On the Effectiveness of Set Covering Formulations for the Vehicle Routing Problem with Time Windows*

Julien Bramel
Graduate School of Business, Columbia University
David Simchi-Levi
Dept. of Industrial Engineering and Management Sciences, Northwestern University

October 1993 Revised: August 1994

Abstract

The Vehicle Routing Problem with Time Windows (VRPTW) is one of the most important problems in distribution and transportation. A classical and recently popular technique that has proven effective for solving these problems is based on formulating them as a *set covering problem.* The method starts by solving its *linear programming relaxation*, via column generation, and then uses a branch and bound strategy to find an integer solution to the set covering problem: a solution to the VRPTW. An empirically observed property is that the optimal solution value of the set covering problem is very close to its linear programming relaxation which makes the branch and bound step extremely efficient. In this paper, we explain this behavior by demonstrating that for any distribution of service times, time windows, customer loads and locations, the relative gap between fractional and integer solutions of the set covering problem becomes arbitrarily small as the number of customers increases.

The Vehicle Routing Problem with Time Windows (VRPTW) can be stated as follows: a set of customers dispersed in a geographic region has to be served by a fleet of vehicles initially located at a given depot. Each customer has a demand that must be picked up and each specifies a period of time, called a *time window*, in which this pick up must occur. The total load carried by each vehicle can be no more than the vehicle capacity. The objective is to find a set of routes for the vehicles, where each route begins and ends at the depot, serves a subset of the customers without violating the capacity and time window constraints, while minimizing the total length of the routes.

Due to the wide applicability and the economic importance of the problem in the service industry, variants of it have been extensively studied in the vehicle routing literature; for a review see

*Research supported in part by ONR Contract N00014-90-J-1649, N00014-95-1-0232, NSF Contracts DDM-8922712 and DDM-9322828.

Solomon and Desrosiers (1988). Since the problem is $\mathcal{NP}$-hard, much of the research has focused on the study of heuristics. These are methods that do not necessarily generate the optimal solution, but tend to produce solutions close to the optimal one. In almost all cases these heuristics are analyzed *empirically*; they are tested on a set of standard test problems and their performance is compared to the performance of other heuristics. For a survey see Solomon (1986) or Fisher (1992).

By contrast, very few papers have appeared that study the problem analytically in an attempt to *characterize* the theoretical behavior of heuristics or the underlying structure of the optimal solution. Exceptions are the recent works of Bramel, Li and Simchi-Levi (1994) who analyzed a special case of the VRPTW, and Bramel and Simchi-Levi (1993) who determine the structure of the asymptotic optimal cost of the VRPTW. This analysis has also led to an algorithm which is attractive from both theoretical and practical points of views, see Bramel and Simchi-Levi (1993). Independently of that work, and using a different technique, Federgruen and van Ryzin (1992) have analyzed the VRPTW as well as some related bin packing problems.

In this paper we study a classical and recently popular technique for solving complex combinatorial problems informally referred to as the *column generation approach*. It is based on formulating the routing problem as a very large-scale linear integer program known as a *set covering problem*. Unfortunately, the mathematical program cannot be solved directly due to its enormous size even for relatively small problems. To overcome this difficulty, the linear programming relaxation of the set covering formulation is solved using a standard column generation approach, and then a branch and bound routine attempts to "close the gap" between integer and fractional solutions. The size of this gap usually determines whether or not the method will work well.

Fortunately, this method has proven successful in solving vehicle routing problems (e.g., Desrochers, Desrosiers and Solomon (1992)) as well as crew scheduling problems (e.g., Hoffman and Padberg (1993)). In this paper, we try to analytically characterize the effectiveness of the method. Our objective is to better understand why this method works so well in practice. *Indeed, we show that for essentially any distribution of customer locations, time windows and service times, the relative gap between fractional and integer solutions of the set covering formulation of the VRPTW decreases to zero as the number of customers increases.*

In the next section we discuss the model and our main result, while in Section 2 we present a simplified model which captures the essential ideas of the proof. In Section 3 we provide a formal proof for the results. In Section 4 we discuss some issues relating to the implementation of a column generation technique. Finally, in Section 5 we show that other formulations of the VRPTW can lead to diverging fractional and integer solution values.

1 The Model and the Main Result

To formally present the vehicle routing model, let the index set of the n customers be denoted $N = \{1, 2, \ldots, n\}$ and let $Y_i \in \mathbb{R}^2$ be the location of customer $i \in N$. Assume, without loss of generality, that the depot is at the origin and, by rescaling, that the vehicle capacity is 1 and the length of the working day is 1. Associated with customer $i \in N$, are four values w_i, e_i, s_i and l_i, called *customer parameters*, which represent, respectively, the load that must be picked up, the earliest starting time for service, the time required to complete the service, called the *service time*, and the latest time service can end. Clearly, feasibility requires that $e_i + s_i \leq l_i$ and $w_i, e_i, s_i, l_i \in [0, 1]$, for each $i \in N$.

