5-2012

Patient and Nurse Considerations in Home Health Routing with Remote Monitoring Devices

Jessica Spicer
University of Arkansas, Fayetteville

Patient and nurse considerations in home health routing with remote monitoring devices

Patient and nurse considerations in home health routing with remote monitoring devices

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Industrial Engineering

By

Jessica Spicer
University of Arkansas
Bachelor of Science in Mathematics, 2010

May 2012
University of Arkansas

Abstract

We build on the current consistent vehicle routing problem literature by formulating a novel multiobjective mathematical model of the home health scheduling and routing problem that includes the option of assigning some patient visits to remote monitoring devices, with the objectives of minimizing total cost, achieving nurse consistency and creating balanced nurse workloads. A heuristic solution approach that approximates the efficient frontier of this multiobjective problem is presented and validated, and the results of using this methodology to solve several realistic instances are included. We also analyze the assignment of patients to devices and present some managerial insights into making these assignments in practice.

This thesis is approved for recommendation to the Graduate Council.

Thesis Director:

Dr. Ashlea Milburn

Thesis Committee:

Dr. Ed Pohl

Dr. Chase Rainwater

Thesis Duplication Release

I hereby authorize the University of Arkansas Libraries to duplicate this thesis when needed for research and/or scholarship.

Agreed ______________________________

Jessica Spicer

Contents

1 Introduction and motivation

Home health is a huge and growing component of the healthcare industry. There were an estimated 33,000 home health providers in the United States in 2009, which served a combined 12 million patients that year (National Association for Home Care and Hospice [2010]). It is expected that by 2040, the number of people 65 or older will quadruple (US Census [2004]), creating additional strain on an already overburdened healthcare system. Additionally, estimates hold that close to one half of US adults suffer from at least one chronic medical condition (CDC [2009]). Home health is an attractive option for these groups of patients, as it can provide care to the elderly and those with recurring conditions at a fraction of the cost of traditional hospital care ($132 a day (NAHC [2007]) versus $1889 a day in the hospital (Agency for Healthcare Research and Quality [2009])). As the population in the United States ages and suffers more chronic disease, there is greater demand for cost effective home health services; as a result, home health jobs are expected grow almost 50% by 2018, which represents greater growth than any other healthcare sector (United States Department of Labor [2010]).

In addition to being a fast-growing component of the nation's healthcare system, the home health sector stands to gain from operations research techniques, particularly in creating daily routes and schedules for nurses. Home health workers drive an estimated 5 billion miles per year to provide care to patients, over twice the total amount driven by all UPS drivers annually (NAHC [2009]). Additionally, many home health agencies do not use scheduling software, but rely instead on various ad hoc methods to assign patients to nurses (Datalytics LLC [2010]), and let nurses create their own routes. The potential for improvement to routes and schedules through the use of decision support technology is enormous, considering the multi-objective and combinatorial nature of the underlying problems. Daily operations in the home health sector require the generation of routes over a specified planning horizon (e.g., number of weeks), where each patient requires a prescribed number of weekly visits for a prescribed number of weeks. While home health agencies are subject to budgetary concerns and thus will be concerned with the total cost and travel time associated with their nurse routes, favorable patient outcomes and nurse satisfaction are top priorities for both for-profit and non-profit agencies. These competing objectives, described in Section 1.1, lead to complex routing problem variants. Technological advances in the home monitoring market, described in Section 1.2, may help home care agencies to simultaneously satisfy these objectives. The goal of this thesis is to evaluate the tradeoffs between the competing objectives when home monitoring technologies are used.

1.1 Competing objectives in home health nurse routing and scheduling

Competing objectives in home health nurse routing and scheduling considered in this thesis include minimizing travel cost, maximizing patient satisfaction, and maximizing nurse satisfaction. The motivation for considering patient and nurse satisfaction when making routing and scheduling decisions is described in this section. Additionally, the proxies used as satisfaction measures in our model are presented. Travel cost is approximated simply as the total travel distance required to perform patient visits.

The number of home health agencies in the United States has almost doubled from 18,000 agencies in 2005 to 33,000 in 2009 (National Association for Home Care and Hospice [2010]). With the increasing number of home health care agencies available to choose from, patient satisfaction and increasing levels of transparency will become even more crucial in attracting and keeping new patients for those agencies that do operate for a profit (Steeg [2008]). Starting in October of 2012, new Medicare reimbursement regulations will take effect that will give bonuses to health care providers that score above average on patient satisfaction surveys, providing yet another incentive for home health agencies to be concerned with patient satisfaction (Rau [2011]). This satisfaction level is often associated with consistency of the care provider, as well as predictable, consistent service times. Thus, patient satisfaction may be measured using nurse consistency (related to continuity of care) and time consistency. Studies indicate that higher levels of continuity of care (that is, treatment administered by the same nurse or small set of nurses) lead to more favorable patient outcomes, high patient satisfaction, and fewer emotional and behavioral issues at discharge (Russell et al. [2011], D'Errico and Lewis [2010]).

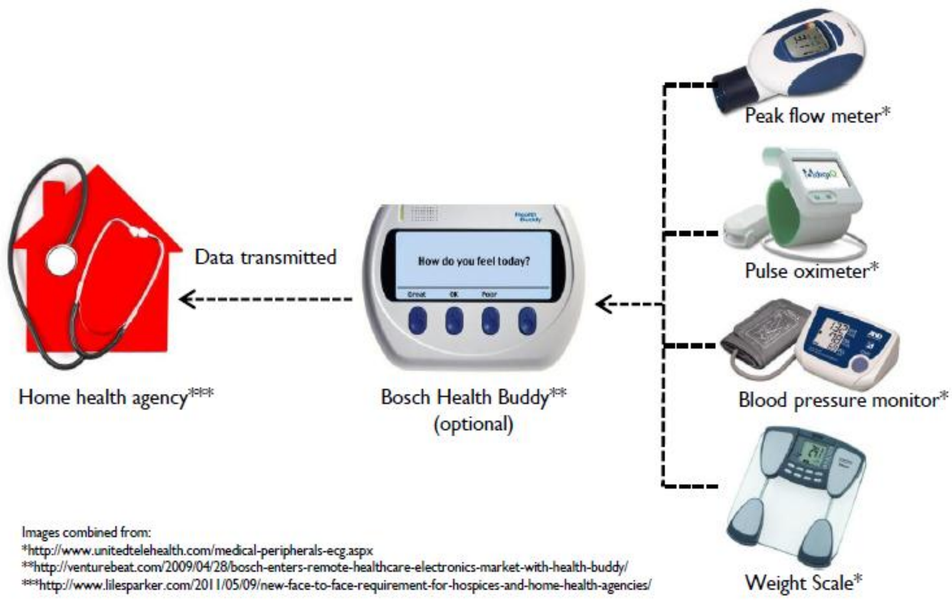
It is predicted that by 2025, there will be an expected national system-wide shortage of 260,000 nurses (Buerhaus et al. [2009]). Therefore, nurse satisfaction will also necessarily become an important goal of home health agencies that wish to attract and maintain adequate nurse staffing levels. Additionally, the literature suggests that nurse satisfaction affects the quality of patient outcomes, in particular, that strenuous workloads negatively affect patient outcomes (Navaie-Waliser et al.

[2004]). Although the literature is replete with nonquantitative factors that affect nurse and home health nurse job satisfaction (Best and Thurston [2004]), there are two measurable objectives that may be considered in this problem setting: balanced nurse workload, and idle time. Excessive workload is mentioned frequently as a dissatisfier of nurses (Navaie-Waliser et al. [2004], McNeese-Smith [1999]), and it is apparent through conversations with home health professionals that nurses are very concerned that the burden of workload be shared equally among all the nurses. We choose to incorporate the balanced workload objective and assess the fairness of workload assignments by evaluating the sum of all pairwise differences over the set of nurses in number of patient visits completed over the planning horizon. The goal of minimizing idle time may be chosen as a counterweight to the time consistency objective, so that nurses do not need to wait idly for a long period of their workday for an appointment time scheduled later in the day.

1.2 Technological advances in home monitoring market

Remote monitoring devices are a relatively new telehealth technology that allow home health nurses to collect and monitor vital sign data without making an in-person visit to collect the data. There are a wide variety of such devices on the market today, including both simple, single-purpose devices such as scales, blood glucose monitors, pulse oximeters, and peak flow meters, and sophisticated systems such as the Bosch Health Buddy and Philips TeleStation, which may be programmed to address the needs of patients suffering from a number of different health conditions (Philips website [2011], Bosch website [2011]). These more holistic systems take and transmit input from single-purpose devices, and also interact with patients through questionnaires and promote education through condition-specific lifestyle recommendations. See Figure 1 for an illustration of the relationship between these two types of devices. The daily measurements and responses from devices are then transmitted back to the agency and/or directly to the patient's nurse for monitoring. Typically, home health agencies own the devices, which they then distribute to patients for the duration of their episode of care. Agencies may differ greatly in the extent to which they employ remote monitoring devices, and the specific configuration of devices used.

Figure 1: Example device illustration



Remote monitoring technology has been used to monitor myriad chronic health conditions, including hypertension, congestive heart failure, diabetes, and coronary artery disease (Bosch website [2011]). Management of these chronic diseases is very sensitive to small changes in various vital sign data. Consistent monitoring of this data can alert care providers early if the patient's condition has suffered, and allow timely interventions that improve quality of life and save money. Studies have consistently found that use of remote monitoring systems results in lower costs and improved outcomes for patients. One such study found that over a six month period, the rate of discharge from home health to more intensive care such as hospital or nursing home was only 15% for those patients issued a remote monitoring system, versus 42% under a traditional home health care paradigm. These improved clinical outcomes came at a cost per visit that was about 70% of the cost of traditional visits (Finkelstein et al. [2006]). Another study conducted by Intel and health insurance company Aetna found that 164 out of 315 heart failure patients avoided potential hospital stays through use of a remote monitoring device (Horowitz [2010]). In a survey administered by Philips, 76.6% of home health agencies that use these remote monitoring devices reported a reduction in unplanned hospitalizations, while 77.2% reported a reduction in emergency room visits for patients issued devices. 49.7% of these agencies saw a reduction in the number of face-to-face visits through use of a remote monitoring system, and 88.6% reported improvement in quality outcomes for patients (Philips national study on the future of technology & telehealth in home care [2008]). Use of remote monitoring device systems is advantageous to both patients and care providers, improving quality of life for patients and increasing both efficiency and effectiveness of home health agencies.

Results of the Philips survey suggest that around 17% of home health agencies currently employ some form of remote monitoring system (Philips national study on the future of technology & telehealth in home care [2008]). Although current Medicare/Medicaid guidelines enforce stringent eligibility requirements for telehealth technology reimbursal which exclude many reasonable applications of the devices, including any store-and-forward technologies like those of most remote monitoring devices (Department of Health and Human Services [2009]), there has been recent movement to relax remote monitoring reimbursement restrictions (Department of Health and Human Services, Centers for Medicare and Medicaid Services [2011]). Perception of cost has discouraged many home health providers from adopting new systems in the past (Philips national study on the future of technology & telehealth in home care [2008]). But remote monitoring solutions are becoming increasingly attractive as providers are becoming aware of cost benefits and government reimbursement is becoming more inclusive, and as a result, the market is predicted to grow substantially in the near future, from a total market value of around \$3 billion in 2009 to \$7.7 billion by the end of 2012 (King [2010]).

We wish to take advantage of remote monitoring technology to simultaneously achieve good performance on the competing objectives of travel cost, patient satisfaction, and nurse satisfaction. Allowing some patient visits to be satisfied by assignment to a remote monitoring device is, abstractly, consistent with the experience of home health agencies that saw a reduction in the number of in-person visits through use of such a system. We assume that device assignments do not need to be assigned a particular start time, do not negatively affect a patient's nurse consistency score, and do not add to any nurse's workload. While a device assignment for a particular visit will clearly reduce travel costs since it eliminates the need for a nurse to travel to the patient's home, it may also improve patient and nurse satisfaction objectives by decreasing the opportunity for variability in nurse assignment, time of visit, or nurse workload assignments.

1.3 Small package shipping application

Although we are most concerned with home health routing applications, consistency of service provider and service times are also important determinants of customer satisfaction in other freight transportation applications. A critical task for small package delivery companies is the construction of low-cost routes for the delivery and pick-up of packages from customers in local service areas. Many of these companies strive to deliver and pickup packages using the same driver at approximately the same time every day. Wong [2008] explains that such consistency may be crucial to customer satisfaction, and that customers grow to depend on consistent deliveries. He gives the example of a video rental store that hires a part-time morning employee to deal with packages received in the morning. If the delivery is late, it may adversely affect the video rental's ability to process the shipments (Wong [2008]). Some customers may have service contracts with the shipper that establish a specific time window during which items should be delivered. Failure to abide by those time windows may lead to decreased customer satisfaction, loss of the contract, as well as opportunity costs associated with the loss of future business or referrals from that customer. UPS, the United Parcel Service, is one company that uses consistency to maintain a competitive advantage. UPS's biggest competitor, FedEx, focuses almost exclusively on efficiency, automation, and cost-reduction. In contrast, UPS emphasizes customer relationships, and as a result, their customers are more likely to know to know their driver by name and retain loyalty to the UPS company (Peppers and Rogers [2004]).

2 Problem description and research questions

We wish to provide a solution approach that finds the efficient frontier using three of the several aforementioned potential objectives: total routing cost, nurse consistency, and balanced workload. The reasons for choosing these particular objectives were twofold. First, it is clear through discussions with home health professionals that these three objectives are routinely prioritized and that agencies attempt to create routes that perform well with respect to all three. This set of objectives has the added benefit of representing three possibly conflicting interests: those of the agency (total cost), patients (nurse consistency), and nurse workforce (balanced workload). The second reason for choosing these three objectives is that it results in a combinatorial problem well-suited to discrete neighborhood moves and a multiobjective heuristic solution approach. We hope to later incorporate the complicating time related objectives like time consistency and minimization of nurse idle time to this base set of three objectives.

We wish to answer the following questions regarding the problem of creating routes for home health nurses:

- What are the trade-offs among objectives associated with creating nurse routes and schedules? That is, can we find or approximate the efficient frontier for problem instances of realistic size? Finding or approximating the efficient frontier allows decision makers from different agencies who may have differing priorities to choose the solution best suited to their agency's situation.

- How should we choose which patients and/or patient visits to assign to remote monitoring devices to create routes that perform well across multiple objectives?

To answer these questions, we define the problem of creating nurse schedules and routes to serve patient demand with the option of assigning some visits to remote monitoring devices. In our multiobjective home health nurse routing and scheduling problem, a home health patient requires a physician-mandated number of weekly visits for a prescribed number of consecutive weeks during a planning horizon. Furthermore, the days on which those visits occur must repeat periodically. Given a set of patients and their required visits, each visit must be assigned to a day, and must also be assigned to either a nurse or remote monitoring device. These assignments are subject to workday length limits for each nurse, daily and horizon-oriented device capacity constraints, and limits on the maximum number of times a specific patient's visits can be performed by a device. This ensures each patient still receives an acceptable number of in-person visits over the planning horizon. Once visit day and nurse or device assignment decisions have been made, daily nurse routes that start and end at each nurse's home must be created, and start times must be assigned for each patient visit. A set of visit start times corresponding to a single nurse for a single day are feasible if, for each pair of consecutive visits $(i, i+1)$, it is possible to start visit i at the assigned time, complete the visit, and travel to visit $i+1$ no later than it is scheduled to begin. While the concept of patient start times does not directly affect any of the three chosen objectives, we will see in Section 4 that they are important in preventing subtours in each nurse's daily route. Patient start times would also be required if time consistency and idle time are incorporated as additional objectives in future work.

Characteristics of a good set of nurse routes and device assignments include low travel cost, a high level of nurse consistency, and balanced nurse schedules across the planning horizon. Nurse consistency is achieved when each patient receives as many of their visits as possible from either the same nurse or by assignment to a remote monitoring device. This patient satisfaction objective is balanced by the nurse satisfaction goal to create routes that give all the nurses approximately the same amount of work over the planning period (for more detail, see Section 4.2). Specifically, we calculate each of these objectives in the following manner:

- **Transportation cost**: total travel cost incurred by all nurses over the planning horizon.
- **Nurse consistency**: the sum over all patients of the number of different nurses that visit each patient.
- **Balanced workload**: the sum of all pairwise differences among nurses of the total number of patient visits completed over the planning horizon.

The problem is to create a set of visit day assignments and nurse or device assignments and daily routes for each nurse that satisfy the set of patient demands over the planning horizon, subject to the aforementioned constraints, while creating good solutions with respect to total distance traveled by all nurses, patient satisfaction, and nurse satisfaction objectives. This problem can be modeled as a multiobjective multidepot vehicle routing problem variant with periodic and consistent requirements. The problem studied in this thesis does not consider the periodic element of complexity, for the reasons described below.

Depending on their prognosis, the specific days that a patient receives visits each week may be subject to a number of scheduling restrictions, including:

- Patients may have a requirement that their visits each week are not scheduled on consecutive days. This may occur if, for example, the patient needs physical therapy that is most effective when the visits are spaced evenly throughout the week.
- Patients may have specific days they must receive care. For example, a patient needing precise doses of IV medications to be administered three times per week may require those doses be

administered at the same time each Monday, Wednesday, and Friday.

In addition to patients requiring specific visit day assignments for medical reasons as described above, patients may also have preferences for which days of the week their visits occur, so visit day assignments are often negotiated with the patient upon admission to home health care. For these reasons, we assume visit day assignment decisions are exogenous, and treat them as input to the planning and scheduling model considered in this thesis. This removes the periodic element of complexity from the resulting routing models, allowing us to focus on providing insights into the consistency requirements and tradeoffs between stated objectives.

