reducing the size of the MIP model.

5.5 Adapted Problem Instances

In order to use the 5 data sets mentioned in section 5.1, we made some adaptations. In this section we describe such adaptations with reference to the key WSRP features discussed in the survey part of this paper. The prefixes used to identify each of the data sets are as follows: *Sol* refers to the 168 instances from Solomon (1987), *Mov* refers to the 15 instances from Castro-Gutierrez et al (2011), *HHC* refers to the 11 instances from Rasmussen et al (2012), *Sec* refers to the 180 instances from Misir et al (2011) and *Tech* refers to the single instance from Günther and Nissen (2012), for a total of 375 problem instances.

Number of Employees and Activities

One of our objectives is to asses the difficulty of instances as employees and activities increase. The Mov, HHC, Sec and Tech data sets have a predefined number of activities and employees.The Sol data set has only a number of activities, so we set the number of employees in those instances to be a fifth of the number of activities. This value was decided following our conversations with a service

organisation, and it also matches the assumption by Bredström and Rönnqvist (2008). The organisation told us that their average visit duration is currently 50 minutes. Current shift is 8 hours (480 minutes) with a break of 1 hour (60 minutes), or two breaks (30 minutes). Taking this into account, we determine the mean number of visits per employee per day x by solving the simple equation $50x + 30(x+1) + 60 = 480$, where $x+1$ is the number of trips in a route including the last trip to the employee's final destination. And, 30 minutes is the average time between visits, which include travelling and idle time. The result is $x = 4.875$ rounded to 5.

Table 2: Number of activities per employee in each data set. [a] Only one instance so no standard deviation provided.

Data Set	Min	Max	Average	StdDev
Sol	5	5	5	0
Mov	1.31	1.46	1.411	0.06
HHC	8.57	11.77	11.05	1.31
Sec	4	4.42	4.07	0.00
Tech	1.73	1.73	1.73	-[a]

Table 2 indicates the number of activities per employee in each group of instances. Notice that within the Mov instances, it appears to be one employee per activity (average 1.411). In this group there are a few employees on training and cannot be assigned on their own to any activity due to lack of skills, but remain included in the workforce. We acknowledge that leaving unskilled employees in the workforce adds variables and constraints to the model but those are identified and removed in the pre-processing state of the solver. Figure 3 gives an indication of the ratio between the number of activities and number of employees. The plots at the bottom of the figure belong to instances in HHC, Sec and Sol, where it is clear that an employee has to perform more activities when compared to Mov and Tech.

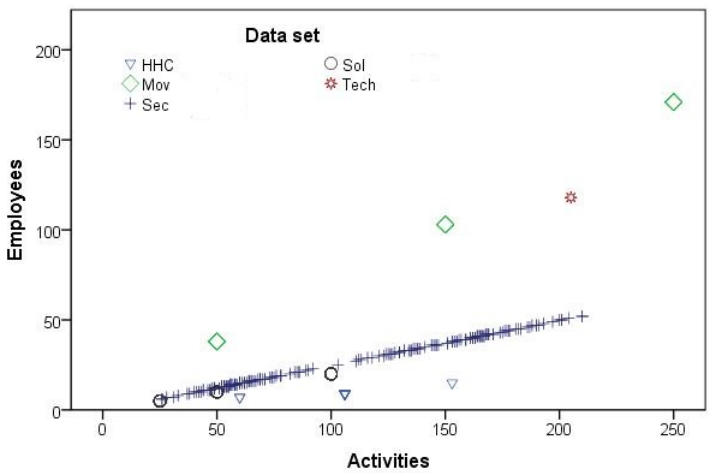

Fig. 3: Ratio between the number of activities and the number of employees in each data set.

We identify groups of instances based on the number of employees as shown in Figure 4, where the number of instances in each group is given. For example, HHC_9 refers to the 8 instances from HHC data set that have 9 employees. Because the Sec instances have greater variability in the number of employees (see Figure 3) we group them in 10 clusters (Sec_10 to Sec_55) and assign each instance to the closest upper cluster e.g. instances with 6 or 7 employees are assigned to cluster Sec_10 and instances with 11 employees are assigned to Sec_15.

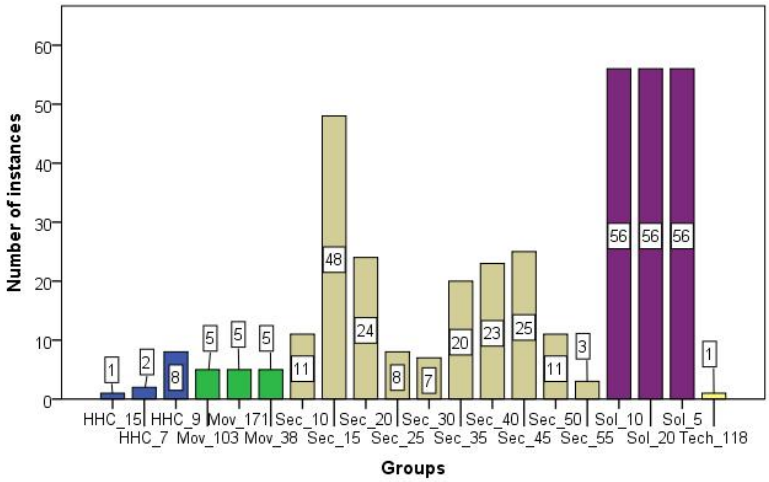

Fig. 4: Groups of instances based on number of employees.