A classical method, suggested by Balinski and Quandt (1964), is based on formulating the VRPTW as a set covering problem. The idea is as follows. Let the index set of all feasible routes be $\{1, 2, \ldots, R\}$, and let c_r be the length of route r. Define

$$\alpha_{ir} = \begin{cases} 1, & \text{if customer } i \text{ is served in route } r, \\ 0, & \text{otherwise,} \end{cases}$$

for each $i = 1, 2, \ldots, n$ and $r = 1, 2, \ldots, R$. Finally, for $r = 1, 2, \ldots, R$, let

$$y_r = \begin{cases} 1, & \text{if route } r \text{ is in the optimal solution} \\ 0, & \text{otherwise.} \end{cases}$$

In the *Set Covering formulation* of the VRPTW, the objective is to select a minimum cost set of feasible routes such that each customer is included in some route. It is the following integer program, which we call problem P:

$$\min \Big\{ \sum_{r=1}^{R} y_r c_r \Big| \sum_{r=1}^{R} y_r \alpha_{ir} \geq 1, \forall i = 1, 2, \ldots, n, \quad y_r \in \{0, 1\}, \forall r = 1, 2, \ldots, R \Big\}.$$

This formulation was first used successfully by Cullen, Jarvis and Ratliff (1981) to design heuristic methods for some routing problems. Recently, Desrochers, Desrosiers and Solomon (1992) have used it in conjunction with other methods to generate optimal or near optimal solutions to the VRPTW.

Of course, the set of all feasible routes is extremely large and one cannot expect to generate it completely. Even if this set is given, it is not clear how to solve the set covering problem since it is a large scale integer program. To overcome the first difficulty, Desrochers, Desrosiers and Solomon use the celebrated column generation technique, which makes it possible to solve the linear programming relaxation of P without having to enumerate all the routes. This is done by enumerating a portion of all possible routes, and solving the resulting linear programming relaxation with this partial route set. The solution to the linear program is then used to determine if there are any routes not included which can reduce the solution value. This is the *column generation* step. Using the values of the optimal dual variables (with respect to the partial route set), we generate a new route and resolve the linear programming relaxation of the set covering problem. This is repeated until one can show that an optimal solution to the linear program is found; one that is optimal for the complete route set. Finally, to get an integer solution to the set covering problem, the linear program is combined in a branch and bound routine.

It is well known that a branch and bound strategy works well only if the lower bound used in the bounding step is very tight. Fortunately, many researchers have reported that the linear programming relaxation of the set covering problem provides a solution close to the optimal integer solution, see e.g. Desrochers, Desrosiers and Solomon. That is, the solution to the linear programming relaxation of P provides a very strong lower bound on the solution to the VRPTW. In this paper we demonstrate why this is true in general.

For this purpose, we perform a probabilistic analysis of the linear programming relaxation of P. Consider the customer locations to be distributed according to a distribution μ with compact support in $\mathbb{R}^2$. Let the customer parameters $\{(w_i, e_i, s_i, l_i) : i \in N\}$ be drawn from a joint distribution Φ with a density ϕ that is Liptshitz continuous of order $q \geq 1$ on $[0,1]^4$. (For $x \notin [0,1]^4$, $\phi(x) \equiv 0$.) A function ϕ is Liptshitz continuous of order q on $S \subseteq \mathbb{R}^4$ if there exists a K such that

$$|\phi(x) - \phi(y)| \leq K|x - y|^q, \ \forall x, y \in S.$$

This implies that there exists a K_0 such that $\phi(x) \leq K_0$, for all $x \in [0,1]^4$. Note that the support of ϕ is a subset of $\mathcal{C} \equiv \{(w,e,s,l) \in [0,1]^4 : e+s \leq l\}$. Each customer is therefore represented by its location in the Euclidean plane along with a point in $\mathcal{C}$. Finally, we assume that a customer's location and its parameters are *independent* of each other.

Our main theorem is the following:

Theorem 1.1 *Let the customer locations be independently and identically distributed according to a distribution μ with compact support in $\mathbb{R}^2$. Let the customer parameters be independently and identically distributed like Φ with support in $[0,1]^4$. Let Z^{LP} be the value of the optimal fractional solution to P, and let Z^* be the value of the optimal integer solution to P; that is, the value of the optimal solution to the VRPTW. Then,*

$$\lim_{n\to\infty} \frac{1}{n} Z^{\mathrm{LP}} = \lim_{n\to\infty} \frac{1}{n} Z^* \quad \text{(a.s.)}.$$

It is interesting to note that the value of these limits is also known. Bramel and Simchi-Levi (1993) show that as the number of customers increases, the quantity Z^*/n tends almost surely to $2\gamma E(d)$, where $E(d)$ is simply the customer's expected distance to the depot and γ is a constant depending on Φ, the joint distribution of the loads, time windows and service times. As pointed out by one of the referees, even though the relative error between the fractional and integer solutions to P decreases to zero as the number of customers increases, the *absolute* error may well increase. We will see through the analysis that this absolute error is in fact $O(n^{8/9})$.