3 Literature review

Because the problem studied in this thesis can be modeled as a multiobjective consistent vehicle routing problem variant, we review the relevant literature on consistent vehicle routing problems and multiobjective vehicle routing problem approaches. Home health routing and scheduling problems that have been modeled as consistent vehicle routing problems are discussed in Section 3.1.

3.1 Consistent vehicle routing problem

Development of the consistent vehicle routing problem represents an important and relatively recent extension of the traditional vehicle routing problem. The incorporation of consistency in the routing problems associated with both home health and small package delivery can have favorable effects on company and customer outcomes. Objectives of customer satisfaction, or consistency, may conflict with cost objectives such as total time or distance traveled, but as will be seen in the applications reviewed, explicitly modeling consistency can result in high measures of consistency with relatively little increase in total route cost. This represents a contribution in the vehicle routing problem literature, as previous attempts to incorporate consistency involved using fixed routes that led to inefficiencies and capacity violations. Smilowitz et al. [2009] demonstrate the improvements afforded by these models by showing that a two-phase approach that first minimizes distance and then maximizes consistency does not achieve consistency as well as integrated approaches that simultaneously consider both distance minimization and consistency maximization.

Each variant of the consistent vehicle routing problem includes a set of customers, N, with known demands over a period of days, D. There is a set of drivers, V, for which routes must be generated to service this demand over the planning period. In some cases the drivers begin from their respective homes, as is often the case in home health contexts, and in some cases the drivers begin from a common depot, as in small package shipping contexts. Regardless, the routes are generated for a complete graph on customer and depot locations that satisfies the triangle inequality property. Costs of traveling between two locations i and j, as well as the time to do so are given by c_{ij} and t_{ij}, respectively. Some variants assume the set of drivers V is homogeneous, while others allow for a heterogeneous fleet, with service requests that require various degrees of qualification. Route lengths must be less than some time constraint, and in applications where goods are delivered or received from customers there are also capacity constraints on each vehicle.

Formulations of the consistent vehicle routing problem focus on different measures of consistency, and some seek solutions that perform well for multiple of the following measures. Table 1 shows the metrics used in the various papers reviewed here. Some of the more common measures are:

- **Time consistency**: Routes are constructed in which the customer is serviced at about the same time every day that demand is requested. This may be modeled by defining consecutive time windows of equal length for each day, and then beginning service to the customer within the same time window each day, or in some formulations, during the patient's preferred window. Another method defines time consistency on the basis of the maximum difference in start times for each patient over the planning period.

- **Driver consistency**: Routes are created that minimize the total number of different drivers that service each customer's demands over the entire planning horizon; the lower the number of different drivers, the better the consistency.

- **Customer familiarity**: Preferred routes maximize the number of times a driver visits a customer, based on the idea that there is a customer access cost for each customer that decreases with customer familiarity, or increased frequency of visits. This goal is similar to the driver consistency goal, although it is not identical - the customer familiarity problem solves the drive consistency problem, while the converse does not necessarily hold (Smilowitz et al. [2009]).

- **Region familiarity**: Ideal routes will maximize the total number of times a driver visits a predefined region, or subset of customers, over the planning horizon. This goal is based on the idea that if a driver visits the same region of customers repeatedly then they exhibit a learning behavior that decreases travel time and costs as they gain increased familiarity.

The aforementioned measures of consistency may be incorporated in different ways into the consistent vehicle routing problem model. Some approaches (Groër et al. [2009]) model consistency as a set of constraints, requiring a certain degree of time consistency, as well as perfect driver consistency. Their metric is then the classic minimization of total travel costs for all routes. Others (MacDonald et al. [2009], Smilowitz et al. [2009]) incorporate consistency as a "soft" constraint by placing it in the classic travel cost objective and adding penalties for consistency violations.

Table 1: Measures of consistency

	Time Consistency	Driver Consistency	Customer Familiarity	Region Familiarity
Groër et al. [2009]	✓	✓		
Smilowitz et al. [2009]		✓	✓	✓
Zhong et al. [2007]				✓
MacDonald et al. [2009]	✓	✓		
Francis et al. [2007]	✓	✓		✓
Steeg [2008]	✓	✓		

3.2 Solution methodologies

Groër, Golden and Wasil Groër et al. [2009] model the consistent VRP as a set of additional constraints requiring driver and time consistency, while maintaining the traditional VRP constraints and objective of minimizing total distance traveled. They develop a two-stage heuristic for solving this problem that is based on a modified record-to-record algorithm of Li et al. [2005]. The first stage of the algorithm involves using a modified Clarke and Wright algorithm to generate a template using all the customers that need service on multiple days. In the second stage, for each day, customers that need service only that day are added and those that do not need service are removed. The templates are then subjected to a series of diversification and improvement steps. Once no more improvements have been found for a specified number of iterations, the entire process is repeated up to three times, and the feasible solution with the lowest total routing cost is returned.

The authors perform several computational experiments, first assessing their algorithm's performance by comparing it to small instances that may be solved to optimality. Their heuristic performs well on these small instances containing 10-12 customers, achieving optimality in most cases, and at worst experiencing a 6% gap. The authors note that the computing times for optimally evaluating even these small instances were on the order of several days. They also assess the effect of the consistency constraints on total travel time for large instances by comparing the total travel cost obtained from their heuristic to travel cost obtained from the generic record-to-record (RTR) algorithm presented by Li et al. [2005] without consistency constraints. For the two instance types the authors develop, they find that travel time is no more than 13.5% longer for the consistent version of the vehicle routing problem. However, when the algorithm was applied to a real-world instance containing 3,715 customers with demands over five weeks of five days each, the algorithm performed very well, resulting in a slight 1% increase in total travel cost when compared to the results of running a generic RTR algorithm without incorporation of consistency on the same instance (Groër et al. [2009]).

Smilowitz, Nowak, and Jiang Smilowitz et al. [2009] present a modified Tabu search as a heuristic for their three variants of the consistent vehicle routing problem dealing with driver, customer, and region consistency. They begin by constructing an initial solution using a sweep algorithm. They then allow the set of possible moves to be those that differ in the assignment of customers to drivers by one customer. For a baseline comparison, Smilowitz et al. [2009] test their heuristic using just the traditional cost objective on the set of instances developed by Groër et al. [2009], to demonstrate that their heuristic comes within 4% of the generic RTR when seeking minimum total cost on these instances. The authors use the Tabu search on the three variants, each with an objective of the weighted sum of travel cost and consistency measure. They compare these results to the problem with an only travel cost objective, as well as two problems where routes minimizing travel cost are found in the first step, and then a post-processing step attempts to achieve customer or region consistency by assigning drivers to the routes created in the first step. They then show for a range of instance styles their several models that focus on driver, customer, and region consistency increase total routing costs by at most 5.3% on those instances, and make the observation that the models that focus on the consistency metrics do a better job of finding low-cost routes than the two-step models that focus on low-cost do with finding high-consistency routes(Smilowitz et al. [2009]).

Zhong, Hall, and Dessouky Zhong et al. [2007] develop a two phase approach that seeks both high route efficiency and driver familiarity within a service region. They group customers into "cells" by postal code, and then treat the cells as the new customers in the model. They do this both to decrease the size of the network, and to better model driver learning, as they feel drivers "learn" more by frequent visits to a given neighborhood than to a given customer.

The first stage of their approach is the strategic planning model, in which "core areas", or groups of cells, are created to ensure that a portion of each driver's route is the same over multiple days. They define a "flex zone" as a user-defined percentage of the cells closest to the common depot that are not assigned to core areas. In the initial stage the authors use a Tabu search to assign cells to core areas using a learning function of expected driver performance as the objective. The second stage involves creating driver routes to visit each of the cells based on partial routes among cells in each of the core areas. The route creation method is based on a parallel insertion heuristic already developed by UPS to account for the cell routing and driver learning effects.

They test their first strategic planning stage method on ten randomly generated 500 customer instances, and report that their results are all within 3.5% of the lower bound taken to be the solution of the linearized generalized assignment problem. The authors then demonstrate the value of the strategic planning stage by comparing solutions generated in the second stage of their method to those generated without the use of core areas. After incorporating driver learning into the no-core area method, they find that their method uses on average 4% fewer drivers, routes that are 4% shorter, and a 78% visiting frequency for the highest frequency drivers (28% higher than that of no-core area method). In effect, by explicitly modeling driver learning and its effect on travel and service time, the authors show that more consistent routes may in fact be less expensive routes (Zhong et al. [2007]).

MacDonald, Dörner, Gandibleux MacDonald et al. [2009] formulate a consistent vehicle routing problem in the context of home health care. They require that patient visits be performed during the patient's preferred time window, and explicitly model varying levels of qualification of the “nurses”, where service requests submitted by patients must be served by a qualified or overqualified nurse. They model nurse consistency as a soft constraint, and seek to find routes over the planning period that minimize the weighted sum of the total distance traveled and the maximum number of nurses assigned to any client, where the consistency portion is weighted more heavily than the cost portion.

MacDonald et al. [2009] present a large neighborhood search metaheuristic to solve this problem. They initialize by using a regret insertion heuristic to construct an initial set of routes. The following improvement phase is based on a Simulated Annealing method, wherein algorithms are randomly chosen to delete and then reinsert service requests from the solution. Greedy insertion, random insertion, and regret insertion are used to choose service requests to insert into the solution, and the authors use a consistency deletion operator, in which all requests belonging to the patient that is least consistent are removed from the solution, random deletion operator, which does the same thing for a random patient, and the distance deletion operator, which removes requests with the largest average distance from any neighbor. The authors run this algorithm on the some of the larger instances (up to 150 customers and 544 service requests) presented by Groër et al. [2009], and define their objective such that each time a client is served by an additional server beyond its first, a penalty of 1000 is added. They find that their method produces solutions with less, although still relatively high, driver consistency (since it is in objective function versus constraint) but overall lower cost routes than Groër et al. [2009] (MacDonald et al. [2009]).

Francis, Smilowitz, and Tzur Francis et al. [2007] do not model consistency outright, but instead evaluate the trade-offs between cost reduction via operational flexibility, or the level of constraint, and operational complexity, as defined from the point of view of the service providers and their customers by the time and complexity involved in implementing a given solution. They discuss several methods of incorporating operational flexibility, including the ability to decide how many visits are scheduled above a customer’s minimum visit requirement, the ability to allow a customer to be visited by multiple drivers over the planning horizon, the ability to increase the number of possible scheduling options, and the ability to decide how much is delivered each visit. The authors then evaluate the effect of incorporating operational flexibility on three measures of complexity: arrival span, or difference in the time at which customers are served over the planning horizon, driver coverage, or the percent of the total service region visited by a given driver over the planning period (which is related to Zhong et al. [2007]'s idea of driver learning associated with several visits to the same region), and finally, crew size, or the number of different drivers that serve a given customer over the planning period.

The authors develop a Tabu search algorithm which is used to find solutions that minimize the traditional metric of total routing cost for various levels of flexibility for several instances with 200 customers over a five day period. The solutions are then used to evaluate a second set of metrics corresponding to operational complexity, and the trade-off between flexibility and complexity is analyzed. The authors find that additional operational flexibility leads to increased complexity, with specific effects depending on the geographic distribution of the customers. They do conclude that restricting crew flexibility, related to our discussion of driver consistency, tends to correspond to a very limited increase in cost, and as a general rule driver consistency may be achieved without significant cost increases (Francis et al. [2007]).

Steeg In what he terms the Home Health Care Problem, Steeg [2008] incorporates consistency as a "soft" constraint in the objective by seeking to minimize the sum of number of drivers to visit each customer, overtime costs for nurses, total travel costs, and unassigned tasks, which he allows through use of a dummy driver. Since the factors of the objective are on different scales, Steeg normalizes them by weighting each with a parameter such that the sum of these parameters is one. Steeg [2008] also models qualification level of nurses, and enforces time consistency through hard time windows for each patient visit. He constructs a simple routing heuristic to create an initial feasible solution that is "good" with respect to all components of the objective except nurse-patient loyalty. Steeg then uses an adaptive large neighborhood search mechanism in which patient requests are deleted from and reinserted into the solution using random, shift combination, and worst removal operations in conjunction with in-order, greedy, and regret insertion operations. Solutions with improving objective function values are kept, and operations that led to acceptance of a new solution are given a higher preference value for use in future iterations (hence the adaptive LNS).

Steeg used his algorithm to compute solutions for two instances using real data from two home health agencies in Germany. However, he found the ANLS portion was unable to make much headway in improving consistency, as more weight was placed on the overtime portion of the objective for those instances. The home health agencies from which the instance data was obtained also did not report their nurse consistency, so a comparison was impossible (Steeg [2008]).

3.3 Multiobjective vehicle routing problem

The most popular approach taken to solve multiobjective vehicle routing problems is the scalar method, which includes weighted linear aggregation, goal programming, and the ϵ-constraint method. The first of these involves combining all objectives into a single objective that is a weighted sum of the others, which incurs the difficulty of choosing appropriate weights. A weighted linear aggregation approach will in general find some, but not all of the Pareto-optimal solutions. Goal programming involves choosing some target threshold for each of the objectives, and then minimizing the total distance from the objectives values to their respective goals. This approach also brings the challenge of choosing appropriate goal values for the objectives. In the ϵ-constraint method, one objective is optimized, while the rest are subject to the constraint that each objective i may be no worse than ϵ_i. The ϵ values as well as the objective chosen for optimization is then varied to produce multiple solutions. Any of these scalar methods has the advantage that all traditional optimization techniques and/or heuristics may be employed to solve the modified problem (Jozefowiez et al. [2008], Talbi [2009]).

Another popular approach to multiobjective vehicle routing problems relies on the concept of Pareto dominance. These Pareto methods are often used in evolutionary algorithms and hybrid evolutionary methods, and are based on the idea of assigning fitness scores to solutions that reflect their quality as compared to the overall population on the basis of dominance among the multiple objectives. Other methods include alternating which objective is under current consideration, ranking objectives and solving the problem in rank order (where previously optimized objectives become constraints for subsequent objectives), and various heuristic-specific methods, such as the use of two types of pheromones based on the two objective functions in an ant colony algorithm (Jozefowiez et al. [2008], Talbi [2009]).

3.4 Contribution to existing literature

Incorporating remote monitoring devices and nurse satisfaction objectives (in this case, balanced workload) into the consistent vehicle routing problem is a novel addition particularly well suited to contribute to the home health application area. We are also interested in the Pareto optimal frontier that results from our multiple objectives, while those who formulated this problem in the past (Smilowitz et al. [2009], Steeg [2008], MacDonald et al. [2009]) simply combine the objectives as a weighted sum and do not explicitly examine the trade-offs that result from consideration of various consistency objectives, nor do they consider consistency measures as individual objectives.

4 Problem definition

We first define the notation used to express the home health nurse routing and scheduling problem studied in this thesis.

4.1 Notation

We assume a set of nurses, V, available to serve a set of patients, N, over a set of days in a planning horizon, D, where each patient $n \in N$ and nurse $v \in V$ is associated with a geographical location in the service area representing their home. Because home health care nurses typically begin and end their day in their own home, we allow each nurse v to have its own depot. Let T represent the set of remote monitoring devices. Then the set of servers, S, that may serve patient demand is defined as $S := T \cup V$. The parameter L is the limit on the number of patients each device can serve each day, and the parameter G is the limit on the number of patients the device can serve over the course of the entire planning horizon. We let K represent the limit on the number of visits via device each patient may incur over the planning horizon. We define the home health nurse routing and scheduling problem on a complete network denoted by the underlying graph $G = (N^0, A)$, with node set N^0 representing the customers in set N and nurse depots V, and arc set A connecting nodes in N^0 with nonnegative travel costs $c_{ij},\ (i,j) \in A$, as well as nonnegative travel times, $t_{ij},\ (i,j) \in A$. We define $N\{v\} := N \cup v$ to be the set of patient locations along with nurse v's depot, or the set of nodes that a given nurse v may visit over the planning period.

For a given day $d \in D$, the set of customers to visit and their respective demands are known *a* *priori.* Let r_n^d denote the demand of patient $n \in N$ on day $d \in D$, which in our case is the time (in minutes) needed for a nurse to care for that patient on that day.

We define the continuous variables $s_n^d, n \in N, d \in D$, to represent the time when care for patient n begins on day d, and e_v^d to represent the time at which nurse v returns home on day d. We model a day as beginning at time 0, assume all nurses are available at that time, and that all nurses must return home within Q, an input parameter representing workday length, minutes.

We define the following day-oriented binary variables:

$$x_{ijv}^d = \begin{cases} 1 & \text{if nurse } v \text{ traverses arc } (i,j) \text{ on day } d \\ 0 & \text{otherwise} \end{cases}$$

$$y_{no}^d = \begin{cases} 1 & \text{if patient } n \text{ visited by server } o \in S \text{ on day } d \\ 0 & \text{otherwise} \end{cases}$$

The following horizon-oriented binary variables are used in the nurse consistency objective function.

$$y_{nv} = \begin{cases} 1 & \text{if patient } n \text{ visited by nurse } v \in V \text{ during planning horizon,} \\ 0 & \text{otherwise} \end{cases}$$

The integer valued bookkeeping variables z_v represent the number of patient visits that nurse v completes over the planning horizon; $z_v = \sum\limits_{d \in D} \sum\limits_{n \in N} y_{nv}^d$ for each nurse $v \in V$.

4.2 Objectives

To study the tradeoffs among cost-effectiveness, patient satisfaction, and nurse satisfaction, we define the following five objectives:

- **Transportation Cost**:

  $$\min \ f_1 = \min \sum_{(i,j) \in A} \sum_{v \in V} \sum_{d \in D} c_{ij} x_{ijv}^d$$

  This objective represents the total travel cost of a solution.