Activities Duration

The activities duration affects the number of activities that can be performed during the scheduling period. All data sets include the duration for every activity so this feature is not adapted. In the Sec instances, the average duration is between just above 5 hours for instances with up to 25 employees and 8 hours for instances with more than 25 employees. All other data sets have an average activity duration of 20 to 30 minutes. Figures 5 and 6 show the range of values for each data set activities duration. We can see that most instances have a wide range of activities durations with the exception of the Mov ones, where duration varies between 10 and 30 minutes only.

Time Horizon

The time horizon values for all instances in Sol, Mov, HHC and Tech are given in the original data set. The time horizon in the original 6 instances of the Sec data set is 1 month. Hence, we adapted these scenarios by splitting each 1-month instance into 30 1-day instances (180 instances in total). Figure 7 shows the range of time horizon lengths in the instances arising from each data set. We can see that the majority of instances have a time horizon of 1 day (1440 min) or less and only 24 instances have a time horizon of just over two days (3390 min).

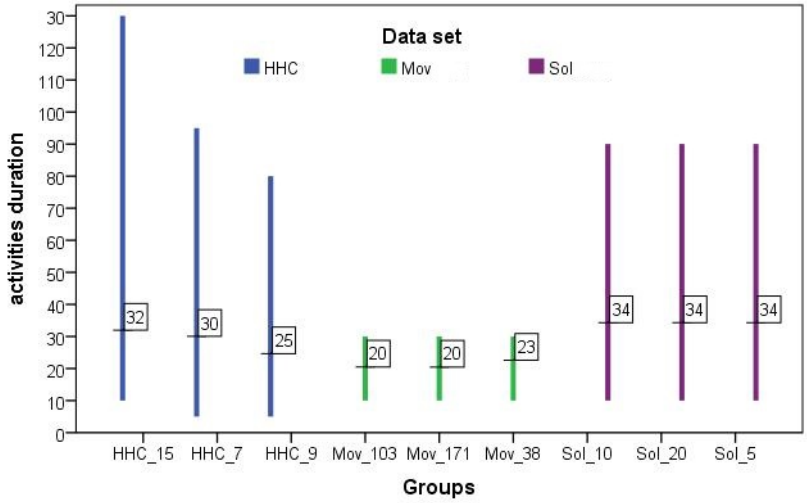

Fig. 5: Range of activities duration (in minutes) values for HHC, Mov and Sol instances. The figure show the minimum duration represented by the bottom of each bar, the average duration contained in the label and the maximum duration at top of each bar.

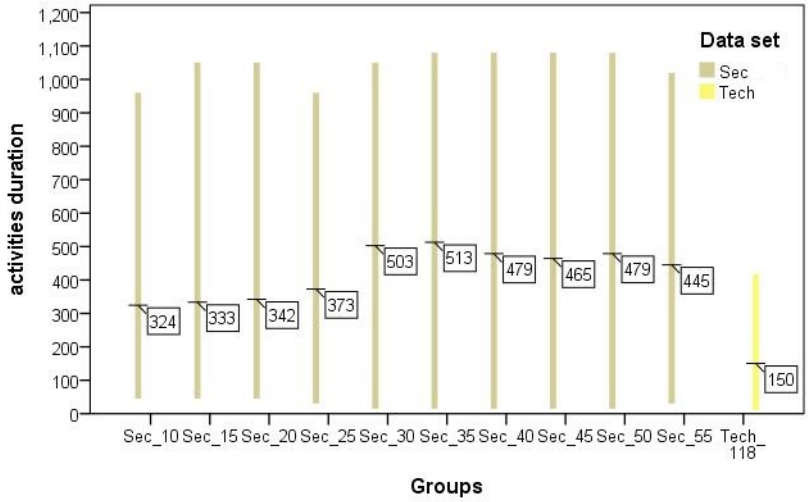

Fig. 6: Range of activities duration (in minutes) values for Sec and Tech instances. The figure show the minimum duration represented by the bottom of each bar, the average duration contained in the label and the maximum duration at top of each bar.