We also note that a similar result holds for the classical Capacitated Vehicle Routing Problem, see Simchi-Levi (1996). Here, service times and time windows are not present. Again, the relative error between fractional and integer optimal solution values to the set covering problem associated with the routing problem decreases to zero as the number of customers increases; the absolute error in this model is $O(n^{3/4})$. Of course, in that case the proof is much simpler.

Observe that an alternative formulation of problem P is obtained by replacing the inequality constraints by equality constraints. We refer to this as problem *PE* and remark that it is no more than a *set-partitioning* formulation of the VRPTW. Since the distances satisfy the triangle inequality, each customer in the optimal solution to problem P will be visited *exactly* once. It is therefore obvious that the optimal solution values to problems P and *PE* are identical. It is also easy to prove that the linear programming relaxations of P and *PE* have the same optimal solution values. Hence, our results hold for both the set covering and set partitioning formulations.

2 Motivation

To motivate the analysis and to help understand why Theorem 1.1 holds, we first consider the following simpler *discrete* model. Define a *customer type* to be a location $Y \in \mathbb{R}^2$ and customer parameters $(w, e, s, l) \in \mathcal{C}$. That is, two customers of the same type are located at the same location and have identical customer parameters. Consider a *discretized* vehicle routing model in which there is a *finite* number, W, of customer types. Let n_i be the number of customers of type i, for $i = 1, 2, \ldots, W$, and let $n = \sum_{i=1}^{W} n_i$ be the total number of customers. Clearly, this discretized VRPTW can be solved by formulating it as a set covering problem. To obtain some intuition about the linear programming relaxation of P, we introduce another formulation of the VRPTW closely related to P.

Let a *vehicle assignment* be a vector $(a_1, a_2, \ldots, a_W)$, where $a_i \geq 0$ are integers, such that a single vehicle can feasibly serve a_i customers of type i, for each $i = 1, 2, \ldots, W$, without violating the load or time window constraints. Index all the possible vehicle assignments $1, 2, \ldots, R'$ and let c_r be the total length of the shortest feasible route serving the customers in vehicle assignment r. (Note R' is independent of n.) The VRPTW can be formulated as follows: let

$$A_{ir} = \text{number of customers of type } i \text{ in vehicle assignment } r,$$

for each $i = 1, 2, \ldots, W$ and $r = 1, 2, \ldots, R'$. Let

$$y_r = \text{number of times vehicle assignment } r \text{ is used in the optimal solution.}$$

Then problem P_D is:

$$\min \Big\{ \sum_{r=1}^{R'} y_r c_r \Big| \sum_{r=1}^{R'} y_r A_{ir} \geq n_i, \forall i = 1, 2, \ldots, W, \quad y_r \in \{0, 1\}, \forall r = 1, 2, \ldots, R' \Big\}.$$

Let Z_D^* be the optimal solution value of P_D and let Z_D^{LP} be the optimal solution to its linear relaxation. Clearly, problems P and P_D have the same optimal solution values, i.e., $Z^* = Z_D^*$. Observe that a feasible solution to the linear programming relaxation of P can be used to construct a feasible solution to the linear programming relaxation of P_D, and therefore

$$Z_D^{\mathrm{LP}} \le Z^{\mathrm{LP}}. \tag{1}$$

Define $\overline{c} = \max_{r=1,2,\ldots,R'}\{c_r\}$, i.e., $\overline{c}$ is the length of the longest route among the R' vehicle assignments. Then,

Lemma 2.1

$$Z^{\mathrm{LP}} \le Z^* \le Z_D^{\mathrm{LP}} + W\overline{c} \le Z^{\mathrm{LP}} + W\overline{c}.$$

Proof. The left most inequality is trivial while the right most inequality is due to equation (1). To prove the central inequality note that in P_D there are W constraints (one for each customer type). Let $\overline{y}_r$, for $r = 1, 2, \ldots, R'$, be an optimal solution to the linear programming relaxation of P_D and observe that there exists such an optimal solution with at most W positive variables; one for each constraint. We construct a feasible solution to P_D by rounding the linear programming solution up; that is, for each $r = 1, 2, \ldots, R'$ with $\overline{y}_r > 0$ we make $y_r = 1$ and for each $r = 1, 2, \ldots, R'$ with $\overline{y}_r = 0$ we make $y_r = 0$. The increase in the objective function is therefore at most W times the largest possible cost of a route, $\overline{c}$. ∎

Observe that the upper bound on Z^* obtained in Lemma 2.1 consists of two terms. The first, Z^{LP}, is a lower bound on Z^*, which clearly grows with the number of customers, n. The second term ($W\overline{c}$) is the product of two numbers that are fixed and independent of n. Therefore, the upper bound on Z^* of Lemma 2.1 is dominated by Z^{LP} and consequently we see that for large n, $Z^* \approx Z^{\mathrm{LP}}$, exactly what is implied by Theorem 1.1. Indeed, much of the proof of the following section is concerned with approximating the distributions μ and Φ with discrete distributions and forcing the number of different customer types to be independent of n.