- **Nurse Consistency**:

  $$\min \ f_2 = \min \sum_{v \in V} \sum_{n \in N} y_{nv}$$

This objective represents the total number of different nurses seen by all patients.

- **Balanced Workload**:

$$\min\ f_3 = \min \sum_{v \in V} \sum_{u \in V: u > v} |z_v - z_u|$$

This objective represents the sum of pairwise differences in total nurse workloads over the planning horizon. Note that this objective is easily linearized.

4.3 Model constraints

The constraints defining our feasible region are as follows:

$$\sum_{o \in S} r_n^d y_{no}^d \geq r_n^d \ \forall n \in N, d \in D, \tag{1}$$

$$\sum_{j \in N\{v\}} x_{njv}^d \leq r_n^d \ \forall n \in N, v \in V, d \in D, \tag{2}$$

$$\sum_{j \in N\{v\}} x_{njv}^d = y_{nv}^d \ \forall n \in N, v \in V, d \in D, \tag{3}$$

$$\sum_{j \in N\{v\}} x_{vjv}^d \leq 1 \ \forall d \in D, v \in V, \tag{4}$$

$$\sum_{j \in N\{v\}} x_{ijv}^d = \sum_{j \in N\{v\}} x_{jiv}^d \ \forall i \in N\{v\}, v \in V, d \in D, \tag{5}$$

$$y_{nv}^d \leq y_{nv} \ \forall n \in N, d \in D, v \in V, \tag{6}$$

$$y_{nv} \leq \sum_{d \in D} y_{nv}^d \ \forall n \in N, v \in V \tag{7}$$

$$\sum_{v \in V} t_{vj} x_{vjv}^d \leq s_j^d \ \forall j \in N, d \in D, \tag{8}$$

$$s_i^d + r_i^d + t_{ij} + M \sum_{v \in V} x_{ijv}^d \leq s_j^d + M \ \forall i \in N, j \in N, d \in D, \tag{9}$$

$$s_i^d + r_i^d + t_{in_v} + M x_{ivv}^d \leq e_v^d + M \ \forall v \in V, i \in N, d \in D, \tag{10}$$

$$e_v^d \leq Q \ \forall v \in V, d \in D, \tag{11}$$

$$z_v = \sum_{d \in D} \sum_{n \in N} y_{nv}^d \ \forall v \in V, \tag{12}$$

$$\sum_{n \in N} y_{nw}^d \leq L \ \forall o \in T,\ d \in D, \tag{13}$$

$$\sum_{o \in T} \sum_{d \in D} y_{nw}^d \leq K \ \forall n \in N, \tag{14}$$

$$\sum_{d \in D} \sum_{n \in N} y_{no}^d \leq G \ \forall o \in T, \tag{15}$$

$$x_{ijv}^d \in \{0,1\} \ \forall i,j \in N\{v\}, v \in V, d \in D, \tag{16}$$

$$y_{no}^d \in \{0,1\} \ \forall n \in N, o \in S, d \in D, \tag{17}$$

$$y_{nv} \in \{0,1\} \ \forall n \in N, v \in V, d \in D, \tag{18}$$

$$z_v \in \mathbb{Z} \ \forall v \in V, \tag{19}$$

$$s_i^d \geq 0 \ \forall i \in N, d \in D, \tag{20}$$

$$e_v^d \in [0, Q] \ \forall v \in V, d \in D, \tag{21}$$

Constraints (1) ensure that customer i is serviced on any day service is requested, either by a nurse or a device. Constraints (2) ensure that a patient is not visited on a given day unless its demand for that day is nonzero. Constraints (3) connect the x and y variables. Constraints (4) ensure that there is exactly one nurse per depot and that nurses do not visit depots which are not their own. Constraints (5) ensure flow conservation through all nodes. Constraints (6) and (7) relate the horizon-oriented nurse assignment variables, y_{nv}, to the day-oriented variables y_{nv}^d. If a patient $n \in N$ is visited by nurse $v \in V$ at least once during the time period, then y_{nv} is set to 1, otherwise, it is set to 0 if nurse v does not visit patient n. Note that there is no need for horizon-oriented device assignment variables, since a device assignment does not contribute negatively to the nurse consistency objective.

Constraints (8) help calculate the time s_i^d, when care for the first patient of the day for each nurse begins. Constraints (9) ensure that the time, s_j^d, when care for patient j (other than the first patient visited for the day) may begin is based on the time, s_i^d, that patient i's care began, the time required to care for this patient, r_i^d, and the time required to travel between the two patients, t_{ij}. Constraints (10) calculate the time at which nurse v can end their day based on the last patient they saw that day and (11) ensure that the length of nurse v's work day is less than Q. These constraints, taken together, prevent subtours by ensuring that the start times of patient care are increasing for each successive patient of a nurse's route. Constraints (12) link the bookkeeping variables $z_v, v \in V$ to the horizon-oriented nurse assignment variables. Although the $z_v, v \in V$ variables are not technically necessary, conceptually they represent the total number of patient visits assigned to nurse v over the planning horizon, and make the calculation of the balanced workload objective more intuitive.

Constraints (13) enforce a daily capacity on each remote monitoring device, while (14) restricts the number of each patient's visits that may be served by a device visit. Constraints (15) enforce a limit on the number of patient visits that may be assigned to a device over the planning horizon. Assignments to the device incur no routing cost. The aforementioned objectives remain the same, and thus patient assignment to a device may only improve objective values for all five objectives. This seems reasonable, as we take as a given that devices will be used, and are most interested in how their use affects the cost, patient satisfaction, and nurse satisfaction objectives.

Finally, constraints (16) and (17) define the day-oriented binary variables for assignment and routing, constraints (18) define the associated horizon-oriented binary variables, constraints (19) define the integer valued bookkeeping variables, and (20) and (21) define the continuous variables for when patient care begins and a nurse's day ends respectively.

5 Methodology

We first attempted to solve the five replications of an instance of realistic size optimally with respect to each of the three objectives (this instance is detailed later in Section 6). We used Cplex via Ampl with a five hour runtime limit for each of these five replications for each of the cost, nurse consistency, and balanced workload objectives. In all fifteen combinations of replication number and objective, Cplex failed to find a feasible solution, and we observed anecdotally that the linear relaxation of our problem took around thirty seconds to run to optimality. This is not surprising, since the generic vehicle routing problem, an *NP*-hard problem, is a special case of our problem, and the size of the instance style is relatively large. Since generating the efficient frontier would involve optimizing multiple times over the feasible set, we chose to pursue a heuristic approach to approximate the efficient frontier.

5.1 MOAMP

We here detail a solution approach for approximating the efficient frontier using our three objectives: total routing cost, nurse consistency, and balanced workload. We model our multiobjective heuristic solution approach after that of Caballero et al. [2007] (hereafter referred to as MOAMP, as it is known in the literature). The authors describe an algorithm based on two phases of Tabu searches. The first phase consists of a linked series of Tabu searches, where each of n single objectives is minimized in turn, and the first objective is then minimized a second time. At each iteration of each of these $n + 1$ Tabu searches, the current solution is checked against the set of efficient points, which is updated as new nondominated points are found. This approach is based on the premise that, in general, nondominated solutions may be found within a neighborhood or reasonable neighborhood search of one another. The second phase then involves a series of Tabu searches which seek to find compromise points that perform reasonably well with respect to all objectives. At each Tabu search, we randomly generate normalized weights for the objectives, and the weighted sum of the objectives normalized over their range in the current nondominated set is minimized, while new nondominated solutions are added to the nondominated list as before. The number of Tabu searches carried out in the second phase is determined by an input parameter that specifies the number of searches that may be undertaken without any change in the efficient set.

Figure 2 shows an abstracted depiction of this two-phase process. The first phase is illustrated with solid arrows, as the heuristic attempts to optimize the three objectives individually, and then returns to optimize the first objective again, completing the cycle. The second phase is shown with dotted arrows, where a series of linked Tabu searches are carried out to identify compromise solutions. Figure 3 shows the nondominated points collected during this two-phase method. As the search moves from one point (representing the best found solution with respect to the objective currently being optimized) to another, each solution along the way is checked against the nondominated set. Figure 3 shows that not only are the best found solutions for individual and weighted sums of objectives added to the approximation set, but also many of the solutions found along the way to optimizing the next objective. This entire two-phase process may be carried out multiple times, although the authors only complete one cycle in their implementation.

Figure 2: Linked Tabu searches

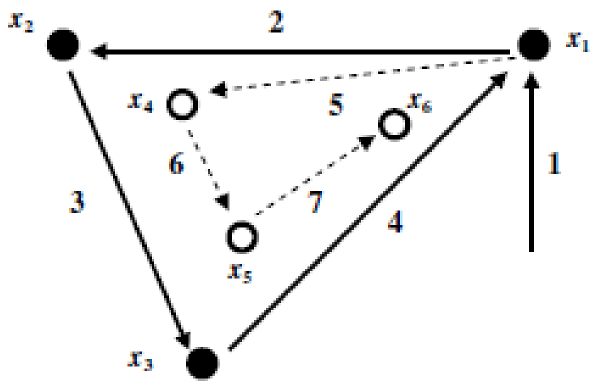
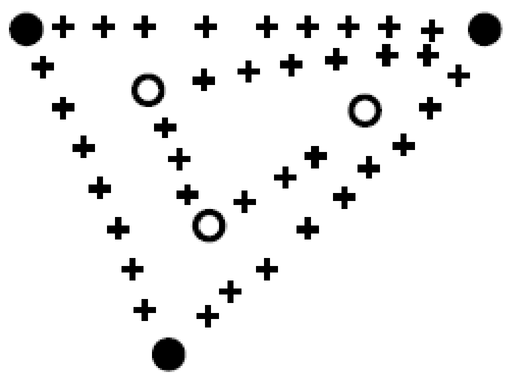


Figure 3: Adding nondominated points



5.2 MOAMP advantages

We chose to use this particular metaheuristic approach for several reasons. First, it has the advantage of being easy to follow and having a clear implementation strategy. It is also general enough to be used with multiple objectives, and requires only the design of a Tabu search capable of making the appropriate neighborhood moves. MOAMP is general enough to be applied to any problem that may be reasonably solved in the single objective using a Tabu search - it does not require any special problem structure (as in Pacheco and Marti [2006], where one integer-valued objective may be used as an input to the solution procedure). In our case, this means that after implementing the initial Tabu strategy for the three chosen objectives, the heuristic may be expanded at a later time to include additional time-related objectives. Additionally, Caballero et al. [2007] and García et al. [2011] achieved impressive results that either exceeded the results of previous methods or achieved results of similar quality in less computational time on a wide variety of problem types. They demonstrate that this solution procedure performs well on a variety of multiobjective combinatorial optimization problems, including biobjective knapsack, assignment, set packing, location routing, and problems involving the optimizing of both routes and inventory levels.

The MOAMP solution approach is also attractive because it is based on a Tabu search heuristic framework. There is more room to contribute to the multiobjective combinatorial optimization problem heuristic literature in this area, as opposed to an evolutionary approach which has been well-studied in the multiobjective setting (Jozefowiez et al. [2008]). For the consistent vehicle routing problems detailed in Section 3.1, variable neighborhood searches have also been successfully used in Groër et al. [2009], Smilowitz et al. [2009], Steeg [2008] and MacDonald et al. [2009] as a means of exploring the single (or weighted) objective versions of the problem. It is therefore natural to extend this methodology to the multiobjective consistent vehicle routing problem. As compared to other multiobjective Tabu search procedures, MOAMP has the potential to be more computationally efficient, as it requires moving only a single solution through improving neighborhoods of the feasible set to find nondominated solutions, in contrast with other approaches which involve the movement of a set of several solutions through the feasible space (Hansen [1997]).

5.3 Neighborhood moves

Our neighborhood moves involve the movement of patients among both nurse and device routes. A given solution to our home health scheduling and routing problem has a set of nurse and device routes for each day of the planning horizon. Each day, each nurse's route contains an ordered list of patient locations that are to be visited on the given day. Device routes are not physical routes that incur any travel cost, but instead represent the set of patients assigned to a given device on a given day. In our instances, devices may only replace one visit for up to one patient over the course of the planning horizon, as a result, most of the device routes for a given day are empty.

We establish two move types often used in vehicle routing contexts: remove-and-reinserts and swaps. Remove-and-reinserts involve removing a patient from either a nurse or device route and placing it in a new place in the set of nurse/device routes. We allow patients to be reinserted in either their original, or a new, route as long as the patient is not placed in the exact position from which it was removed. Swaps occur when two chosen patients are removed from their respective places in the route set and replaced in the original place of the other patient. These swap moves defined in our setting require that the the two patients are from different nurse/device routes, since the rearrangement of patients within a route may be achieved using the remove-and-reinsert operator.

5.4 Our implementation

Instead of examining all possible remove-and-reinsert and swap moves at each iteration of the algorithm, we attempt to take advantage of the underlying problem structure to develop move strategies well-suited to improve each of our three objectives: cost, nurse consistency, and balanced workload. In Phase I, where we search for the optimal of each objective in turn, the algorithm exclusively uses the objective move strategy of the current objective to be optimized. In Phase II, where we seek compromise solutions, the algorithm chooses randomly among the three strategies with an equal probability of selection.

5.4.1 Cost move strategy

In the cost move strategy, a list of all the arcs used in all days of the planning horizon for the set of current routes is maintained, along with the associated arc cost. At each iteration, this list is sorted based on increasing arc cost. One arc is chosen at random in the most costly α percent of arcs. The patient from whom this arc emanates is selected, as is one of either the remove-and-reinsert or swap moves (the move is selected randomly with equal probability). All possible remove-and-reinserts or swaps involving this patient are evaluated, and the feasible move resulting in the minimal objective value is taken. The arc list is then updated by removing arcs present only in the previous route set and adding new arcs resulting from the recent remove-and-reinsert or swap, and the process is repeated.

More formally, suppose arc (i, j) traversed on day d is chosen from the sorted arc list. Then, patient i is removed from between locations l and k in its route on day d and the algorithm randomly chooses between a remove-and-reinsert or a swap move. If remove-and-reinsert is chosen, then all possible reinsertions of patient i in all feasible places in all routes are evaluated, and the one resulting in the best objective value is selected and implemented. If patient i is reinserted in a nurse route between locations g and h, then arcs (l, i), (i, k), and (g, h) are removed, while arcs (l, k), (g, i), and (i, h) are added to the arc list. There are no arcs associated with device routes, so if i is reinserted in a device route, then only arcs (l, i) and (i, k) are removed, and only arc (l, k) is added to the arc list. If a swap move is chosen, then all possible swaps involving patient i are evaluated, and the best is taken. Assuming patient i is swapped with patient p, the arc list updates are those that correspond to patient i being removed and reinserted in patient p's place, and patient p being removed and reinserted in patient i's place. The arc list must be updated in this way at each move of the Tabu search, regardless of the move strategy used.

5.4.2 Nurse consistency move strategy

In the move strategy motivated by the nurse consistency objective, a list of patients is maintained throughout the algorithm. Each patient is associated with their individual nurse consistency score (recall that this score must be at least 1 and is bounded above by either the number of visits that particular patient requires over the course of the planning horizon, or the total number of nurses, whichever is smaller). As in the cost strategy, this list is sorted based on increasing (less desirable) nurse consistency scores, and one patient, say patient j, is chosen at random within the top β percent of this list. The nurse route and day to be involved in the move are chosen based on their incremental effect on total nurse consistency score, that is, the nurse will be chosen that visits the given patient the fewest number of times over the planning horizon. For example, if nurse i visits patient j only one time over the planning horizon, moving patient j to another nurse that already visits the patient on a different day will improve the total nurse consistency objective. Once the patient, nurse, and day have been chosen, a remove-and-reinsert or swap move is chosen randomly, with total nurse consistency as the determining objective. If remove-and-reinsert is chosen, then only patient j will require an update in the patient list. Patient j's nurse consistency score in the patient list is decremented if the new solution requires one less nurse to visit patient j over the planning horizon, and incremented if one new nurse is added to the set of nurses that visit patient j. If a swap move takes place, the update process occurs both with respect to patient j, and with respect to the other swapped patient, patient p. Visits from a device do not contribute to an individual patient's nurse consistency score. As with the arc list, the patient list must be updated after every iteration of the Tabu search.

5.4.3 Balanced workload move strategy

In the balanced workload move strategy, a list of all nurses is maintained, along with their respective number of assigned appointments over the entire course of the planning period. This list is sorted at each iteration, and the nurse with the most assigned appointments is chosen, say nurse k. For that nurse, a day in the planning horizon is chosen randomly; the probability of choosing each day is proportional to the number of patients the nurse visits that day. Once the day is selected, a patient is chosen randomly from the nurse's route that day and is removed from the route and replaced in a new nurse's route so to minimize the balanced workload objective (note that a swap would not change the balanced workload objective value). Nurse k's balanced workload score, or number of assigned appointments, is decremented each time a remove-and-reinsert is made. If the patient is reinserted in another nurse route then that nurse's balanced workload score is incremented. Swaps do not affect any nurse's balanced workload score.

5.4.4 Device moves

At predetermined intervals in the Tabu search, a device is randomly chosen and the patient assigned to it is removed from the device route and reinserted in a nurse route. This is done to introduce diversity in the search. While swapping a device patient with a patient in a nurse route may improve the overall objective value, a neighborhood move that merely removes a patient from a device and reinserts the patient in a nurse route would very rarely be improving (only possibly in the case of the balanced workload objective), and is very rarely selected in cost and nurse consistency move strategies. Forcing the removal of a patient from a device route creates the opportunity for other patients to be assigned to the newly open device, or even for the same patient to be assigned to the device on a different day.