Employee Skills

An homogeneous highly-skilled workforce tends to facilitate the allocation of tasks because an employee can easily be replaced by another one with the same or equivalent set of skills. Such workforce could be more expensive than having a workforce with a wide range of skill levels. Figure 8 shows the average coverage of activities that an employee has given his skills. For example, in the group of

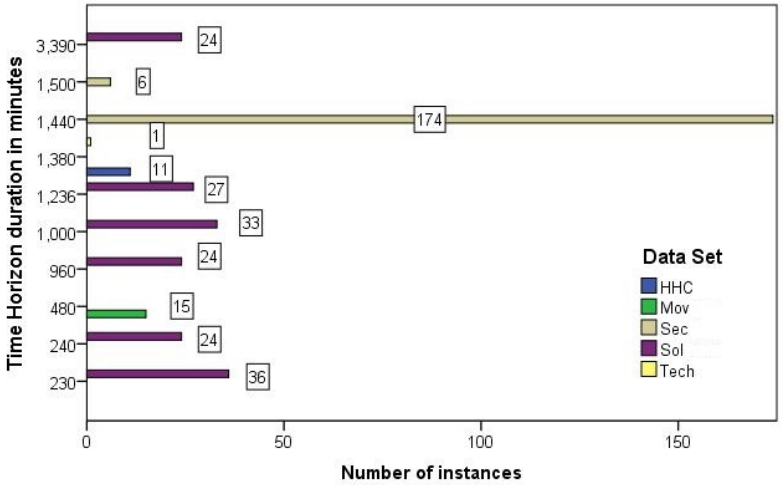

Fig. 7: Time horizons' lengths for instances in each data set (after adapting the Sec ones). The majority of instances are a day or less 1,444 min.

instances Mov_171 the average employee can perform 68% of the activities due to skill matching. Nevertheless, in the same group there are also employees that cannot perform any activity and those that can perform any of them. We can see that in most of the instances there is at least one employee that cannot perform any activity when their skills are taken into account. The HHC instances represent scenarios where most of the employees have the skills that allow them to perform a high percentage (above 80%) of the activities.

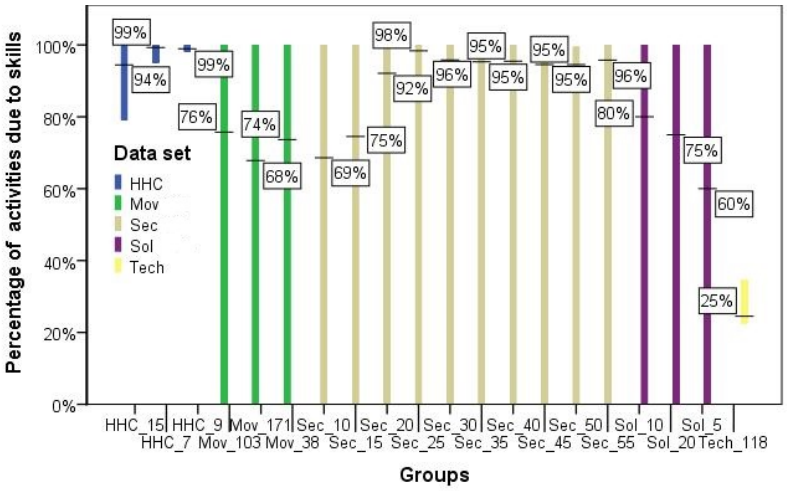

Fig. 8: Percentage of activities that employees can perform given their set of skills. The label in each bar indicates the percentage of activities that the average employee can perform.

Time Windows

Instances in the considered data sets have a range of time window sizes. Figure 9 depicts the minimum, average and maximum time windows size for each group. The Sol instances have the greatest variability of time window sizes ranging from a strict starting time, time window duration 0, to a time window duration of the size of the time horizon. Nevertheless, some instances in the Sol groups have the same time window size for all the activities in the instance. For example, Figure 10 shows the time window duration for individual instances in the group Sol_20. In some cases, all activities have the same time window size (no bar just a line in the graph).

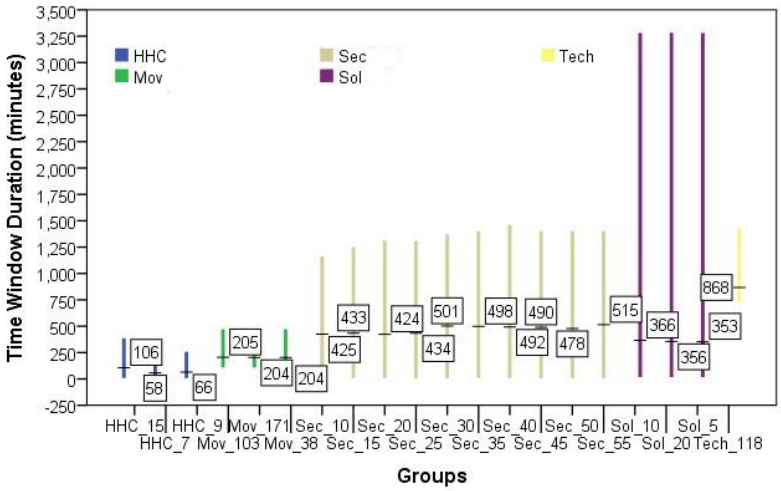

Fig. 9: Average, minimum and maximum time window duration (in minutes) for HHC, Mov, Sec, Sol and Tech groups.