3 Proof of Theorem 1.1

It is clear that $Z^{\mathrm{LP}} \le Z^*$ and therefore, almost surely, $\underline{\lim}_{n\to\infty} \frac{1}{n}(Z^* - Z^{\mathrm{LP}}) \ge 0$. The interesting part is to find an upper bound on Z^* that involves Z^{LP} and to show that $\overline{\lim}_{n\to\infty} \frac{1}{n}(Z^* - Z^{\mathrm{LP}}) \le 0$, almost surely. We do this in essentially the same way as in Section 2. We successively discretize the problem by introducing a sequence of routing problems whose optimal solutions are "relatively" close to Z^*. The last problem is a discrete problem which therefore, as in Section 2, can be directly related to the linear programming relaxation of its set covering formulation. This linear program is also shown to have an optimal solution "close" to Z^{LP}.

To prove the upper bound, let N be the index set of customers, with $|N| = n$. Let A be the compact support of the distribution of the customer locations (μ), and define $d_{\max} = \sup\{\|y\| : y \in A\}$. Pick a *fixed* $k > 1$, and let $N_k \subseteq N$ be the subset of customers with demand less than $1/k$, that is, $N_k \equiv \{i \in N : w_i < 1/k\}$. The vehicle routing problem P_k is defined on only the customers in $N \setminus N_k$. Let Z_k^* be the optimal solution to P_k. Then, it is clear that

$$Z^* \leq Z_k^* + 2d_{\max}|N_k|. \tag{2}$$

Discretization of the Locations

We start by constructing a VRPTW with discrete locations. Define $\Delta \equiv 1/k^2$ and let $G(\Delta)$ be an infinite grid of squares of *diagonal* Δ, that is, of side $\Delta/\sqrt{2}$, with edges parallel to the system coordinates. Let $A_1, A_2, \ldots, A_{m(\Delta)}$ be the subregions of $G(\Delta)$ that intersect A and have $\mu(A_i) > 0$. Since A is bounded, $m(\Delta)$ is finite for each $\Delta > 0$. For convenience, we omit the dependence of m on Δ in the notation. For each subregion, define the *centroid* of subregion A_i as the point at the center of the grid square containing A_i. This defines m centroids and note that a customer is at most $\Delta/2$ units from the centroid of the subregion it is located in.

We define an operation on the set of parameters called *shifting* the parameters. Shifting customer i's parameters (w_i, e_i, s_i, l_i) results in the parameters $(w_i, e_i - \Delta/2, s_i + \Delta, l_i + \Delta/2)$. Let

$$N^0 = \{i \in N \setminus N_k : e_i < \Delta/2\} \cup \{i \in N \setminus N_k : s_i > 1 - \Delta\} \cup \{i \in N \setminus N_k : l_i > 1 - \Delta/2\},$$

that is, N^0 are the customers of $N \setminus N_k$ who's shifted parameters are no longer in $[0,1]^4$.

Construct a new VRPTW, called $P_k(m)$, defined only on the customers of $N \setminus (N_k \cup N^0)$. Each of the customers in $N \setminus (N_k \cup N^0)$ has its parameters shifted as above. In addition, each customer in $N \setminus (N_k \cup N^0)$ is moved to the centroid of the subregion it is located in. Let $Z_k^*(m)$ be the optimal solution to $P_k(m)$. Then,

Lemma 3.1

$$Z_k^* \leq Z_k^*(m) + n\Delta + 2d_{\max}|N^0|.$$

Proof. We construct a feasible solution to P_k using an optimal solution to $P_k(m)$ and by serving the customers of N^0 individually. To serve the customers in $N \setminus (N_k \cup N^0)$ note that every feasible tour in $P_k(m)$ can be modified so that after the vehicle arrives at a centroid, it travels to and from each customer in the tour that is in the subregion. Repeating this for each route cannot increase the total distance traveled by more than $n\Delta$. This solution is feasible for P_k, since the extra Δ units of service time in $P_k(m)$ enable the vehicles to travel to and from the centroid and still serve each customer in its time window. ∎

From here on we assume that all customers in $N \setminus (N_k \cup N^0)$ have parameters that are shifted as above. To simplify notation, we use the following conventions:

$$e_i' \equiv e_i - \frac{\Delta}{2}, \quad s_i' \equiv s_i + \Delta, \text{ and } l_i' \equiv l_i + \frac{\Delta}{2}.$$