5.4.5 Tabu definition

Once a remove-and-reinsert or swap neighborhood move has been made, the reverse move may not be implemented for a defined number of Tabu search iterations. Researchers have used various definitions of reverse moves in the context of vehicle routing problems, but we adopt a commonly used definition (Cordeau et al. [2001], Smilowitz et al. [2009], Pacheco and Marti [2006]) that associates reverse moves with moves that place the patient anywhere in its original route. At the conclusion of a remove-and-reinsert, the recently moved patient, along with its previous care provider (be it nurse or device), and the day of the planning horizon this move took place, are added to the tabu list. The patient then may not be moved back to its previous care provider for care on that same day of the planning horizon (although it may be moved to that care provider's route a different day of the planning period) while it remains on the tabu list. At the conclusion of a swap move, two elements are added to the tabu list, one preventing each patient from being returned to their previous route on the given day.

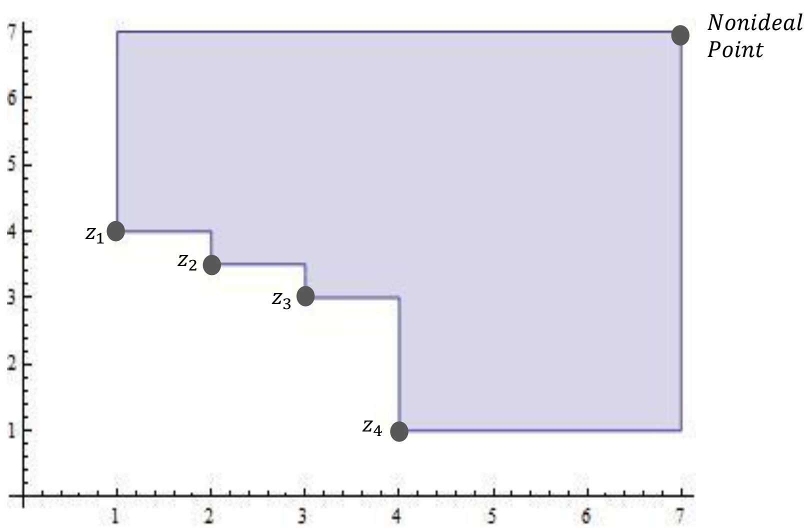
5.5 Assessing solution quality

In order to assess solution quality, and to choose an appropriate set of heuristic parameters for use in running all twelve of our instance styles, we looked at both the hypervolume metric and lower bounds on the best found solutions for some of the individual objectives.

5.5.1 Hypervolume metric

Hypervolume is a common metric used to assess the quality of an approximated efficient set found by a multiobjective optimization heuristic scheme. It requires the user to choose a nonideal point in objective space that has the property that each component is greater than the maximum individual objective value that may be achieved within the feasible set of solutions (for a minimization problem). For a problem with n objectives, the hypervolume metric measures the n-dimensional volume of the set of solutions dominated by an approximated efficient set in objective space and bounded by this nonideal point. An illustration of this for a minimization problem with two objectives is given in Figure 4, where the area of the shaded region gives the hypervolume measurement of the set of solutions dominated by the approximation set $\{z_1, z_2, z_3, z_4\}$. Since the hypervolume metric measures the size of objective space dominated by the efficient set approximation, the approximation with the greater hypervolume measurement is taken to be the best approximation when multiple approximations are compared. Hypervolume has several nice properties that make it a natural choice for use in comparing the quality of two different efficient set approximations.

Figure 4: Hypervolume example



Advantages of using hypervolume metric The most useful property of the hypervolume metric is that it is the only known unary quality measure (that is, its value may be calculated for any approximation set independent of any other set) that can indicate that one approximation set is not worse than another (Zitzler et al. [2003]). Put another way, the hypervolume metric has the property that if set A dominates set B, then the hypervolume metric associated with set A will be greater than that associated with set B. Set A is said to dominate set B if every solution in B is either in A or is dominated by a solution in A.

This is particularly interesting in the context of our MOAMP heuristic solution approach to approximate the efficient frontier. Since the current approximation set will dominate any previous approximation set at any point in the heuristic run, the hypervolume metric will monotonically increase as the algorithm proceeds. Because the hypervolume metric assigns a scalar value to each approximation set, we may use the sequence of increasing hypervolume measurements to assess the relative improvement to the approximation set over various iterations of our MOAMP implementation. Put simply, it would be reasonable to allow the heuristic implementation to run until the hypervolume metric seems to converge; halting the search process while each iteration still produces marked improvements in that metric would be premature. This is especially true in light of the second useful property of hypervolume: the metric will be maximized if and only if the approximation set is (exactly) the Pareto optimal set (Fleischer [2003]). This has motivated some researchers to use maximization of hypervolume as a guiding strategy in heuristic solution approaches (Emmerich et al. [2005], Bradstreet et al. [2006]).

Calculating hypervolume Unfortunately, there does not exist an algorithm that is capable of calculating hypervolume efficiently. While some methods have been developed that run relatively quickly in practice, all such methods are exponential in either number of solutions in the input approximation set, or number of objectives in the worst case (While et al. [2006]). We used one such method known as HSO, or the Hypervolume by Slicing Objectives algorithm, to find the hypervolume associated with our various approximation sets. We then used this metric to assess the effectiveness of various choices of parameters for our MOAMP implementation (see Section 7).

HSO works by iterating through each objective, and making slices in the set dominated by the approximation set with respect to the current objective: each slice then represents a section of hypervolume in the dimension of the remaining objectives. These slices are then again sliced with respect to the next objective, in an iterative fashion. This process partitions the points in the approximation set based on slices, and retains the weight associated with each slice as cuts are made in more objectives. The result is a list of one dimensional slices, each with a weight that represents the cumulative multiple of volume in the previously processed objectives. The total sum of all one dimensional objective values multiplied by their weights gives the final hypervolume measurement. The algorithm we implemented for this process, developed and detailed in While et al. [2006] is given in Figure 5 (graphic from While et al. [2006]).

Figure 5: HSO pseudocode

```
hso (ps):
  pl = sort ps worsening in Objective 1
  s = {(1, pl)}
  for k = 1 to n-1
    s' = {}
    for each (x, ql) in s
      for each (x', ql') in slice (ql, k)
        add (x * x', ql') into s'
    s = s'
  vol = 0
  for each (x, ql) in s
    vol = vol + x * |head (ql)[n] - refPoint[n]|
  return vol

slice (pl, k):
  p = head (pl)
  pl = tail (pl)
  ql = []
  s = {}
  while pl /= []
    ql = insert (p, k+1, ql)
    p' = head (pl)
    add (|p[k] - p'[k]|, ql) into s
    p = p'
    pl = tail (pl)
  ql = insert (p, k+1, ql)
  add (|p[k] - refPoint[k]|, ql) into s
  return s

insert (p, k, pl):
  ql = []
  while pl /= [] && head (pl)[k] beats p[k]
    append head (pl) to ql
    pl = tail (pl)
  append p to ql
  while pl /= []
    if not (dominates (p, head (pl), k))
      append head (pl) to ql
    pl = tail (pl)
  return ql

dominates (p, q, k) returns True iff the point p
dominates the point q in Objectives k..n
```

5.5.2 Comparison to lower bounds

A second way we assess the quality of our approximation sets is to compare the best solutions for each individual objective to lower bounds on those objectives. While this measure does not provide much information about the quality of the approximation set as a whole, it may indicate the quality of some members of that set, as the true Pareto frontier would certainly include solutions yielding the optimal values for each individual objective.

Due to the nature of our home health routing and scheduling problem, we can easily generate intuitive lower bounds for both nurse consistency and balanced workload objectives. Recall that nurse consistency is the sum over all patients of the total number of different nurses that visit the given patient over the planning horizon. Since we require that each patient receive at least one in-person during the planning period, we may conclude that the best possible routing scenario with respect to nurse consistency would correspond to each patient being visited by exactly one unique nurse over the planning horizon. Due to constraints on workday length, and the specific nature of the patient demand schedule, it may not be possible to achieve this in reality, but the nurse consistency objective value will be bounded below by this ideal case. In other words, nurse consistency will always be greater than or equal to the number of patients in the problem.

The balanced workload objective is calculated by summing the pairwise differences among all nurses in total number of patient appointments served over the course of the planning horizon. The ideal scenario, in terms of the balanced workload objective, is that all nurses perform exactly the same total number of visits over the planning period. In this case, all pairwise differences in total number of visits for any two nurses will be zero, as will the total balanced workload objective. Therefore, the optimal value of the balanced workload objective will be bounded below by zero for all instance types.

6 Realistic instance development

We developed twelve instance styles that use one each of two geographic region sizes, three patient location distributions, and two patient visit distributions. Each instance style has five replications using the chosen parameters. These instances were developed to be as realistic as possible with regard to all possible parameters. Information from National Association for Home Care and Hospice [2010] was used to determine nurse staffing levels and daily average nurse productivity, while the Medicare.gov [2010] database provided information for all agencies regarding their service areas. We also used data available from U.S. Census Bureau [2012] to find the area in square miles of various agency service areas, and the IRS [2011] for information regarding standard mileage reimbursement rates.

6.1 Common parameters

Each instance style has several parameters in common. We assume a ten day planning horizon for all styles. Although home health patients often receive care for at least 60 days, the schedule of each individual patient's needed visits typically repeats itself. Since we are studying the deterministic home health scheduling and routing problem, where all patient demand is known in advance, we can assume that the complete schedule of all needed patient visits also has this property. Two weeks is therefore a reasonable amount of time over which to create a set of nurse routes to be repeated for both cost and nurse consistency objectives, as well as the set of patient visit assignments to create balanced schedules for nurses. We assume that the agency has ten devices which may each be assigned to exactly one patient over the planning horizon, and which may replace exactly one in-person visit for that patient over the ten day period. Each nurse is given a full workday's length of nine hours in which to complete all patient visits for the day and return home. All instance styles are constructed assuming that the agency employs nine full-time nurses (National Association for Home Care and Hospice [2010]).

Table 2: Instance parameters

Parameter	Description	Value
D	number of days in planning horizon	10
T	number of remote monitoring devices	10
K	planning period device visit limit per patient	1
Q	workday length limit (minutes)	540
V	number of nurses	9

6.2 Instance style dimensions

We define the following geographic region sizes, patient location distributions, and patient visit distribution types.

Geographic region sizes We used two geographic regions sizes, which we denote as large and small. Assuming that the service area of home health agencies may be represented with a square, we used data available from the Home Health Compare database (Medicare.gov [2010]) to find the list of zip codes serviced by each home health agency and cross-referenced this with the zip code tabulation area data available from the U.S. Census Bureau [2012]. We used this data to estimate the size of service area in total square miles for each agency and took the 25th and 75th percentile service area sizes in the list of all calculated agency service areas for our small and large geographic regions, respectively. These correspond to a small geographic region that is 17 miles square, and a large region that is 37 miles square.

Patient location distributions Within each of these two sizes of geographic service areas, we chose to generate instances with three types of patient location distribution styles: uniformly distributed (Figure 6), clustered (Figure 7), and clustered with uniformly distributed (Figure 8). In the uniformly distributed style, all patient and nurse home locations are generated randomly throughout the square service area from a continuous uniform distribution over both coordinates. In the clustered style, three to five seed locations were chosen (depending on the size of the service area), and patient and nurse home locations were generated randomly within a given radius of those seed locations, resulting in three to five "clusters" of locations. Finally, the clustered with uniformly distributed style generated half of patient and nurse locations within clusters, and the other half uniformly randomly throughout the service area. Patient and nurse locations are not a direct input to our model; instead, these locations were used to create both travel cost and travel time matrices, assuming Euclidean distance between any two points, costs equal to the IRS mileage reimbursement rate (IRS [2011]) for each trip, and travel times that assume travel will occur at 35 miles per hour on average.

Figure 6: Example uniform patient location distribution for small geographic region

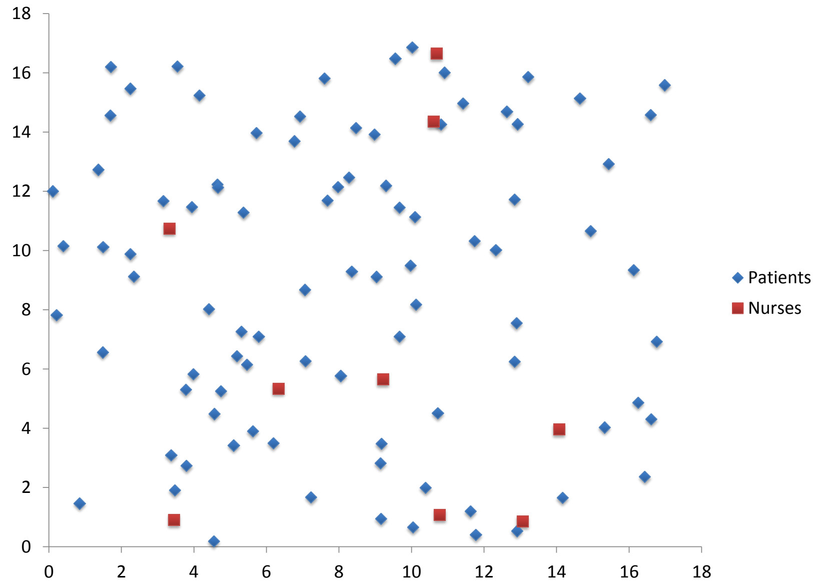


Figure 7: Example clustered patient location distribution for small geographic region

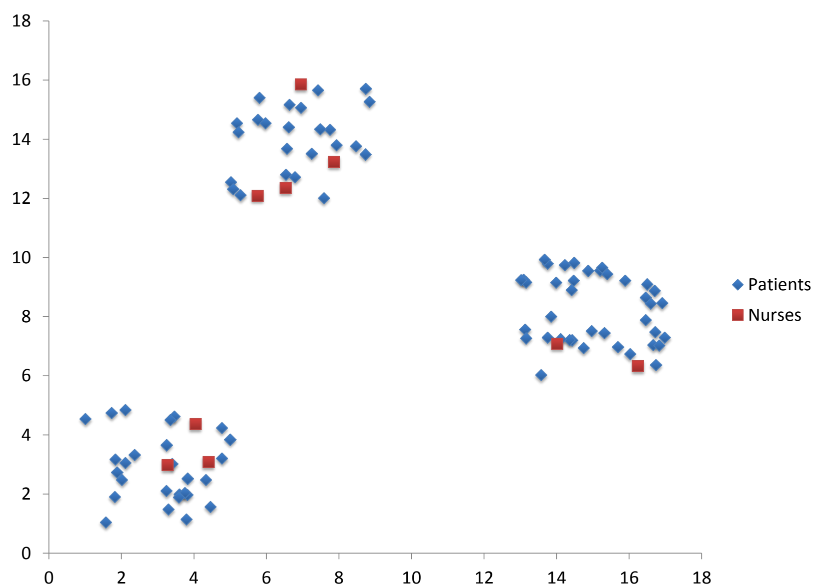
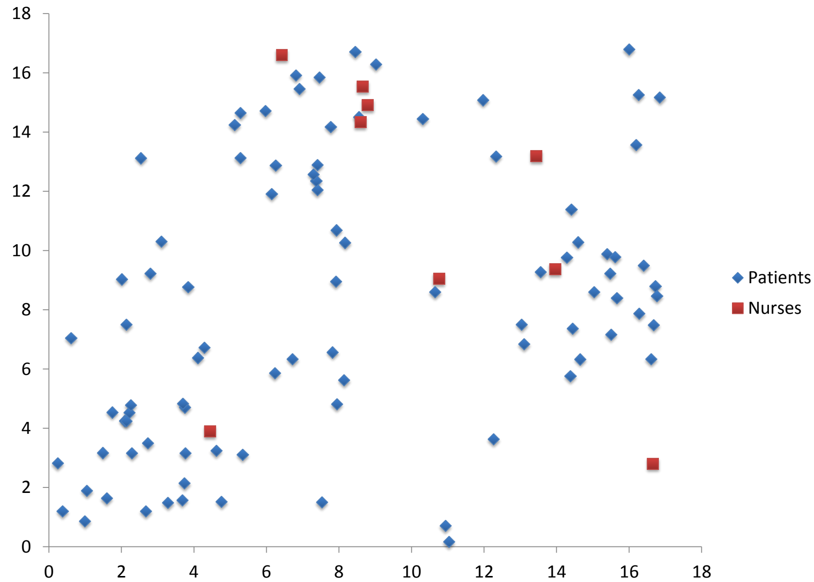


Figure 8: Example clustered with uniform patient location distribution for small geographic region



Patient visit distributions To create our two different patient demand schedules, we developed two sets of patient visit distributions. We assume that over the course of the ten day planning horizon that there will be patients requiring anywhere from two to eight visits. The first style is based on the assumption that most patients will require two to three visits per week, or around five visits over the planning horizon, while the distribution of the second style corresponds to roughly equal numbers of patients requiring each of $2, 3, \ldots, 8$ visits over the ten days. More formally, let p_i be the proportion of the N total patients that require exactly i visits over the course of the planning horizon, where $i = 0, 1, 2, \ldots, 10$. The two patient visit distributions used for our instances are given in Table 3, along with the associated number of patients that correspond to each distribution.