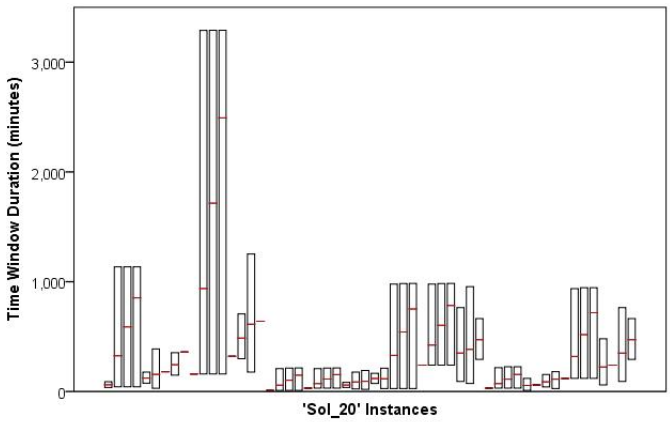

Fig. 10: Average, minimum and maximum time window duration (in minutes) for the 56 instances in the group Sol_20. Some instances do not have diversity in their time windows duration.

Teaming

Teaming is used when activities require more than one employee. The original data sets did not provide information regarding this feature although it is often mentioned in the literature. Our approach for adapting the data sets to include this feature was to set the need for having 2 employees to perform an activity with a probability of 0.10 for the Sol and Mov instances and, a probability of 0.20 for the Sec instances. The HHC and Tech instances were left requiring one employee only per activity. Then on average, activities in the Mov, Sec and Sol instances require more than 1 employee and this is illustrated in Figure 11.

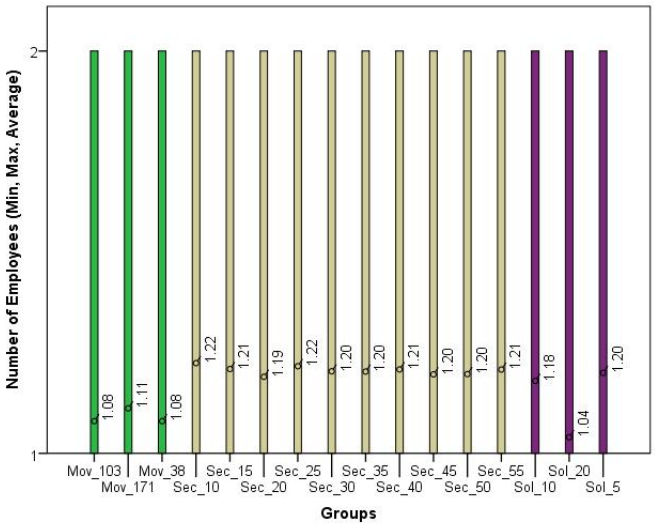

Fig. 11: Mov, Sec and Sol groups of instances (where teaming was generated). This graph shows the average, minimum and maximum number of employees needed per activity.

Connected Activities

As explained earlier in section 2.2, connected activities arise when there is a dependency between two activities. Such dependencies, although present in the real world, are not incorporated in most instances from the literature with the exception of those by Rasmussen et al (2012). Since the HHC instances already incorporate connected activities, no adaptation was necessary. However, since the other 4 data sets do not contain this information, we adopt the 5 types of precedence constraints: synchronisation, overlap, minimum difference, maximum difference and minimum-maximum difference proposed in (Rasmussen et al, 2012, p. 601) and include them in the groups of instances Sol, Mov, Sec and Tech.

In the procedure to generate connected activities we needed to consider the already given time windows for each activity. Otherwise, we could create constraints not possible to satisfy (as experienced in our first set of experiments). For each activity, the procedure seeks to set a connection with another activity with a probability of 0.25 and then the type of connection (precedence constraint) to establish

is also selected with some probability. If the selected activities to be connected by the selected precedence represent a constraint that makes the instance infeasible (e.g. due to time window clashes), then the proposed constraint is discarded and the next one generated by the procedure is tried. The probability of each type of constraint in every data set varies as shown in Table 3. Such information is useful if experiments are to be reproduced or compared given a different probability for the type of connected activity constraint. As it can be noticed the Sol and Mov instances favour the synchronisation and overlapping constraints, each of them with 0.35, whereas Sec and Tech instances used the same probability (0.2) when generating the constraints.

Table 3: Probabilities for adding each type of precedence constraint when generating connected activities in the Sol, Mov and Sec instances.

Connected Activity	Sol	Mov	Sec	Tech
Synchronisation	0.35	0.35	0.2	0.2
Overlap	0.35	0.35	0.2	0.2
Minimum difference	0.1	0.1	0.2	0.2
Maximum difference	0.1	0.1	0.2	0.2
Min-max differnece	0.1	0.1	0.2	0.2

5.6 MIP Experiment Results

This section presents the results of our experiments using the MIP model (see section 5.4) to tackle all the instances adapted as described above. The Gurobi solver version 5.1 was used with all default parameters left unchanged, with exception of the computing time limit set to 2 hours maximum, a value obtained previously (see Fig. 1), for each instance. Out of the 375 instances, the solver found feasible solutions for 356 of them. For 11 of the remaining 19 instances, it was not possible to load the model with our settings. We observe that all these instances are the larger ones in the Mov group (those with 150 or more activities). However, this is not a definitive indication of the instance difficulty because some scenarios with more than 150 activities were solved, although in those cases the number of employees was less than 100. The scenario with the largest number of activities and employees for which Gurobi found a solution, has 210 activities and 52 employees. For the other 8 of the 19 instances, the solver reported an out of memory error very early in the process. These instances are part of the Sec group.