Discretization of the Customer Parameters

We now describe a VRPTW where the customer parameters are also discretized. Let $\epsilon \equiv 1/k$ and divide each of the four coordinate axes of $[-2\epsilon, 1+\epsilon]$ into intervals of size ϵ. This produces $(k+3)^4$ four-dimensional *hypercubes* of equal dimension $(\epsilon\times\epsilon\times\epsilon\times\epsilon)$. This also creates $I \equiv (k+1)^4$ points in $[0,1]^4$ which correspond to the corners of the hypercubes. We use the term *corners* to describe these points. Corner i has coordinates (w^i, e^i, s^i, l^i), for each $i = 1, 2, \ldots, I$.

We now describe a method to construct a fully discretized VRPTW, one in which there are a *fixed* number of customer locations and customer parameters. We refer to each centroid–corner pair as a customer type. Each centroid defines a customer location and each corner defines exactly the four customer parameters. It is clear that there are mI possible customer types. An instance of a fully discretized VRPTW is then defined by specifying the number of customers of each of the mI types.

For each corner $i = 1, 2, \ldots, I$ (at coordinates (w^i, e^i, s^i, l^i)), let $\mathcal{E}_i$ be the union of the 16 hypercubes immediately adjacent to (surrounding) the corner and the 16 hypercubes that are adjacent to the corner situated ϵ units directly below corner i in the l–dimension. (Note this is a total of 24 hypercubes since the middle 8 are counted twice.) That is,

$$\mathcal{E}_i = \Big\{(w,e,s,l) \in [w^i-\epsilon, w^i+\epsilon] \times [e^i-\epsilon, e^i+\epsilon] \times [s^i-\epsilon, s^i+\epsilon] \times [l^i-2\epsilon, l^i+\epsilon]\Big\}.$$

For each $i = 1,2,\ldots,I$ and $j = 1,2,\ldots,m$, let

$$E_{ij} \equiv \Big\{h \in N \setminus (N_k \cup N^0) : (w_h, e'_h, s'_h, l'_h) \in \mathcal{E}_i \text{ and } Y_h \in A_j\Big\}.$$

The customers in the set E_{ij} can be partitioned into 24 subsets, each one corresponding to the set of customers in one of the 24 hypercubes in the definition of $\mathcal{E}_i$. Let n_{ij} be the number of customers in the smallest (in cardinality) of these 24 subsets.

We now define a fully discretized VRPTW called $\hat{P}_k(m)$, whose optimal solution value is denoted $\hat{Z}^*_k(m)$. The vehicle routing problem $\hat{P}_k(m)$ has n_{ij} customers located at centroid j with customer parameters defined by the coordinates of corner i, for each $i = 1,2,\ldots,I$ and $j = 1,2,\ldots,m$.

For each corner $i = 1,2,\ldots,I$, define the single hypercube $\mathcal{R}_i \subset \mathcal{E}_i$ adjacent to corner i as:

$$\mathcal{R}_i \equiv \Big\{(w,e,s,l) \in [w^i-\epsilon, w^i] \times [e^i-\epsilon, e^i] \times [s^i-\epsilon, s^i] \times [l^i, l^i+\epsilon]\Big\}.$$

Also define, for each $i = 1,2,\ldots,I$ and $j = 1,2,\ldots,m$,

$$R_{ij} \equiv \Big\{h \in N \setminus (N_k \cup N^0) : (w_h, e'_h, s'_h, l'_h) \in \mathcal{R}_i \text{ and } Y_h \in A_j\Big\},$$

and $r_{ij} \equiv |R_{ij}|$. We have the following:

Lemma 3.2

$$Z^*_k(m) \leq \hat{Z}^*_k(m) + 2d_{\max} \sum_{j=1}^{m} \sum_{i=1}^{I} (r_{ij} - n_{ij}).$$

Proof. Observe:

(i) For each $i = 1,2,\ldots,I$ and $j = 1,2,\ldots,m$, $r_{ij} \geq n_{ij}$. This follows since $\mathcal{R}_i \subset \mathcal{E}_i$ and n_{ij} is the smallest number of customers among the hypercubes in $\mathcal{E}_i$.

(ii) In $P_k(m)$, the customers in R_{ij} have smaller loads, smaller earliest time for service, smaller service times and later due dates than the customers of $\hat{P}_k(m)$ at centroid j with parameters defined by corner i.