Table 3: Patient visit distributions

	p_2	p_3	p_4	p_5	p_6	p_7	p_8	N
Distribution 1	0.05	0.125	0.125	0.4	0.125	0.125	0.05	90
Distribution 2	0.14	0.14	0.14	0.16	0.14	0.14	0.14	92

6.3 Schedule generation

In addition to cost and travel time matrices, the third model input is a complete schedule of patient demand over the planning horizon. All patient visits are assumed to take 45 minutes to complete. To generate a schedule of required patient visits from the patient visit distributions, we develop a simple algorithm capable of making such assignments. Data suggests that there should be enough patients that require visits each day that each nurse may be assigned 5 patient visits per day, on average (National Association for Home Care and Hospice [2010]). If each home health agency employs 9 full-time nurses on average, this corresponds to approximately 450 total needed patient visits over a ten day period. The algorithm first uses the given patient visit distributions to calculate N, the total number of patients in the model (note that some rounding may occur if the patient visit distributions do not divide into the total number of patient visits). Note that different patient visit distributions will correspond to different total numbers of patients in the instance. For example, if all patients in the instance style require 8 visits over the planning horizon, there will be fewer total patients needed to create a schedule with an average of 5 patient visits per nurse per day than if all patients in the instance style require 2 visits over the planning horizon. Starting with the patients that require the most visits and ending with those that require the fewest visits over the planning horizon, the algorithm then randomly chooses which days the patient will require visits (recall the number of days the patient will be visited is determined by the patient visit distribution). Each day has a maximum number of patient visits that may be assigned to it, based on the average number of needed patient visits per day (e.g., in our instances, each day should have 45 patient visits per day, so this maximum was allowed to be 48 for each day). The result is a schedule with approximately 45 patient visits assigned to each day, and relative proportions of patients that match the given patient visit distribution requirements.

Table 4: Instance styles

Code	Patient Location Distribution	Geographic Region Size	Patient Visit Distribution
$US1$	Uniform	Small	Distribution 1
$US2$	Uniform	Small	Distribution 2
$UL1$	Uniform	Large	Distribution 1
$UL2$	Uniform	Large	Distribution 2
$CS1$	Clustered	Small	Distribution 1
$CS2$	Clustered	Small	Distribution 2
$CL1$	Clustered	Large	Distribution 1
$CL2$	Clustered	Large	Distribution 2
$UCS1$	Uniform and Clustered	Small	Distribution 1
$UCS2$	Uniform and Clustered	Small	Distribution 2
$UCL1$	Uniform and Clustered	Large	Distribution 1
$UCL2$	Uniform and Clustered	Large	Distribution 2

7 Heuristic tuning and validation

The three heuristic parameters we are interested in optimizing are Tabu tenure, or number of iterations that each reverse move is prohibited, the Tabu stopping condition, or the maximum number of iterations that are allowed to occur without improvement to the best known objective value of the particular Tabu search, and N, the maximum number of Tabu searches allowed to occur in Phase II of the algorithm without any change to the nondominated set. To promote diversity in the neighborhood search, we set α to 0.2 and β to 0.25 (see Sections 5.4.1 and 5.4.2). We also tune the heuristic using a relaxed version of the instance styles, that allows patients to be assigned to multiple devices over the course of the ten day planning horizon. We may in the future wish to run instances with a more abstract concept of devices that allows for the assignment of patients to multiple devices or to take the place of multiple in-person visits. Since this more relaxed version corresponds to a combinatorially more complex problem, if the heuristic converges using a given set of parameters for this problem, it should also converge for the restricted version.

7.1 Heuristic parameter tuning

To assess the selection of heuristic parameter values, we examined their effect on runtime, hypervolume, and best found solutions with respect to each of the objective values. These experiments were carried out using two of our twelve instances styles: $CL2$ and $UL2$ (see Table 4). For each of the two instance styles, we compared the results of eight heuristic runs resulting from each possible combination of two Tabu tenure lengths (5 and 25) two Tabu stopping conditions (300 and 500) and two values for N (35 and 50). The averaged results over each instance style's five replications are given in Tables 5 and 6. For example, when the five replications of $UL2$ are run with a Tabu tenure of 25, a Tabu stopping condition of 500, and an N value of 35, the average runtime experienced is 2742.4 seconds, the average number of Tabu searches carried out in Phase II is 347.6, the average of the hypervolume metrics is 5.714006E+09, an average of 239.2 solutions were in the final approximation set, and the average of best found objective values over the five replications are 1672.07, 101, and 0, respectively.

Table 5: $UL2$ Heuristic parameter experiments

Tabu Tenure	Tabu Stop Condition	N	Runtime (seconds)	No. Phase II Searches	Hypervolume	Approx. Set Size	Best Cost	Best NC	Best BW
25	300	35	1170	230	5.634168E+09	188	1722.14	103.6	0
		50	1841.6	370	5.686672E+09	207.6	1706.5	102	0
	500	35	2742.4	347.6	5.714006E+09	239.2	1672.07	101	0
		50	3797.4	492	5.702654E+09	245	1656.55	103	0
5	300	35	921	167.4	5.774850E+09	202	1634.15	100.8	0
		50	1520.6	277.8	5.772374E+09	223.6	1631.82	101.8	0
	500	35	1640.8	196.8	5.803192E+09	249.4	1595.3	99.8	0
		50	3150.6	365	5.816430E+09	257.4	1578.23	100.6	0

Table 6: $CL2$ Heuristic parameter experiments

Tabu Tenure	Tabu Stop Condition	N	Runtime (seconds)	No. Phase II Searches	Hypervolume	Approx. Set Size	Best Cost	Best NC	Best BW
25	300	35	1638.8	293.4	6.252764E+09	231.2	870.91	102.6	0
		50	3088	648.6	6.239812E+09	207.6	898.33	104.2	0
	500	35	2814.6	357.8	6.303204E+09	251.4	825	101.4	0
		50	3370	424.8	6.289336E+09	265.8	830.33	101.2	0
5	300	35	1320	255.8	6.391530E+09	246.2	796.63	100	0
		50	2664.6	502.2	6.388714E+09	283.2	820.73	100.2	0
	500	35	2649.8	308	6.402850E+09	298.6	768.34	100.6	0
		50	4752.4	596.8	6.460438E+09	324.6	754.83	98.8	0

It is clear in that in terms of average runtime for comparable Tabu stopping conditions and N values, averaged hypervolume metrics, and averaged best found solutions for the three objectives, that the shorter Tabu tenure yields better results than the longer Tabu tenure for both $UL2$ and $CL2$ instance styles. We therefore conclude that the heuristic yields better results with the shorter Tabu tenure, and we use a Tabu tenure length of 5 when we run the full set of twelve instance styles. Deciding the appropriate values of the Tabu stopping condition and N is not quite as straightforward. It is clear from Tables 5 and 6 that, assuming a short Tabu tenure, a Tabu stopping condition of 500 with an N value of 50 results in the best average hypervolume metric, and best average individual objective values (with the exception of nurse consistency for the $UL2$ instance set) of the four combinations of heuristic parameters. However, this improved performance comes at the expense of significant increases in average runtime for both instance styles. Tables 7 and 8 show the tradeoff in hypervolume improvement versus increase in average runtimes for the four combinations of heuristic parameters assuming a short Tabu tenure. The percent improvement in hypervolume and the runtime multipliers are calculated from the base case of a Tabu stopping condition of 300 and an N of 35, so Table 8 illustrates that for the $CL2$ instance set, a Tabu stopping condition of 500 with an N of 35 results in approximation sets that have 0.177% larger hypervolume on average, but that take 2.01 times as long to run on average. Even when the runtime is almost quadrupled, average hypervolume improves a mere 1.078%, on average. Since the base case runs of $UL2$ and $CL2$ replications already average 15 to 20 minutes in runtime, respectively, we chose to use these parameters (Tabu stopping condition of 300 and N of 35) to run the full set of instances.

Table 7: $UL2$ Hypervolume improvement versus runtime increase

Tabu Stop Condition	N	Runtime (seconds)	Hypervolume	Percent Increase in Hypervolume	Runtime Multiplier
300	35	921	5.774850E+09		
	50	1520.6	5.772374E+09	-0.0429%	1.65
500	35	1640.8	5.803192E+09	0.491%	1.78
	50	3150.6	5.816430E+09	0.720%	3.42

Table 8: $CL2$ Hypervolume improvement versus runtime increase

Tabu Stop Condition	N	Runtime (seconds)	Hypervolume	Percent Increase in Hypervolume	Runtime Multiplier
300	35	1320	6.391530E+09		
	50	2664.6	6.388714E+09	-0.0441%	2.02
500	35	2649.8	6.402850E+09	0.177%	2.01
	50	4752.4	6.460438E+09	1.078%	3.6

7.2 Heuristic convergence and comparison to lower bounds

As noted in Section 5.5.1, we can be more confident about our heuristic implementation if the hypervolume metric converges before the heuristic run comes to an end. To evaluate this, we graph the hypervolume metric as a function of the the number of Tabu searches that have been carried out for some example heuristic runs with a Tabu tenure length of 5. Refer to Figures 9 and 10 depicting a single replication from the $UL2$ instance set, and Figures 11 and 12 depicting a single replication from the $CL2$ instance set, where hypervolume is first calculated after the tabu search corresponding to the first objective, travel cost, in Phase I of the heuristic. It increases dramatically through the tabu searches corresponding to other objectives in Phase I of the heuristic, then converges after approximately 25 to 100 Phase II tabu searches, depending on choice of Tabu stopping condition and N.

Figure 9: Example replication of $UL2$

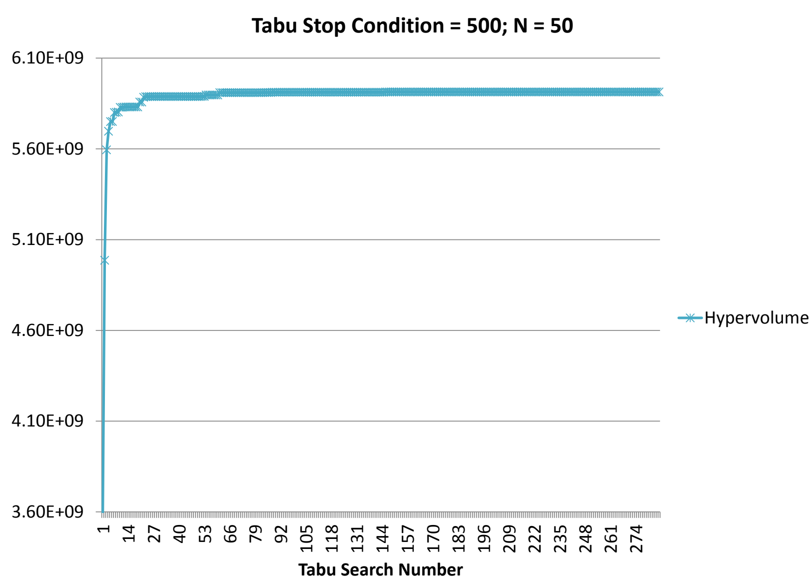


Figure 10: Example replication of $UL2$

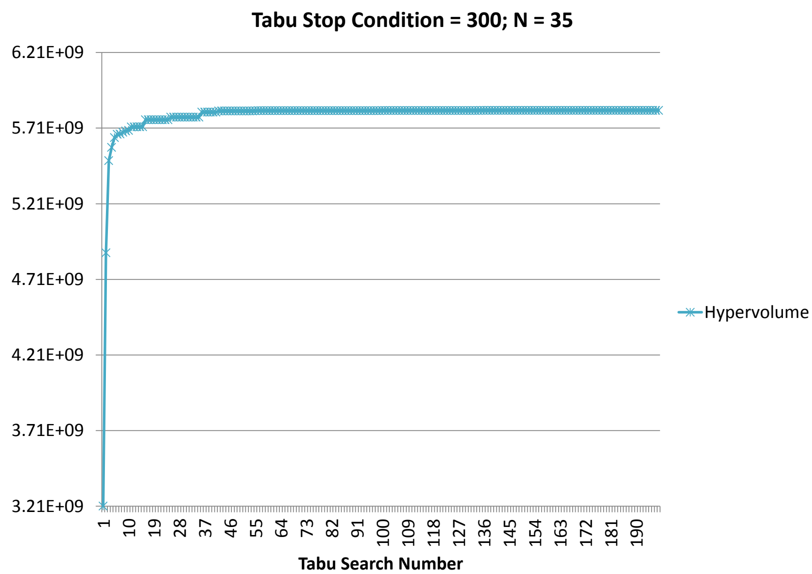


Figure 11: Example replication of $CL2$

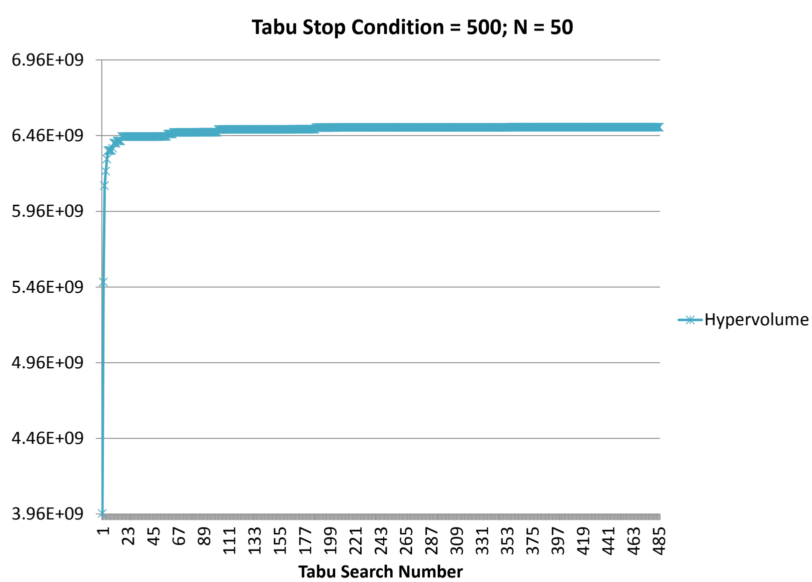
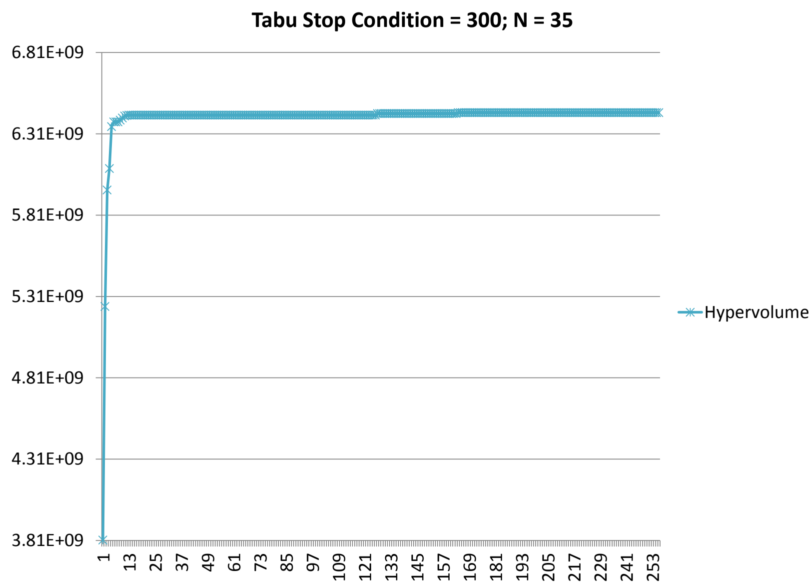


Figure 12: Example replication of $CL2$



We also compare the quality of the averaged best found solutions with respect to each of the nurse consistency and balanced workload objectives with the lower bounds discussed in Section 5.5.2. We do not make any comparison regarding the total cost objective. Tables 9 and 10 show this comparison across replications of $UL2$ and $CL2$ for the four combinations of heuristic parameters. Since all replications of both $UL2$ and $CL2$ achieve the lower bound for the balanced workload objective with any heuristic parameter setting, we know that the best solutions found with respect to that objective are in fact optimal for balanced workload. While the best found nurse consistency solutions do not achieve their lower bounds, they do come within 11% of those bounds over all heuristic parameter settings. As an example, for the set of $UL2$ replications run with a Tabu stopping condition of 500 and N of 35, the best found nurse consistency solutions are 8.48% greater on average than the lower bound of 92 for those instances. Since we do not know the optimal nurse consistency objective values for instances of this size, we may only conclude from this comparison that on average, the optimality gap for this set of replications is no worse than 8.48%.

Table 9: $UL2$ comparison to lower bounds

Stopping Condition	N	Averaged Best Nurse Consistency	Lower Bound Nurse Consistency	Gap
300	35	100.8	92	9.57%
300	50	101.8	92	10.65%
500	35	99.8	92	8.48%
500	50	100.6	92	9.35%

Stopping Condition	N	Averaged Best Balanced Workload	Lower Bound Balanced Workload	Gap
300	35	0	0	0.0%
300	50	0	0	0.0%
500	35	0	0	0.0%
500	50	0	0	0.0%

Table 10: $CL2$ comparison to lower bounds

Stopping Condition	N	Averaged Best Nurse Consistency	Lower Bound Nurse Consistency	Gap
300	35	100	92	8.70%
300	50	100.2	92	8.91%
500	35	100.6	92	9.35%
500	50	99.8	92	8.48%

Stopping Condition	N	Averaged Best Balanced Workload	Lower Bound Balanced Workload	Gap
300	35	0	0	0.0%
300	50	0	0	0.0%
500	35	0	0	0.0%
500	50	0	0	0.0%

8 Results

We present the results of our twelve instance styles, including an analysis of the tradeoff among objectives and managerial insights regarding the assignment of devices to patients.