Note on the Number of Employees

The Mov_103 group could not be loaded despite other instances with up to 210 activities were loaded and provided with feasible solutions by the solver. We believe that the large number of employees is the issue here. If most activities in the Mov_103 instances had a duration of more than 50% of the time horizon length, it could be said the number of employees is not too large since an employee could only perform one activity during the time horizon. But this is not the case. Instances in Mov_103 have an average activity duration of 20 minutes as shown in Figure 5 and a time horizon of 480 minutes (see Fig. 7). Therefore, employees in

these instances are able to perform more than two activities per day, resulting in having much more employees than needed. It could also be said that, if on average employees are not skilled enough to perform all activities, then the number of employees is not too large. But this is not the case either because the average employee in these instances can perform 76% of the activities as shown in Figure 8. Despite this, we left the number of employees in all Mov instances unchanged (the deliberate over-staffing in these instances was confirmed by the authors of those Mov instances). Including all employees significantly increases the number of constraints and variables in the model to a degree that such instances could not be loaded into the solver given our hardware settings. Reducing the number of available workers while maintaining all other features unchanged in these scenarios and solving them is kept for future investigation.

Analysis of Feasible Results

During the optimisation process, Gurobi provides information about the gap between the lower bound and the current feasible solution (if it exists). Such Gap for the 356 solved instances is shown in Figure 12, where we can observe a widespread range of values across the instances. The figure uses the id of each instance in the horizontal axis: Sol [1, 168]; Sec [169, 348]; Mov [349, 363]; HHC [364, 374] HHC, Tech [375]. It is noticeable that optimality was achieved for some of the Sol and Mov instances (gap value of zero). On the contrary, for Sec the solver reported a gap of more than 50% in most of the cases (see Fig. 13).

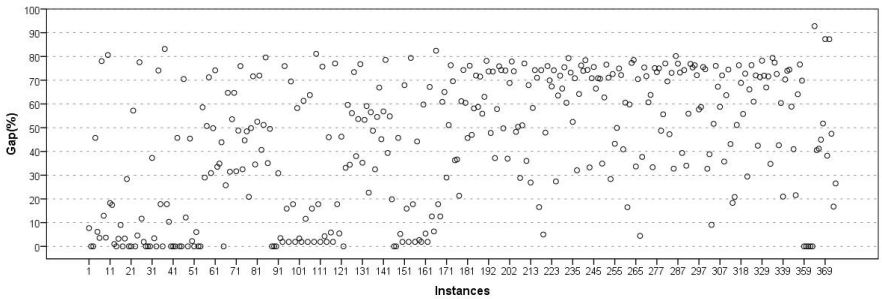

Fig. 12: Percentage gap values for 356 instances where the solver found feasible solutions.

In order to achieve a better understanding of the difficulty of the WSRP, we further analyse the gap results for the instances in each data set. Figure 13 shows the gap results for the instances that were adapted from each of the original data sets (except Tech which could not be solved). It is clearer that the Sec instances are more difficult to solve given the larger number of high gap values reported. This can be confirmed in the corresponding box plot in Figure 14 which shows a median of 67%. On the contrary, low gap values were reported for the Sol instances with a median of just above 20%. For the mayority of HHC instances the gap value is between 40% and 70%. In the HHC and Sec data sets, no instance was solved to optimality. The worst gap value was reported within HHC (92% for hh_00_WSRP). The results shown for the Mov data sets are for the 5 solved instances all with 50 activities and 38 employees. For all of them an optimal solution was found.

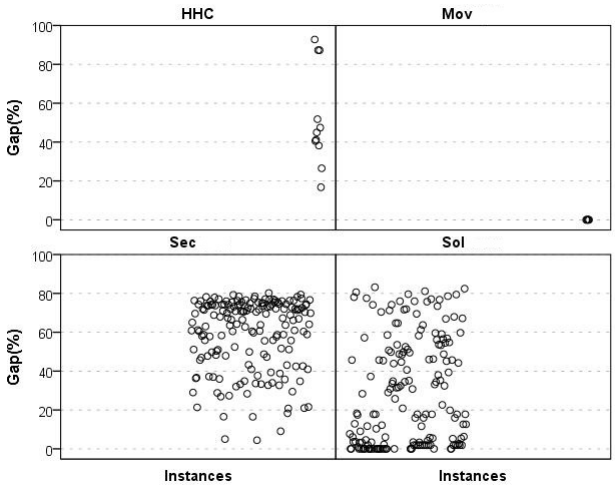

Fig. 13: Distribution of gap percentage values reported for each instance in each data set.