Given an optimal solution to $\hat{P}_k(m)$, we construct a solution to $P_k(m)$. For each corner $i = 1, 2, \ldots, I$ and for each $j = 1, 2, \ldots, m$, we pick any $r_{ij} - n_{ij} \geq 0$ customers from R_{ij} and serve them in individual vehicles. Because of (ii), the remaining n_{ij} customers in R_{ij} can be served with exactly the same vehicle schedules as in $\hat{P}_k(m)$. ∎

We now construct another VRPTW. Define $N^1 \equiv \{i \in N \setminus (N_k \cup N^0) : l'_i > 1 - \epsilon\}$, that is, N^1 is a particular set of customers near the boundary of $\mathcal{C}$. For each customer $i \in N \setminus (N_k \cup N^0 \cup N^1)$, define a new set of parameters as follows:

$$w''_i \leftarrow \epsilon \lfloor w_i/\epsilon \rfloor, \quad e''_i \leftarrow \epsilon \lfloor e'_i/\epsilon \rfloor, \quad s''_i \leftarrow \epsilon \lfloor s'_i/\epsilon \rfloor, \text{ and } l''_i \leftarrow \epsilon \lceil l'_i/\epsilon \rceil + \epsilon.$$

The fully discretized VRPTW, called $P_{k,\epsilon}(m)$, is defined only on the customers of $N \setminus (N_k \cup N^0 \cup N^1)$ where each customer is moved to its centroid and the customer parameters of $N \setminus (N_k \cup N^0 \cup N^1)$ are $(w''_i, e''_i, s''_i, l''_i)$. Let $Z^*_{k,\epsilon}(m)$ be its optimal solution value. Then,

Lemma 3.3

$$\hat{Z}^*_k(m) \leq Z^*_{k,\epsilon}(m) + 2d_{\max}|N^1|.$$

Proof. Note that in the definition of $P_{k,\epsilon}(m)$, the number of customers at corner i and at centroid j is equal to the number of customers in one particular hypercube of the 24 hypercubes "surrounding" corner i. Therefore, $P_{k,\epsilon}(m)$ has at least as many customers of each type as does $P_k(m)$. The customers in N^1 are served individually. ∎

Now $P_{k,\epsilon}(m)$ is a discrete problem and we can analyze it as in Section 2. Let $Z^{\text{LP}}_{k,\epsilon}(m)$ be the optimal solution value of the linear programming relaxation of the set covering formulation of $P_{k,\epsilon}(m)$. Let $\overline{c}$ be defined as in Section 2, that is, it is the cost of the most expensive route in $P_{k,\epsilon}(m)$. A straightforward application of Lemma 2.1 gives the following.

Lemma 3.4

$$Z^*_{k,\epsilon}(m) \leq Z^{\text{LP}}_{k,\epsilon}(m) + mI\overline{c}.$$

Let Z^{LP}_k be the optimal solution value of the linear programming relaxation of the set covering formulation of P_k. Define the following set of customers:

$$T(k,\epsilon) \equiv \Big\{h \in N \setminus (N_k \cup N^0 \cup N^1) : s_h \in \cup_{i=1,2,\ldots,k}[i/k - \Delta, i/k]\Big\}.$$

Then,

Lemma 3.5

$$Z^{\mathrm{LP}}_{k,\epsilon}(m) \le Z^{\mathrm{LP}}_k + n\Delta + 2d_{\max}|T(k,\epsilon)|.$$

Proof. Let $\{\overline{y}_r : r = 1,2,\ldots,R\}$ be an optimal solution to the linear programming relaxation of the set covering formulation of P_k. We can assume $\sum_{r=1}^{R} \overline{y}_r \alpha_{ir} = 1$, for each $i = 1,2,\ldots,n$. We construct a feasible solution to the linear programming relaxation of the set covering formulation of $P_{k,\epsilon}(m)$ using the values $\overline{y}_r$. For each route r with $\overline{y}_r > 0$ we do the following. Delete all customers in $T(k,\epsilon)$, N^0 and N^1 from the route. Construct a route r' feasible for $P_{k,\epsilon}(m)$ that serves the remaining customers of route r. Since in $P_{k,\epsilon}(m)$ the customers are at the centroids instead of at their original locations, we modify the route so that the vehicle travels from the customer to its centroid and back. Thus, the length (cost) of route r' is at most the cost of route r in P_k plus $n_r\Delta$ where n_r is the number of customers in route r. To show that route r' is a feasible route for $P_{k,\epsilon}(m)$, note that the customers in $P_{k,\epsilon}(m)$ all have smaller demands and smaller earliest time for service than their corresponding customer in P_k. Hence, the demands of the customers in route r' can be feasibly packed and the vehicle arrives at each customer no *later* than its earliest time for service. Furthermore, since the customers of $T(k,\epsilon)$ are taken out of route r', each customer in $P_{k,\epsilon}(m)$ also has a smaller service time than its service time in P_k. This is true since each customer h not in $T(k,\epsilon)$ has a service time equal to

$$\epsilon\lfloor s'_h/\epsilon\rfloor = \epsilon\lfloor (s_h + \Delta)/\epsilon\rfloor \le s_h.$$

Finally, compare the time service is *finished* at a customer in route r and route r'. Observe that the difference between the time service is finished at a customer in route r' and route r is at most $\Delta/2$ for the first customer visited in r', at most $3\Delta/2$ for the second customer, and at most $(i-\frac{1}{2})\Delta$ for the ith customer in the route. Since at most k customers can be in any route of P_k, the difference between the time service is finished in route r' and route r for the last customer in the route is no more than $(k - \frac{1}{2})\Delta < \epsilon$. Consequently, the time service is finished for a customer in route r' is less than its latest time service can end. Therefore route r' is a feasible route in $P_{k,\epsilon}(m)$.