8.1 General results

We used the heuristic described in Section 7, with a Tabu tenure of 5, Tabu stopping condition of 300, and N of 35, to run 5 replications of each of our twelve instance styles detailed in Section 6. The average runtime, number of Tabu searches in Phase II, number of nondominated solutions, hypervolume, and best found objective values are given in Table 11 for each instance style. The first thing to note is that the heuristic is able to find optimal solutions with respect to the balanced workload objective for all replications of all twelve instance styles. Table 12 shows the gap between the average best found solutions and their respective lower bounds for the nurse consistency objective. The standard deviation in the hypervolume metric across the five replications for each instance style is given in Table 13.

Table 11: Averaged results of all instance styles

Instance Style	Average Runtime	Average Phase II It.	Approximation Set Size	Average Hypervolume	Average Best Cost	Average Best NC	Average Best BW
$UL1$	1418.6	287.2	183.6	5.807318E+09	1662.06	98.8	0
$UL2$	1067.2	223	159.4	5.757016E+09	1669.94	101.2	0
$US1$	1290.2	283	201.2	6.517416E+09	725.62	97.2	0
$US2$	1253.4	236.4	219	6.393974E+09	777.72	99.8	0
$CL1$	1643.8	333.8	203.6	6.355052E+09	929.43	97.8	0
$CL2$	1875	410.6	243.8	6.342016E+09	869.99	100.8	0
$CS1$	2229.8	475.4	284.4	6.732122E+09	414.92	97.8	0
$CS2$	1787	369	332.6	6.676592E+09	436.09	98.8	0
$UCL1$	1245	247	223.2	6.025572E+09	1411.75	98.4	0
$UCL2$	1783.6	342.8	263.4	6.029044E+09	1316.48	101	0
$UCS1$	1242.4	243.4	272.4	6.549032E+09	628.71	98	0
$UCS2$	1497.4	294.6	239.2	6.530052E+09	643.83	99.4	0

Table 12: Best-case nurse consistency gap

Instance Style	Average Best NC	Lower Bound	Gap
$UL1$	98.8	90	9.78%
$UL2$	101.2	92	10.00%
$US1$	97.2	90	8.00%
$US2$	99.8	92	8.48%
$CL1$	97.8	90	8.67%
$CL2$	100.8	92	9.57%
$CS1$	97.8	90	8.67%
$CS2$	98.8	92	7.39%
$UCL1$	98.4	90	9.33%
$UCL2$	101	92	9.78%
$UCS1$	98	90	8.89%
$UCS2$	99.4	92	8.04%

Table 13: Variance in hypervolume metric

Instance Style	Hypervolume Standard Deviation
$UL1$	8.138E+07
$UL2$	4.511E+07
$US1$	5.239E+07
$US2$	4.991E+07
$CL1$	2.065E+08
$CL2$	1.444E+08
$CS1$	8.548E+07
$CS2$	1.088E+08
$RCL1$	9.967E+07
$RCL2$	1.129E+08
$RCS1$	5.322E+07
$RCS2$	8.259E+07

The purpose of generating an approximated Pareto front is so that decision makers may choose the solution that is most appropriate given their particular agency's priorities. However, even a high level analysis of the results presented in Table 11 show that the extent to which a home health agency may achieve these various objectives of minimal cost, nurse consistency, and balanced workload depends on the configuration of their service network. The instance styles developed in Section 6 are twelve examples of possible configurations resulting from various service area sizes, patient location distributions, and patient visit distributions.

While it is likely intuitive that agencies with smaller service areas can experience much smaller best-case total costs (and indeed is confirmed for small versus large service areas in the results table above), it is interesting to note from Table 11 that patient location distribution also has a significant effect on best-case total cost. For example, the average best-case cost for the uniformly distributed instance set $UL1$ is 1662.06, while the best-case cost for the clustered set of instances, $CL1$, is 929.43, almost a 45% decrease in cost simply as a result of the patients being clustered in the service region as opposed to being uniformly distributed. Not surprisingly, the instance style that combines these two patient distributions, $UCL1$, has a best-case average total cost of 1411.75, which falls between the averaged costs for the instance styles reflecting an only uniform or only clustered patient location distribution. This trend holds for various patient distributions controlling for both small and large geographic region size.

In addition to its obvious effect on cost, agency service area also seems to affect best-case nurse consistency. For each combination of patient location distribution and patient visit distribution style, the instance style with a smaller service region achieves better or equal average best-case nurse consistency than the large service region style. As an example, the large service area with uniform patient distribution style and patient visit distribution 1 ($UL1$) has an average nurse consistency, 98.8, while the corresponding small service area instance set ($US1$) has better best-case nurse consistency of 97.2. As can be seen in Table 12, this comparison between small and large geographic areas holds for all instance sets, controlling for patient location and patient visit distributions (the gap for $UCL1$ is 9.33%, while the gap for $UCS1$ is only 8.89%). This effect is likely due to the reduced travel times among patients in instances with relatively small geographic areas, which allows the nurses freedom to see more patients in a shorter period of time and more likely to achieve nurse consistency. It is also worth noting that the best-case nurse consistency scores appear to be better for some sets of instance styles with patient visit distribution style 1, as can be seen in the comparison of $CL1$ with a nurse consistency of 97.8 to the nurse consistency score of 100.8 for instance set $CL2$. In reality, instance styles with patient visit distribution 1 have fewer patients than than those with distribution 2 (as detailed in Section 6). As shown in Table 12, the best-case nurse consistency gap for instances with patient visit distribution 1 do not seem to perform any better in comparison with their lower bound than those with patient visit distribution

2.

Agency service area size and patient location distribution also seem to affect the variance in hypervolume metric across the five replications of the various instance styles. Observe from Table 13 that controlling for patient location and patient visit distributions, the instance styles with larger service regions have greater variability in hypervolume metric than do those with smaller regions, as can be in seen the comparison of instance style $CL1$ with standard deviation 2.065E+08 with style $CS1$, which has a smaller standard deviation of 8.548E+07. This is likely because a larger region size may contribute to a greater number of patient location and schedule options than a smaller region size, thus increasing the variance in feasibility of various nurse routes (and therefore variance in the approximation set) across the replications. The uniform patient location distribution with patient visit distribution 1 styles are the exception to this trend. Note also that when controlling for service area size and patient visit distribution, instance styles with uniformly distributed patients have the least variability in hypervolume across the five replications, while the clustered styles have the most variability, and the uniform with clustered styles fall between the two. The increased variance in hypervolume metric for instance styles that involve the clustering of some or all patients is likely due to the fact that for each style, there is at least one replication in which at least one cluster does not contain a nurse. When all patient clusters contain a nurse, the clustered instance styles may achieve better best found values for the individual objectives (see Table 11), but replications where not all clusters contain nurses cannot perform as well with respect to the various objectives. This necessarily drives down the hypervolume metric for these replications, and increases variability across the instance style.

Figures 13, 14, and 15 depict examples of the three dimensional graph in objective space of the approximation set generated by the heuristic procedure for three replications of the $CS2$ instance style. These examples, which have an approximately similar shape in the solution space, may give some indication as to the shape of the true efficient frontier for this instance style. They include a number of compromise solutions, and a number of solutions in which near-optimal values for a single objective are coupled with poor values for alternate objectives. For example, in Figure 13, point $(479.173, 129, 132)$ represents a compromise solution with relatively good objective values for total cost, nurse consistency, and balanced workload. Alternatively, point $(398.106, 228, 734)$ is a solution with near-optimal total routing cost and poor nurse consistency and balanced workload.

Figure 13: Approximation set from $CS2$ instance style

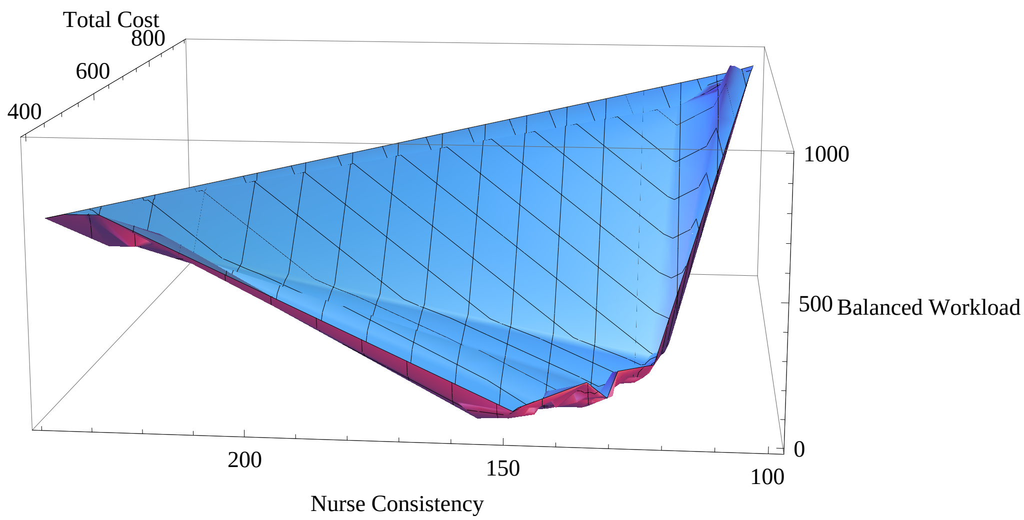


Figure 14: Approximation set from $CS2$ instance style

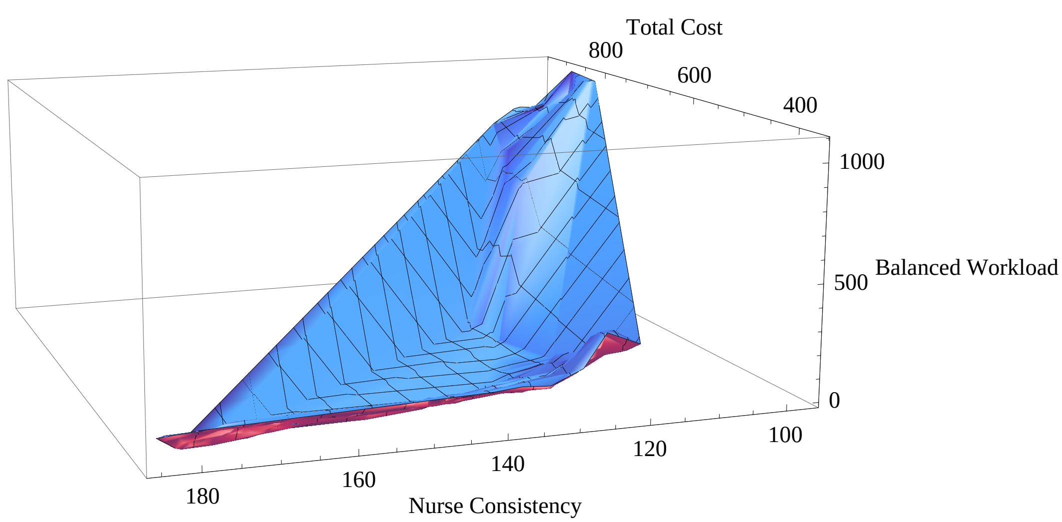
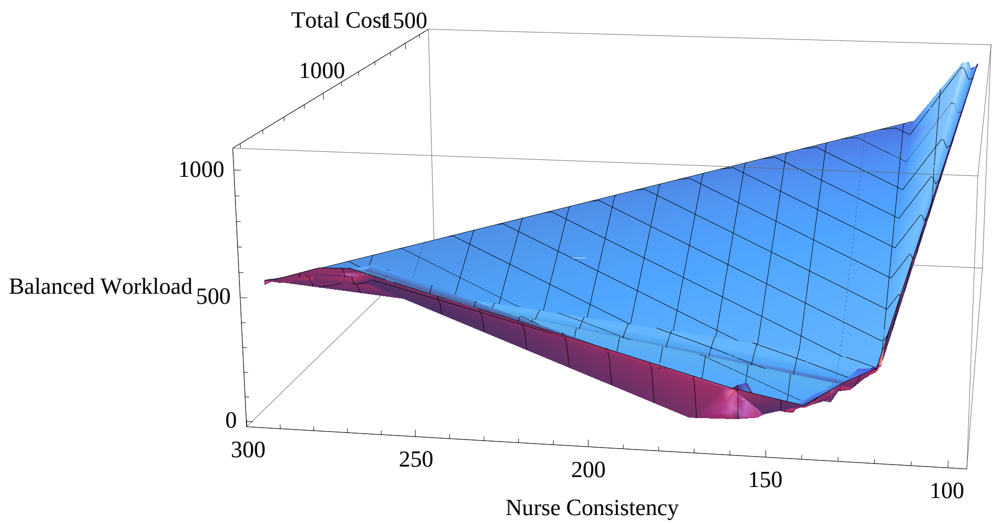


Figure 15: Approximation set from $CS2$ instance style



8.2 Compromise solutions

Decision makers are often interested in finding compromise solutions that perform relatively well in all or multiple of the total cost, nurse consistency, and balanced workload objectives. We analyzed the approximation sets of the five replications of our twelve instance styles to find the number and quality of compromise solutions found by the heuristic. We defined compromise bins based on the percent degradation allowed from the best found objective value, taken at 5% intervals for each of the three objectives. Figures 16, 17, and 18 are histograms depicting the number of replications of the sixty total replications solved by our heuristic that have at least one solution within each of the compromise bins. For example, Figure 16 shows that only one replication of the sixty contained a solution in the approximation set that was within 30% of the best found for cost, 15% of the best found for nurse consistency, and 5% of the best found for balanced workload. The same figure shows that no replications had a solution within 20% of best cost, 15% of best nurse consistency, and 5% of best balanced workload. For each of the three histograms corresponding to balanced workload objectives within 5%, 10%, and 15% of the best found, respectively, all sixty replications had some solution within 50% of the best for both total cost and nurse consistency.

Note that increasing the allowed degradation of the balanced workload objective from 5% in Figure 16 to 15% in Figure 18 does not appear to have a significant impact in compromise bin counts. In fact, the overall shape of the histogram and the magnitudes of the compromise bins remain approximately the same as in Figures 16 and 17, where balanced workload is required to be closer to its best found objective value. We may conclude that it is easier to achieve balanced workload in compromise solutions than to simultaneously achieve low cost and good levels of nurse consistency. Figure 18 shows that over half the replications contained some solution within 25% of the best found for all three objectives. Therefore, we may conclude that using our solution methodology, many agencies will be able to find nurse schedules and device assignments that perform relatively well with respect to all of their total cost, nurse consistency, and balanced workload goals.

Figure 16: Compromise solution histogram (balanced workload 0.05)

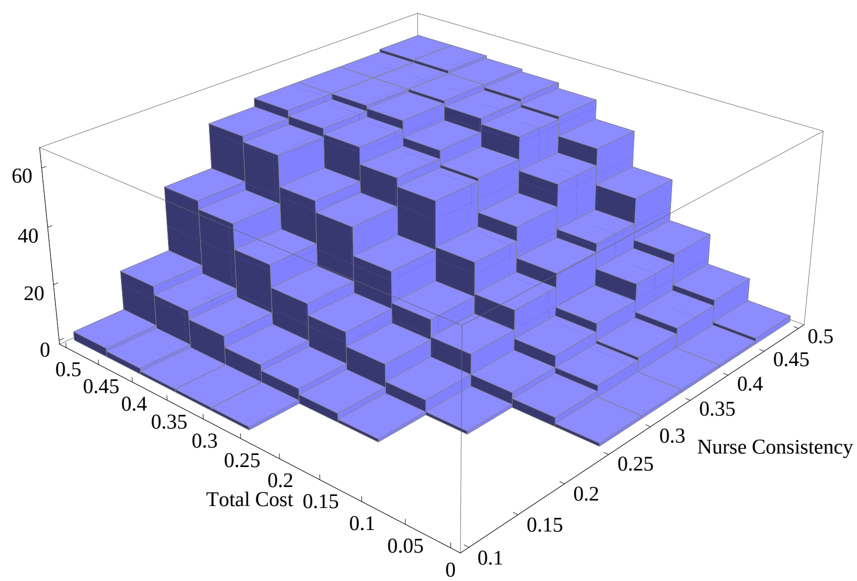


Figure 17: Compromise solution histogram (balanced workload 0.1)

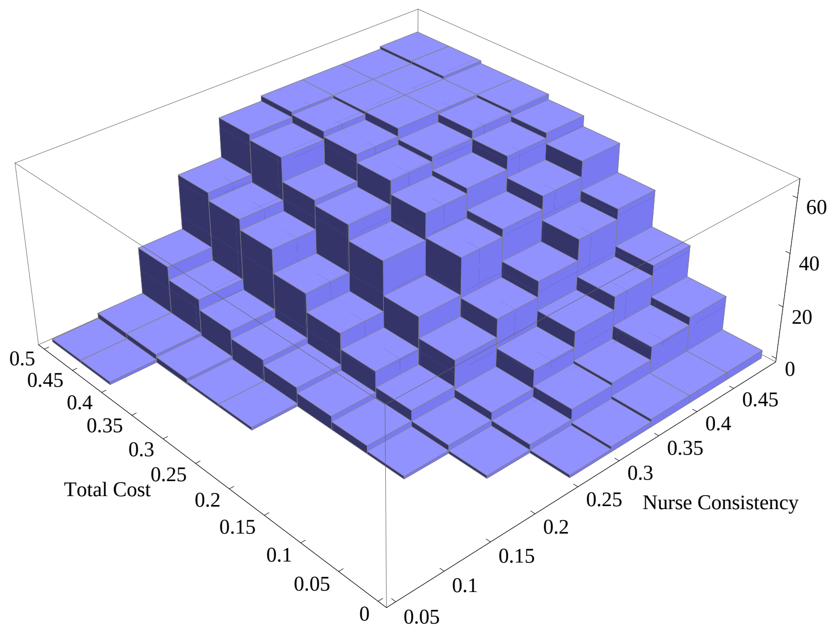
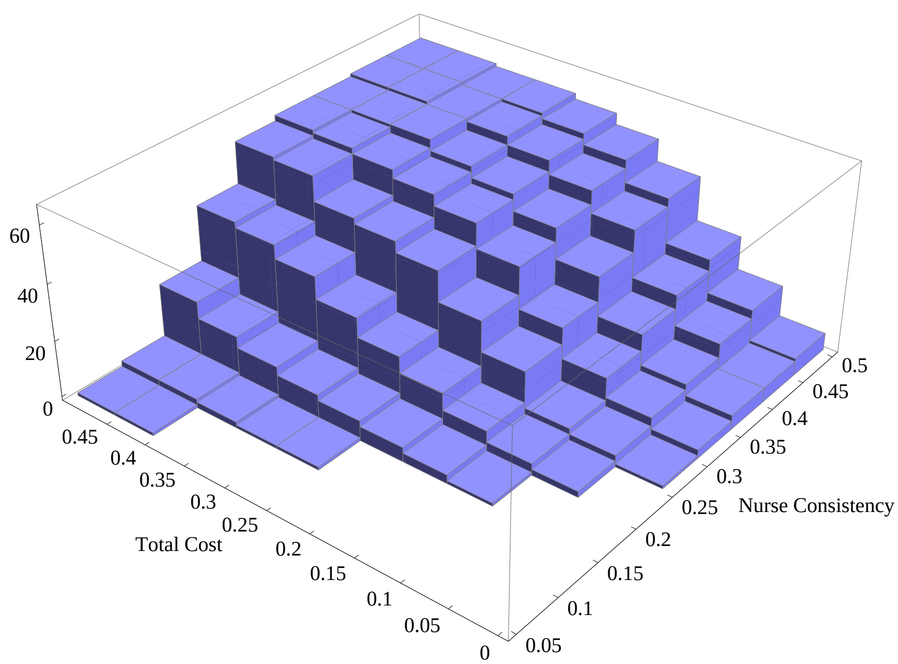


Figure 18: Compromise solution histogram (balanced workload 0.15)



8.3 Device assignments insight

We are interested in providing insight into making device assignments for various instance styles. To this end, we examined two replications, one each of the $UL2$ and $CL2$ instance styles. For each replication, we first find for each patient in the instance replication the proportion of solutions in the approximation set that make a device assignment to that patient. That is, for each patient we have a measure, between 0 and 1, of the percent of nondominated solutions for which that patient is assigned to a device. We then examine the top ten most frequently assigned device patients for the replication to gain insight into the types of patient assignments that characterize the efficient set approximation and hopefully result in good solutions with respect to the three objectives.