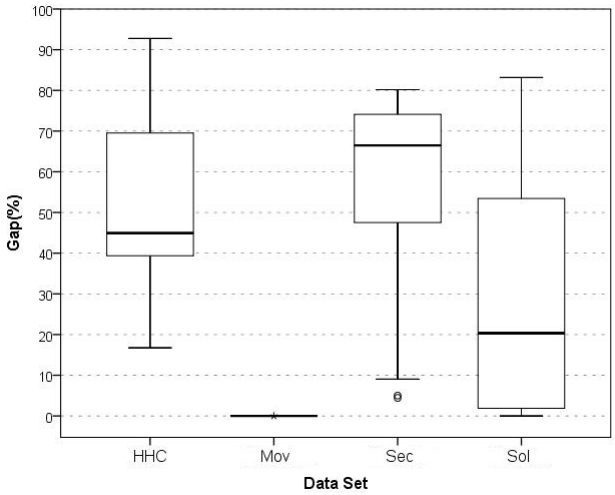

Fig. 14: Aggregated gap percentage values reported for the instances in each data set.

Figure 15 plots the gap values reported for all instances within different groups with respect to the number of employees. Looking at the three HHC groups it is clear that the smaller the size of the workforce, the better the gap value achieved. A similar observation can be made for the Sec instances, the achieved gap value worsens as the number of employees increases (the group Sec_55 is not reported because all three instances are amongst the 19 instances that could not be solved in our experiments). Note that this tendency is not clear for the Sol instances. Looking at the box plot of the Sol group in Figure 16 we can see that 50% of the instances with 10 employees (Sol_10) report a gap value below 5%. Also, 50% of the instances with 5 employees (Sol_5) report a gap of 15%. That is, more instances in Sol_10 report a better gap than instances in Sol_5.

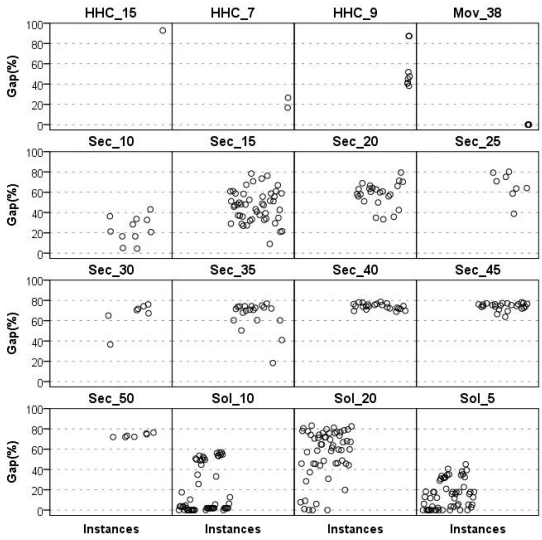

Fig. 15: Distribution of gap percentage values reported for each instance grouped by data set and number of employees.

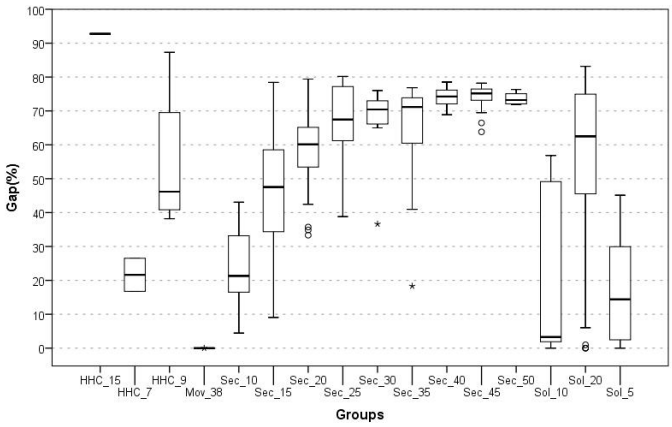

Fig. 16: Aggregate gap percentage values reported for each group of instances according to the number of employees.

6 Conclusion

A workforce scheduling and routing problem (WSRP) refers to any environment in which a skilled diverse workforce should be scheduled to perform a series of activities distributed over geographically different locations. Activities should be performed at specific times or within a given time window. The time window for each activity is usually determined by the customer or recipient of the job. The survey in the fist part of this paper was aimed at identifying problems reported in the literature that can be seen as a WSRP scenario. The problems identified include but are not limited to: home health care, home care, scheduling of technicians, security personnel routing and rostering, and manpower allocation. The

survey also sought to identify similarities between these problems in order to define the core characteristics that any WSRP would have. Those core characteristics include: time windows, transportation modalities, start and end locations, a diverse skilled workforce, activities service time and precedence relationship between activities. In addition, there are other characteristics such as personnel teaming and clusterisation of locations among perhaps others, that sometimes arise in WSRP.

The survey part of this paper also sought to identify the solution methods that have been employed in the literature when tackling WSRP scenarios. Since the WSRP combines employee scheduling and vehicle routing with time windows, there seems to be a clear tendency to apply exact approaches for the routing component of the problem, and to apply heuristics for the matching of employees to activities. Among other approaches for solving WSRP, we found mixed integer linear programming (MIP), integer linear programming (IP) and constraint programming (CP) using set partitioning and multi-commodity network flow models. Also, a variety of meta-heuristics such as tabu search (TS), particle swarm optimisation (PSO) and simulated annealing (SA) have been applied to tackle WSRP.