Using these routes and serving customers of $T(k,\epsilon)$ individually, we get

$$Z_{k,\epsilon}^{\mathrm{LP}}(m) \leq Z_k^{\mathrm{LP}} + \sum_{r=1}^{R} \overline{y}_r n_r \Delta + 2d_{\max}|T(k,\epsilon)| \leq Z_k^{\mathrm{LP}} + n\Delta + 2d_{\max}|T(k,\epsilon)|.$$

∎

Since taking customers out of a solution (e.g. N_k) cannot increase the cost of that solution, we clearly have

$$Z_k^{\mathrm{LP}} \leq Z^{\mathrm{LP}}. \tag{3}$$

We can now prove the main result. Through consecutive application of equation (2), Lemmas 3.1–3.5 and equation (3), we have:

$$Z^* \leq Z^{\mathrm{LP}} + mI\overline{c} + 2d_{\max} \sum_{j=1}^{m} \sum_{i=1}^{I} (r_{ij} - n_{ij}) + 2n\Delta + 2d_{\max}(|N_k| + |N^0| + |N^1| + |T(k,\epsilon)|).$$

Thus, to complete the proof of Theorem 1.1, we need to show that Z^{LP} is the dominant part of this upper bound. We start with the following lemma whose detailed proof can be found in Bramel and Simchi-Levi (1994).

Lemma 3.6 *There exists a constant K_1 such that, almost surely,*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \Big[\sum_{j=1}^{m} \sum_{i=1}^{I} (r_{ij} - n_{ij}) + |N^0| + |N^1| + |T(k,\epsilon)| \Big] \leq \frac{K_1}{k}.$$

Proof. To bound the first term, we note that r_{ij} and n_{ij} both define a number of customers in hypercubes that are relatively close to one another. By Liptshitz continuity and the Strong Law of Large Numbers, a standard probability theory argument shows that $\overline{\lim}_{n\to\infty} \frac{1}{n} \sum_{j=1}^{m} \sum_{i=1}^{I} (r_{ij} - n_{ij}) = O(1/k)$. The other terms can also be bounded because they each represent the number of customers whose parameters occupy regions of $[0,1]^4$ that have very small measure (e.g., measures that are $O(1/k)$). ∎

Finally observe that since each tour in $P_{k,\epsilon}(m)$ can serve no more than k customers, $\overline{c}$ can be no larger than $k+1$ times the length of the diagonal of the compact region A. In fact, the length of any tour can be no more than $1+2d_{\max}$ due to the length of the working day (i.e., 1). Hence, $m\epsilon^{-4}\overline{c} = O(k^8)$, and therefore, almost surely,

$$\overline{\lim_{n\to\infty}} \frac{1}{n}(Z^* - Z^{\mathrm{LP}}) \leq 2d_{\max}K_1/k + 2\Delta + 2d_{\max}\phi([0,1/k])$$

$$\leq 2d_{\max}K_1/k + 2/k^2 + 2d_{\max}K_0/k,$$

where K_0 is defined in Section 1. Since K_1 and K_0 are constants and k was arbitrary, we see that the right-hand side can be made arbitrarily small thus completing the proof of Theorem 1.1. The careful reader will observe that our analysis also implies that the rate of convergence is $O(n^{8/9})$. That is, $E[Z^*] = E[Z^{\mathrm{LP}}] + O(n^{8/9})$, where $E[X]$ represents the expected value of the random variable X.

4 Column Generation in Practice

At this point, we would like to elaborate on the way a number of researchers (e.g., Desrochers, Desrosiers and Solomon (1992), but by no means all, e.g., Noon (1994)) solve the linear programming relaxation of the set covering problem in practice. In the column generation technique, generating a new column for P is $\mathcal{NP}$-hard. For that reason, problem P is modified in such a way that a new column can be generated in pseudo-polynomial time using a dynamic program.