Figure 19 shows the set of patient and nurse locations for the example replication from the $UL2$ instance style, with the top ten patients most frequently assigned to devices in red. Most of the device patients are clearly on the periphery of the agency service area, but those that are relatively centrally located require either five to seven visits over the course of the planning horizon.

Figure 19: $UL2$ device patients

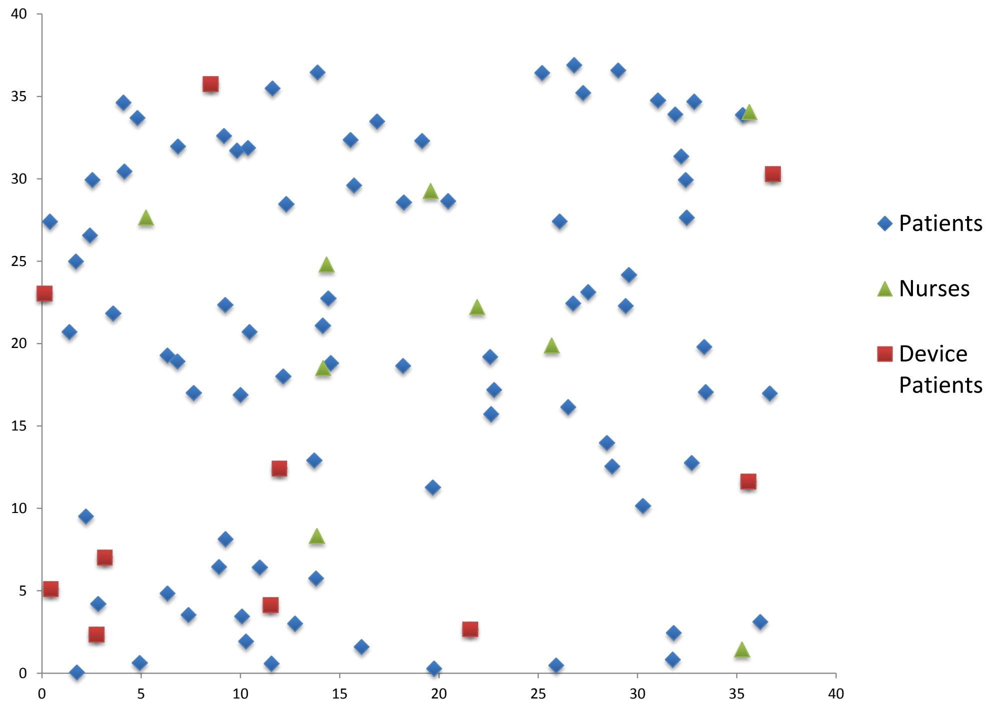
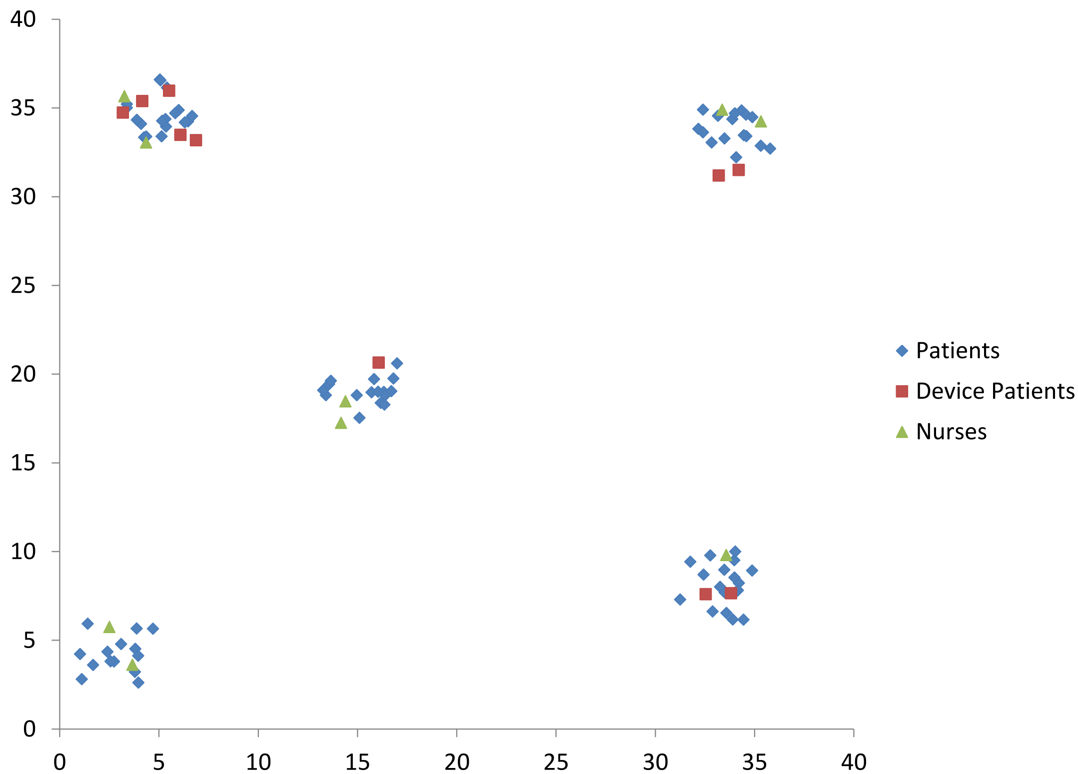


Figure 20 shows the set of patient and nurse locations for the example replication from the $CL2$ instance style, with the top ten patients most frequently assigned to devices in red. Observe that the cluster in the upper left corner of the service region contains five of these ten most frequently assigned patients, and that the cluster in the lower right corner of the region has another three. It is not surprising that the cluster in the lower right corner contains so many device patients, as only one nurse is located in that cluster, device assignments likely help ease the burden of fulfilling patient demand without requiring many costly visits from nurses located in other clusters. Upon analysis of the patient schedules for the various clusters, the reason for so many device assignments in the upper left cluster also becomes clear. It so happens that the patients in that cluster tend to require visits on the same days of the planning horizon, with an average of 11.4 required visits per day. In fact, for three of the ten days of the planning period at least 14 visits are required for patients located in that cluster (well above the average 5 patient visits per nurse per day for the two nurses located there). Contrast this with the cluster located in the lower left of the service region, where an average of 6.6 visits are required per day, with a maximum demand of 9 visits on days three and four. While nurses may serve patients located outside their cluster, it is reasonable to assume that the best solutions, those found in the approximation set, minimize this behavior as it drives up cost and takes travel time, reducing the likelihood of good nurse consistency.

Figure 20: $CL2$ device patients



This analysis, based on two case studies, it does seems to indicate some preferred patient assignment to device strategies when agencies are interested in achieving low cost, high nurse consistency, and balanced workloads for their nurses. It suggests that agencies with patient location distributions most resembling that of a uniform distribution should likely assign devices to patients located on the geographic periphery of the service area (with farther distances to nurses), and/or to patients that require a relatively high number of visits over the planning horizon as compared to other patients. For agencies with clustered patient location distributions, it seems important to assess the unique attributes of each cluster, and assign devices to those patients who do not have as many nurses located in their surrounding cluster, and/or to patients in clusters where patient schedules happen to match frequently. We further investigate two of the hypotheses raised in analysis of the example replications from the $UL2$ and $CL2$ instance styles: namely, that patients with greater distances to the closest nurse will be assigned to a device more frequently, and that more device assignments are made proportionally to patients requiring a greater number of visits over the planning horizon.

8.3.1 Correlation of patient-to-nurse distance and device assignment

We test the first hypothesis by finding the correlation for the six large instance styles between the distance from patients to their closest nurse and device assignment frequency. The data points, represented by ordered pairs of distance and device assignment frequency for each patient, were aggregated over replications for each instance style, and Figures 21 through 26 show the scatter plots of minimum distance to a nurse versus device assignment frequency for all the patients of each instance style. Note that there is a gap along the distance axis for both cluster instances (see Figures 23 and 24). This is because both clustered instance styles contain one or more replications in which at least one cluster of patients does not contain a nurse. Table 14 shows the correlation coefficients for each of the six instance styles; all six correlation coefficients are statistically significant using a t-test, with p values of essentially zero. Not surprisingly, the correlation coefficients are greater for uniformly distributed patient instance styles than for clustered patient styles, since most patients in clustered styles are located very near a nurse. This was also illustrated in the $CL2$ case study above, where distance from closest nurse did not seem to play as great a role in determining patient device assignment as it did in the $UL2$ example replication.