In the second part of this paper we presented our computational study. The study used two models, an IP and an MIP to tackle instances adapted from the literature. The instances belong to problems that could be treated as WSRP after some adaptations such as increasing the number of employees for an activity, and adding skill matching and precedence constraints, etc. We performed an analysis of the characteristics of the instances taking into account: the size of their time horizon which for the majority is less than a day; the duration of the activities which range from a few minutes to up to 8 hours; the abilities of the workforce given their skills; and the size of the activities' time windows.

Results when using the IP model suggest that, if the number of activities reaches 100 and we aim to schedule all of them, a commercial solver running on a conventional personal computer struggles to reach an optimal solution and in some cases even a feasible one. If we allow activities to be unscheduled at a cost, then feasible solutions are found for instances with up to 210 activities. Finally, we suggest that setting an adequate number of available employees should take into account their skill set and the duration of activities. Our results showed that sometimes instances with fewer employees could be more difficult to solve if those employees are less skilled.

We consider some extensions for future work. Firstly, to perform more analysis into the interaction between the WSRP characteristics by varying each data set independently. Secondly, to apply a different MIP model or extend the one by Rasmussen et al (2012), seeking to include other features such as: employees capacity (number of hours allowed to work within the time horizon), employees breaks, and balancing the number of activities in routes. Lastly, in our study we have used exact approaches only, other algorithms such as metaheuristic and evolutionary computation should be explored to tackle our adapted WSRP instances.

Acknowledgements We acknowledge the authors (Castro-Gutierrez et al, 2011; Rasmussen et al, 2012; Misir et al, 2011; Günther and Nissen, 2012) who kindly provided us the original data sets to perform this study.

References

Akjiratikarl C, Yenradee P, Drake PR (2006) An improved particle swarm optimization algorithm for care worker scheduling. In: Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference, Bangkok, Thailand, pp 457–466

Akjiratikarl C, Yenradee P, Drake PR (2007) Pso-based algorithm for home care worker scheduling in the uk. Computers & Industrial Engineering 53(4):559–583

Alfares HK (2004) Survey, categorization and comparison of recent tour scheduling literature. Annals of Operations Research 127(1-4):145–175

Baker KR (1976) Workforce allocation in cyclical scheduling problems: A survey. Journal of the Operational Research Society 27(1):155–167

Barnhart C, Johnson EL, Nemhauser GL, Savelsbergh MW, Vance PH (1998) Branch-and-price: Column generation for solving huge integer programs. Operations Research 46(3):316–329

Begur SV, Miller DM, Weaver JR (1997) An integrated spatial dss for scheduling and routing home-health-care nurses. Interfaces 27(4):35–48

Bertels S, Fahle T (2006) A hybrid setup for a hybrid scenario: combining heuristics for the home health care problem. Computers and Operations Research 33(10):2866–2890

Blais M, Lapierre SD, Laporte G (2003) Solving a home-care districting problem in a urban setting. Journal of the Operational Research Society 54(11):1141–1147

Borsani V, Matta A, Sommaruga F, Beschi G (2006) A home care scheduling model for human resources. In: Proceedings of the International Conference on Service Systems and Service Management, vol 1, pp 449–454

Brandão J, Mercer A (1997) A tabu search algorithm for the multi-trip vehicle routing and scheduling problem. European Journal of Operational Research 100(1):180–191

Brandão J, Mercer A (1998) The multi-trip vehicule routing problem. Journal of the Operational Research Society 49(8):799–805

Bredström D, Rönnqvist M (2007) A branch and price algorithm for the combined vehicle routing and scheduling problem with synchronization constraints. NHH Dept of Finance & Management Science Discussion Paper No 2227/7

Bredström D, Rönnqvist M (2008) Combined vehicle routing and scheduling with temporal precedence and synchronization constraints. European Journal of Operational Research 191(1):19–31

Castro-Gutierrez J, Landa-Silva D, Moreno-Perez J (2011) Nature of real-world multi-objective vehicle routing with evolutionary algorithms. In: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp 257 –264

Cheang B, Li H, Lim A, Rodrigues B (2003) Nurse rostering problems a bibliographic survey. European Journal of Operational Research 151(3):447–460

Cheng E, Rich JL (1998) A home health care routing and scheduling problem

Chuin Lau H, Gunawan A (2012) The patrol scheduling problem. In: Proceedings of the 9th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2012), Son, Norway, pp 175–192

Cordeau JF, Laporte G, Pasin F, Ropke S (2010) Scheduling technicians and tasks in a telecommunications company. Journal of Scheduling 13(4):393–409

De Angelis V (1998) Planning home assistance for aids patients in the city of rome, italy. Interfaces 28(3):75–83

Desaulniers G, Lavigne J, Soumis F (1998) Multi-depot vehicle scheduling problems with time windows and waiting costs. European Journal of Operational Research 111(3):479–494

Desrochers M, Desrosiers J, Solomon M (1992) A new optimization algorithm for the vehicle routing problem with time windows. Operations research 40(2):342–354