Figure 21: $UL1$ Nurse distance vs. device assignment frequency

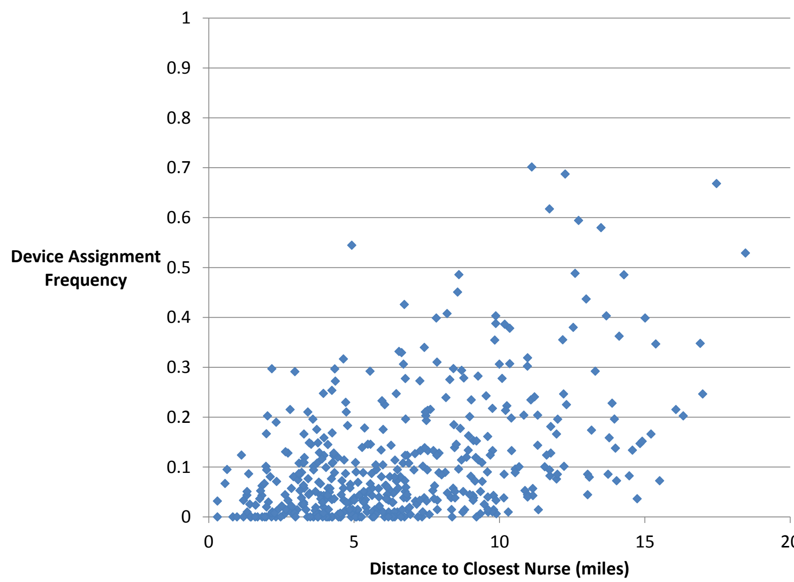


Figure 22: $UL2$ Nurse distance vs. device assignment frequency

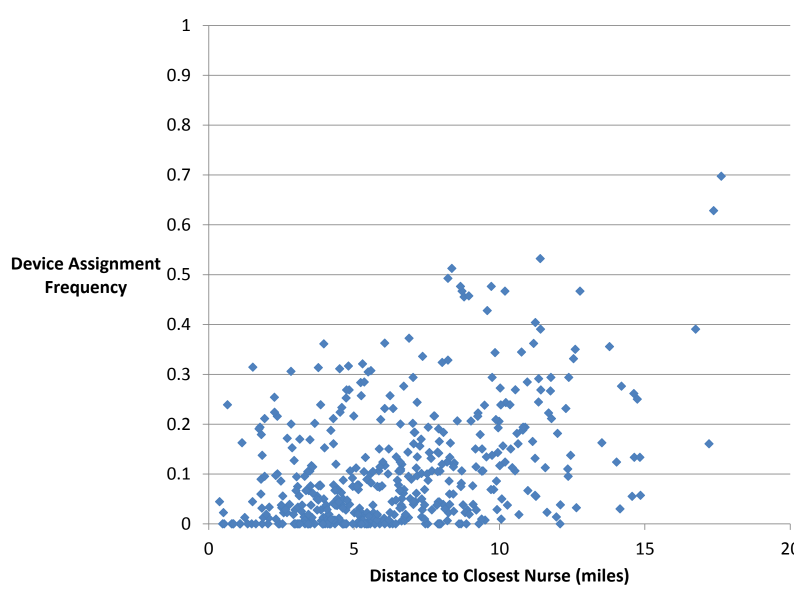


Figure 23: $CL1$ Nurse distance vs. device assignment frequency

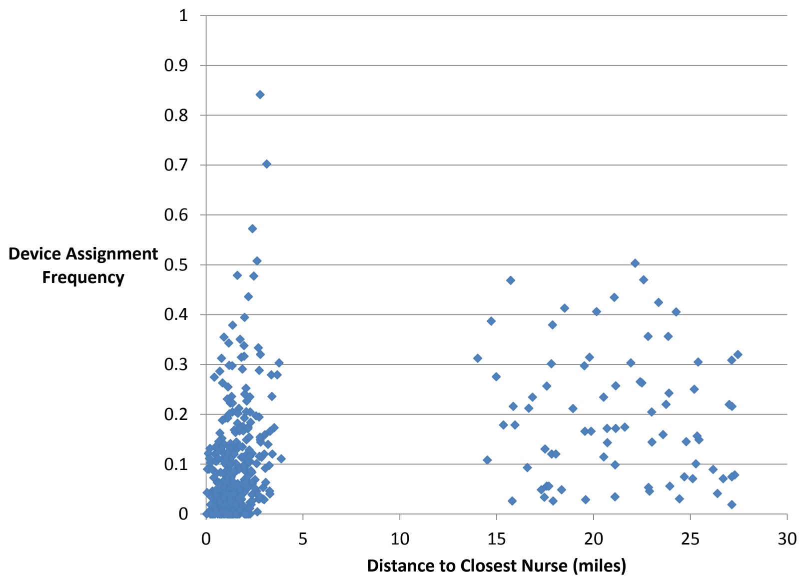


Figure 24: $CL2$ Nurse distance vs. device assignment frequency

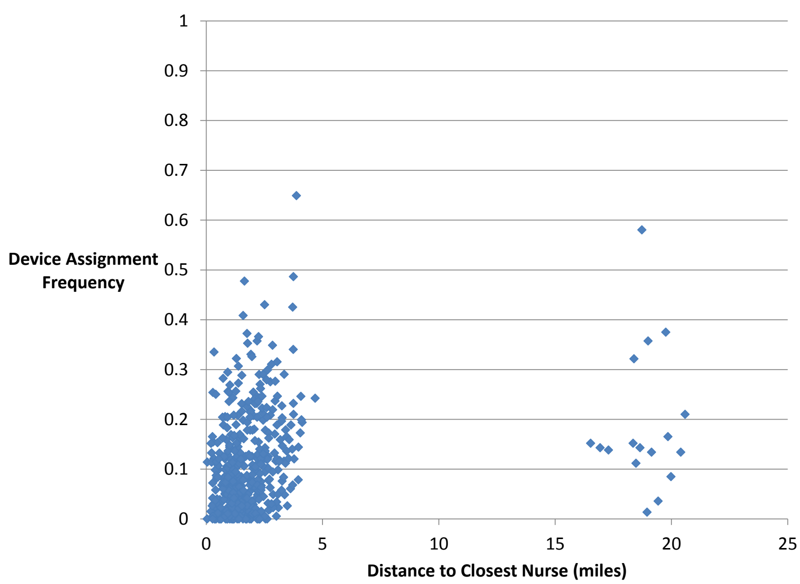


Figure 25: $UCL1$ Nurse distance vs. device assignment frequency

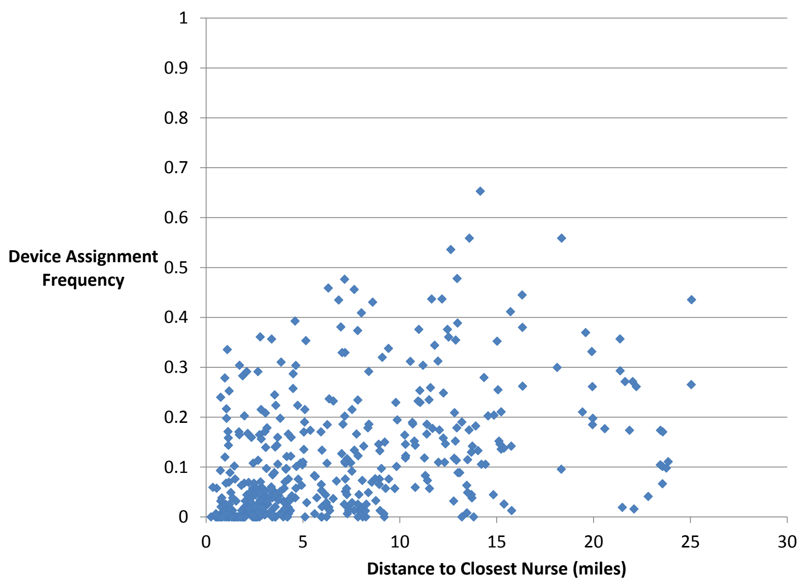


Figure 26: $UCL2$ Nurse distance vs. device assignment frequency

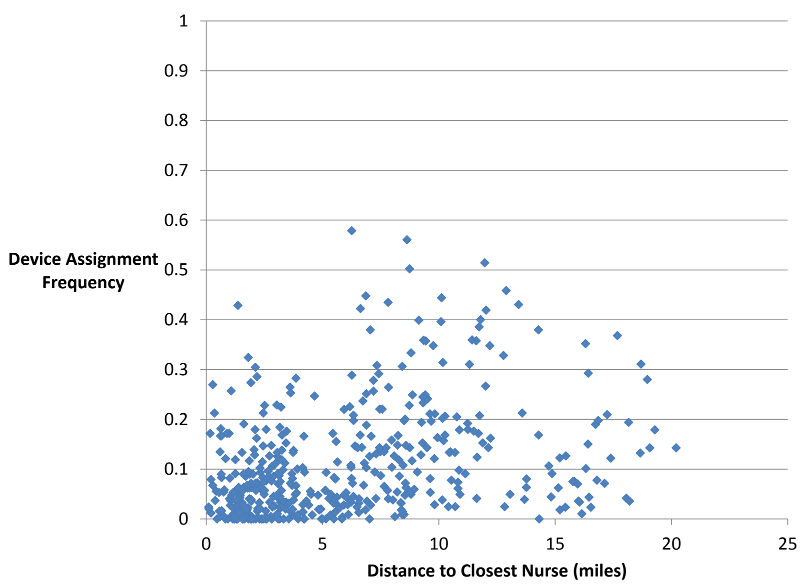


Table 14: Correlation coefficients

Instance Style	Correlation Coefficient
$UL1$	0.473
$UL2$	0.422
$CL1$	0.337
$CL2$	0.238
$UCL1$	0.452
$UCL2$	0.352

8.3.2 Patient demand and device assignment

To investigate our hypothesis that patients with more frequent demand over the planning horizon are proportionally more likely to be assigned to a device, we examine the frequency of device assignment by patient visit type, where visit type is the number of visits required over the planning period. The number of patients in each patient visit type is determined by the distributions discussed in Section 6.2; for example, our instance styles of distribution 2 have approximately the same number of patients in each patient visit type. We find the proportion of device assignments by patient visit type for each of the six large instance styles. If the distribution of device assignments by visit type is close to the distribution of patient types for each of distribution 1 and distribution 2, then we may conclude that patient visit type does not significantly affect device assignment. Let V_i be visit type i, the set of patients requiring i visits during the planning horizon. Let C_j be the number of solutions in the approximation set for which patient j was assigned to a device, T be the number of devices available for assignment (10 for our instances), and A be the number of solutions in the approximation set. Then proportion of device assignments to patient visit type i is found by

$$\frac{\sum\limits_{j \in V_i} C_j}{TA}.$$

Tables 15 and 16 show the proportion of device assignments by visit type averaged over the five replications for instance styles of patient visit distribution 1 and 2, respectively.

While the device assignment proportions by visit type do not always exactly match the patient distributions of the instance styles, they also do not seem to always be skewed towards the patients with higher visit requirements. While we hypothesized that patients requiring more visits would be more likely to be assigned devices, it is not clear from this analysis that this is always the case. Therefore, further investigation is needed to determine whether visit type should be used as a criteria for device assignment in practice. It may instead be more appropriate to use visit type in combination with location distribution or other demand characteristics as criteria for device assignment.

Table 15: Distribution 1 device assignments by visit type

Patient Visit Type	2	3	4	5	6	7	8
Distribution 1	0.05	0.125	0.125	0.4	0.125	0.125	0.05
$UL1$	0.0431	0.1420	0.1203	0.3744	0.1262	0.1185	0.0692
$CL1$	0.0660	0.1166	0.1164	0.4416	0.1166	0.0811	0.0526
$UCL1$	0.0404	0.1404	0.1656	0.4606	0.0641	0.0773	0.0461

Table 16: Distribution 2 device assignments by visit type

Patient Visit Type	2	3	4	5	6	7	8
Distribution 2	0.14	0.14	0.14	0.16	0.14	0.14	0.14
$UL2$	0.1490	0.1085	0.1579	0.1305	0.1628	0.1223	0.1618
$CL2$	0.1399	0.1162	0.1225	0.1800	0.1562	0.1383	0.1439
$UCL2$	0.1258	0.1655	0.1977	0.2015	0.1171	0.1125	0.0760

9 Future research

We propose that future work focus on further assessment of the quality of solutions developed by the heuristic, as well as the incorporation of time-related objectives.

9.1 Further heuristic validation

While our analysis supports our choice of heuristic and heuristic parameters, we would like to compare the approximation set to the exact efficient frontier for instances of relatively small size. One possibility is to create an enumerative approach that generates all possible feasible solutions for a very small instance and then collects the nondominated solutions in a list. Another option is to approximate the efficient set of the linear relaxation of our problem, as in Balachandran and Gero [1985]. To pursue the latter option, some analysis regarding the quality of the lower bounds elicited from the linear relaxation of our problem would be necessary to verify that generating the efficient set of that relaxation would provide a useful comparison to the Pareto front of our original problem.

9.2 Incorporation of additional objectives

As described previously in Section 1, time consistency and idle time are two objectives that may contribute to patient and nurse satisfaction, respectively, and therefore we are interested in incorporating these objectives into our problem. Time consistency is achieved when each patient is visited at approximately the same time every day they require service over the planning horizon. We divide each workday of the planning horizon into nonoverlapping time windows of equal duration, and model time consistency as beginning patients' care in the same time window each time service is required over the planning period. This method of assessing time consistency is new in the consistent vehicle routing problem literature, as can be seen by comparing to the formulations of time consistency detailed in Section 3.1.

Note with the attempt at scheduling patient visits in consistent time windows comes the possibility of idle time in a nurse's schedule, if the nurse arrives to a visit earlier than it is scheduled to begin. We formulate minimizing this idle time over all nurses over the planning horizon as a nurse satisfaction objective to balance the patient satisfaction objective of time consistency. It is modeled as the total amount of time worked by all nurses over the planning horizon, less the number of visits performed and travel time required to traverse the assigned routes.

To add these time-based objectives to our mathematical model in Section 4, we let the decision variables m_v^d represent the time that nurse v's workday begins on day d for $v \in V,\ d \in D$. We assume a set W of non-overlapping appointment time windows (e.g. between 8 a.m. and 9 a.m., 9 a.m. and 10 a.m., etc.) during which a patient may be visited, and index this set by $w \in W$. For each $w \in W$ we let the input parameters $[a_w, b_w]$ represent the start and end times of the window, respectively. We then define the following day-oriented and horizon-oriented binary decision variables:

$$q_n^{wd} = \begin{cases} 1 & \text{if patient } n \text{ is visited during window } w \text{ on day } d \\ 0 & \text{otherwise} \end{cases}$$

$$q_n^{w} = \begin{cases} 1 & \text{if patient } n \text{ visited in window } w \text{ during planning horizon} \\ 0 & \text{otherwise} \end{cases}$$

Then the additional two objectives may be given by:

- **Time Consistency**:

$$\min \ f_4 = \min \sum_{n \in N} \sum_{w \in W} q_n^w$$

This objective represents the total number of different appointment windows during which all patients are seen.

- **Idle Time**:

$$\min \ f_5 = \min \sum_{v \in V} \sum_{d \in D} (e_v^d - m_v^d) - \sum_{(i,j) \in A} \sum_{v \in V} \sum_{d \in D} c_{ij} x_{ijv}^d - \sum_{n \in N} \sum_{d \in D} \sum_{v \in V} r_n^d y_{nv}^d$$

This objective represents the total amount of idle time incurred over the planning horizon; that is, time between the starting and ending times of each nurse's day not spent traveling or serving a patient.

The addition of these two time-related objectives significantly complicates the neighborhood moves established in Section 5.3. While the mathematical model developed in Section 4 calculates the start time of care for each patient each day they are visited by a nurse, the heuristic procedure does not need to directly address this issue when only the original three objectives are present. When time consistency and idle time are added, the heuristic solution approach must now keep track of all patient start times. Additionally, it is not immediately obvious how a remove-and-reinsert or swap move should handle the assignment of patient start times. In the current implementation, nurses visit all patients in succession on their route, one immediately after the other. Now there exists the option of pushing the start time of the recently inserted patient back so that care will begin in a different time window. It is apparent that the neighborhood moves developed for the current three objectives must be expanded to handle the time concerns of these two new objectives, as well move strategies motivated by each of the time consistency and nurse idle time objectives.

References

Agency for Healthcare Research and Quality. National and regional estimates on hospital use for all patients from the HCUP nationwide inpatient sample (NIS): 2007 outcomes by patient and hospital characteristics for all discharges, 2009. Available from the Agency for Healthcare Research and Quality at `http://hcupnet.ahrq.gov/HCUPnet.jsp`, accessed November 3, 2009.

M. Balachandran and J.S. Gero. The noninferior set estimation (NISE) method for three objective problems. Engineering Optimization, 9(2):77–88, 1985.

M.F. Best and N.E. Thurston. Measuring nurse job satisfaction. Journal of Nursing Administration, 34(6):283–290, 2004.

Bosch website, 2011. Available at `http://www.bosch-telehealth.com/`.

L. Bradstreet, L. Barone, and L. While. Maximizing hypervolume for selection in multi-objective evolutionary algorithms. Evolutionary Computation, pages 1744–1751, 2006.

P.I. Buerhaus, D.I. Auerbach, and D.O. Staiger. The recent surge in nurse employment: causes and implications. Health Affairs, 28(4), 2009.

R. Caballero, M. Gonzáleza, F.M. Guerrerob, and J. Molina. Solving a multiobjective location routing problem with a metaheuristic based on tabu search. application to a real case in andalusia. European Journal of Operational Research, 177(3):1751–1763, 2007.

CDC. At a glance 2009: chronic disease-the power to prevent, the call to control. Technical report, Centers for Disease Control and Prevention, National Center fro Chronic Disease Prevention and Health Promotion, 2009.

J.F. Cordeau, G. Laporte, and A. Mercier. A unified tabu search heuristic for vehicle routing problems with time windows. Journal of the Operational Research Society, 52:928–936, 2001.

Datalytics LLC. Solving the scheduling and routing problem: how to cut your transportation costs up to 30%. A White Paper for Home Health Business Managers, 2010.

Department of Health and Human Services. Telehealth services fact sheet. Available at `http://www.telemedicine.com/pdfs/TelehealthSrvcsfctsht.pdf`, July 2009.

Department of Health and Human Services, Centers for Medicare and Medicaid Services. CMS proposals for changes to physician payment policies and rates for calendar year 2012. CMS Fact Sheet, July 2011. Available at `http://interactive.snm.org/docs/FSG.MPFS.2012.NPRM.07%2001%2020111.pdf`.

E.M D'Errico and M.A. Lewis. RN continuity in home health: Does it make a difference? Home Health Care Management & Practice, 22(6):427–434, 2010.

M. Emmerich, N. Beume, and B. Naujoks. An emo algorithm using the hypervolume measure as selection criteria. Evolutionary Multi-criterion Optimization, Lecture Notes in Computer Science, 3410:62–76, 2005.

S. M. Finkelstein, S.M. Speedie, and S. Potthoff. Home telehealth improves clinical outcomes at lower cost for home healthcare. Telemedicine and e-Health, 12(2), 2006.

M. Fleischer. The measure of Pareto optima: Applications to multi-objective metaheuristics, volume 2632 of Lecture Notes in Computer Science, pages 519–533. Springer-Verlag, 2003.

P. Francis, K. Smilowitz, and M. Tzur. Flexibility and complexity in periodic distribution problems. Naval Research Logistics, 54, 2007.

I. García, J. Pacheco, and A Alvarez. Optimizing routes and stock. Accepted to Journal of Heuristics, 2011. Available at http://0-www.springerlink.com.library.uark.edu/content/d680867p0t505022/.

C. Groër, B. Golden, and E. Wasil. The consistent vehicle routing problem. Manufacturing & Service Operations Management, 11(4):630–643, 2009.

M.P. Hansen. Tabu search for multiobjective optimization: MOTS. Conference Paper for MCDM, 1997.

B. T. Horowitz. Intel telehealth system helps heart failure patients. eWeek, June 2010. Available at http://www.eweek.com/c/a/Health-Care-IT/Intel-Telehealth-System-Helps-Heart-Failure-Patients-153973/.

IRS. Irs increases mileage rate to 55.5 cents per mile. IR-2011-69, October 2011. Available at http://www.irs.gov/newsroom/article/0,,id=240903,00.html.

N. Jozefowiez, F. Semet, and E.G. Talbi. Multi-objective vehicle routing problems. European Journal of Operational Research, 189:293–309, 2008.

C. King. To your health: Intel and ge's joint venture. E-Commerce Times, August 2010. Available at http://www.technewsworld.com/story/70586.html.

F. Li, B. Golden, and E. Wasil. Very large-scale vehicle routing: new test problems, algorithms, and results. Computers & Operations Research, 32:1165–1179, 2005.

T. MacDonald, K. Dörner, and X. Gandibleux. Metaheuristics for the consistent nurse scheduling and routing problem. Working Paper, 2009.

D.K. McNeese-Smith. Experience before and throughout the nursing career: a content analysis of staff nurse descriptions of job satisfaction and dissatisfaction. Journal of Advanced Nursing, 29 (6):1332–1341, 1999.

Medicare.gov. Home health compare. Online database, 2010.

NAHC. Basic statistics about home care. Technical Report, National Association for Home Care and Hospice, 2007. Available at http://www.nahc.org/facts/homecareStudy.html.

NAHC. Study shows home health care workers drive nearly five billion miles to serve elderly and disabled patients. Technical Report, National Association for Home Care and Hospice, 2009. Available at http://www.nahc.org/facts/homecareStudy.html.

National Association for Home Care and Hospice. Basic statistics about home care. Technical report, 2010.

M. Navaie-Waliser, P. Lincoln, M. Karuturi, and K. Reisch. Increasing job satisfaction, quality care, and coordination in home health. Journal of Nursing Administration, 34(2), 2004.

J. Pacheco and R. Marti. Tabu search for a multi-objective routing problem. Journal of the Operational Research Society, 57:29–37, 2006.

D. Peppers and M. Rogers. Managing Customer Relationships: A Strategic Framework. John Wiley and Sons, Inc., Hoboken, NJ, 2004.

Philips national study on the future of technology & telehealth in home care, April 2008. Accessed November 2011.

Philips website, 2011. Available at http://www.healthcare.philips.com/usen/products/telehealth/products/telestation.wpd accessed November 2011.

J. Rau. Medicare to begin basing hospital payments on patient-satisfaction scores. Kaiser Health News, April 2011. Available at http://www.kaiserhealthnews.org/Stories/2011/April/28/medicare-hospital-patient-satisfaction.aspx.

D. Russell, R.J. Rosati, P. Rosenfeld, and J.M. Marren. Continuity in home health care: Is consistency in nursing personnel associated with better patient outcomes? Journal for Healthcare Quality, 2011. Forthcoming, available online on January 14, 2011.

K. Smilowitz, M. Nowak, and T. Jiang. Workforce management in periodic delivery operations. Working Paper No. 09-004, 2009.

J. M. Steeg. Mathematical models and algorithms for home health care services. Thesis, 2008.

E.G. Talbi. Metaheuristics: from design to implementation. John Wiley and Sons, Inc., Hoboken, NJ, 2009.

United States Department of Labor. Career guide to industries, 2010–11 edition: Healthcare. Bureau of Labor Statistics, 2010. Available at http://www.bls.gov/oco/cg/cgs035.htm.

US Census. Projected population of the united states, by age and sex: 2000 to 2050, 2004. Available at http://www.census.gov/population/www/projections/usinterimproj.

U.S. Census Bureau. Census 2010 geographic identifiers. American Factfinder Database, 2012.

L. While, P. Hingston, L. Barone, and S. Huband. A faster algorithm for calculating hypervolume. IEEE Transactions on Evolutionary Computation, 10(1):29–38, 2006.

R.T. Wong. Vehicle routing for small package delivery and pickup services. The Vehicle Routing Problem: Latest Advances and New Challenges, 2008.

H. Zhong, R.W. Hall, and M. Dessouky. Territory planning and vehicle dispatching with driver learning. Transportation Science, 41(1):74–89, 2007.

E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca, and V. Grunert da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. IEEE Transactions on Evolutionary Computation, 7(2), 2003.