Desrosiers J, Lübbecke ME (2005) A primer in column generation. In: Desaulniers G, Desrosiers J, Solomon MM (eds) Column Generation, Springer, chap 1, pp 1–32

Dohn A, Rasmussen MS, Justesen T, Larsen J (2008) The home care crew scheduling problem. In: Sheibani K (ed) Proceedings of the 1st International Conference on Applied Operational Research, Lecture Notes in Management Science, vol 1, pp 1–8

Dohn A, Kolind E, Clausen J (2009) The manpower allocation problem with time windows and job-teaming constraints: A branch-and-price approach. Computers & Operations Research 36(4):1145–1157

Eaton SC (2003) if you can use them: Flexibility policies, organizational commitment, and perceived performance. Industrial Relations: A Journal of Economy and Society 42(2):145–167

Ernst A, Jiang H, Krishnamoorthy M, Sier D (2004) Staff scheduling and rostering: A review of applications, methods and models. European journal of operational research 153(1):3–27

Eveborn P, Flisberg P, Rönnqvist M (2006) Laps care an operational system for staff planning of home care. European Journal of Operational Research 171(3):962–976

Eveborn P, Rönnqvist M, Einarsdóttir H, Eklund M, Lidén K, Almroth M (2009) Operations research improves quality and efficiency in home care. Interfaces 39(1):18–34

Fischetti M, Polo C, Scantamburlo M (2004) A local branching heuristic for mixed-integer programs with 2-level variables, with an application to a telecommunication network design problem. Networks 44(2):61–72

Fosgerau M, Engelson L (2011) The value of travel time variance. Transportation Research Part B: Methodological 45(1):1–8

Golembiewski R, Proehl Jr C (1978) A survey of the empirical literature on flexible workhours: Character and consequences of a major innovation. Academy of Management Review 3(4):837–853

Günther M, Nissen V (2012) Application of particle swarm optimization to the british telecom workforce scheduling problem. In: Proceedings of the 9th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2012), Son, Norway, pp 242–256

Itabashi G, Chiba M, Takahashi K, Kato Y (2006) A support system for home care service based on multi-agent system. In: Proceedings of the Fifth International Conference on Information, Communications and Signal Processing 2005, pp 1052–1056

Jara-Díaz S (2000) Allocation and valuation of travel time savings. In: Hensher DA, Button KJ (eds) Handbook of transport modelling, vol 1, Emerald Group Publishing, chap 18, pp 303–318

Kallehauge B, Larsen J, Madsen OBG, Solomon MM (2005) Vehicle routing problem with time windows. In: Desaulniers G, Desrosiers J, Solomon MM (eds) Column Generation, Springer, chap 3, pp 67–98

Kergosien Y, Lente C, Billaut JC (2009) Home health care problem, an extended multiple travelling salesman problem. In: Blazewicz J, Drozdowski M, Kendall G, McCollum B (eds) Proceedings of the 4th Multidisciplinary International Scheduling Conference: Theory and Applications (MISTA 2009), Dublin, Ireland, pp 85–92

Landa-Silva D, Wang Y, Donovan P, Kendall G (2011) Hybrid heuristic for multicarrier transportation plans. In: Proceedins of the 9th Metaheuristics International Conference (MIC2011), Udine, Italy, pp 221–229

Li Y, Lim A, Rodrigues B (2005) Manpower allocation with time windows and job-teaming constraints. Naval Research Logistics 52(4):302–311

Lim A, Rodrigues B, Song L (2004) Manpower allocation with time windows. Journal of the Operational Research Society 55(11):1178–1186

Martínez-Sánchez A, Pérez-Pérez M, Luis-Carnicer P, Vela-Jiménez M (2007) Telework, human resource flexibility and firm performance. New Technology, Work and Employment 22(3):208–223

Miller H (1976) Personnel scheduling in public systems: a survey. Socio-economic planning sciences 10(6):241–249

Misir M, Verbeeck K, De Causmaecker P, Vanden Bergue G (2010) Hyper-heuristics with a dynamic heuristic set for the home care scheduling problem. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC), pp 18–23

Misir M, Smet P, Verbeeck K, Vanden Bergue G (2011) Security personnel routing and rostering: a hyper-heuristic approach. In: Proceedings of the 3rd International Conference on Applied Operational Research, ICAOR11, Istanbul, Turkey, pp 193–206

Raff S (1983) Routing and scheduling of vehicles and crews: The state of the art. Computers & Operations Research 10(2):63–211

Rasmussen MS, Justesen T, Dohn A, Larsen J (2012) The home care crew scheduling problem: Preference-based visit clustering and temporal dependencies. European Journal of Operational Research 219(3):598–610

Ropke S, Pisinger D (2006) An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Transportation Science 40(4):455–472

Salani M, Vaca I (2011) Branch and price for the vehicle routing problem with discrete split deliverables and time windows. European Journal of Operational Research 213(3):470–477

Solomon MM (1987) Algorithms for the vehicle routing and scheduling problems with time windows constraints. Operations Research 35(2):254–265