Problem P is modified by allowing routes to visit the same customer more than once. Therefore the set covering problem that is actually solved in practice is slightly different. This problem, called P_R, is defined as follows. Enumerate all feasible routes, satisfying capacity and time window constraints, that may visit the same customer a number of times within its time window (each such visit increases the total load by the demand of that customer). Let the number of routes (columns) be R'', and let c_r be the total distance traveled in route r. For each $i = 1, 2, \ldots, n$ and $r = 1, 2, \ldots, R''$, let

$$\xi_{ir} = \text{number of times customer } i \text{ is visited in route } r,$$

and, for each $r = 1, 2, \ldots, R''$, define

$$y_r = \begin{cases} 1, & \text{if route } r \text{ is in the optimal solution} \\ 0, & \text{otherwise.} \end{cases}$$

Problem P_R is:

$$\min\Big\{\sum_{r=1}^{R''} y_r c_r \Big| \sum_{r=1}^{R''} y_r \xi_{ir} \geq 1, \forall i = 1, 2, \ldots, n, \quad y_r \in \{0, 1\}, \forall r = 1, 2, \ldots, R''\Big\}.$$

Clearly, the integer optimal solution to P_R is still the optimal solution to the VRPTW. However, the optimal solution values of the linear programming relaxations of P_R and P may be different. Of course, the linear programming relaxation of P_R is a lower bound on the linear programming relaxation of P. Since the proof of Theorem 1.1 can be essentially repeated with P replaced by P_R, we conclude that Theorem 1.1 holds for P_R as well.

5 Other Formulations of the VRPTW

In this section we show that other formulations of the VRPTW may have diverging linear and integer solutions. For this purpose consider the following formulation of the VRPTW.

Given a subset of the customers S, let $b(S)$ be the minimum number of vehicles required to carry the demands of customers in S, i.e., $b(S)$ is the solution to the bin-packing problem defined on the demands of all customers in S. For $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$, let

$$x_{ij} = \begin{cases} 1, & \text{if a vehicle travels directly between points } i \text{ and } j, \\ 0, & \text{otherwise.} \end{cases}$$

Let 0 denote the depot and define c_{ij} as the cost of traveling directly between points i and j, for $i, j = 0, 1, 2, \ldots, n$. Let t_i represent the time a vehicle arrives at the location of customer i and for every i and j, such that $i < j$, define $M_{ij} = \max\{l_i + d_{ij} - e_j, 0\}$ where $d_{ij} \equiv \|Y_i - Y_j\|$. Then, the following is a valid formulation of the VRPTW:

$$\text{Problem } P' : \quad Min \qquad \sum_{i<j} c_{ij} x_{ij}$$

$$s.t. \qquad \sum_{i<j} x_{ij} + \sum_{i>j} x_{ji} = 2, \quad \forall i = 1,2,\ldots,n,$$

$$\sum_{i,j\in S} x_{ij} \le |S| - b(S), \quad \forall S \subset \{1,2,\ldots,n\}, 2 \le |S| \le n-1,$$

$$e_i \le t_i \le l_i - s_i, \quad 1 \le i \le n,$$

$$t_i + s_i + d_{ij} - t_j \le M_{ij}(1 - x_{ij}), \quad 1 \le i < j \le n,$$

$$x_{ij} \in \{0,1\}, \quad 1 \le i < j \le n, \tag{4}$$

$$x_{0j} \in \{0,1,2\}, \quad j = 1,2,\ldots,n. \tag{5}$$

The case $x_{0j} = 2$ corresponds to a vehicle serving only customer j. The linear programming relaxation of P' is obtained by replacing constraints (4) and (5) by their linear equivalents.

Consider now an instance of the VRPTW in which vehicles of capacity $Q \ge 2$, serve n customers each having a demand $w_i = 1/Q$, and parameters $e_i = s_i = 0$ and $l_i = 1$ for all $i = 1,2,\ldots,n$. In addition, assume all customers are located at a single point, 1/2 units distance from the depot, and that $c_{ij} = d_{ij}$ for all i and j. For simplicity, we assume n is a multiple of Q. One can check to see that in this example, fractional and integer solutions to the above linear program, almost surely, do not approach the same value asymptotically, and in fact the relative difference can be as large as 50%. The interested reader can consult Bramel and Simchi-Levi (1994) for more details.

6 References

Balinski, M. L. and R. E. Quandt (1964), On an Integer Program for a Delivery Problem. *Operations Research* **12**, pp. 300-304.

Bramel, J. and D. Simchi-Levi (1993), Probabilistic Analyses and Practical Algorithms for the Vehicle Routing Problem with Time Windows. To appear in *Operations Research.*

Bramel, J. and D. Simchi-Levi (1994), The Set Partitioning Model and the Vehicle Routing Problem. Working Paper, Columbia University.

Bramel, J., C. L. Li and D. Simchi-Levi (1994) Probabilistic Analysis of the Vehicle Routing Problem with Time Windows. *American Journal of Mathematical and Management Sciences* **13**, pp. 267-322.

Cullen, F. , J. Jarvis and D. Ratliff (1981), Set Partitioning Based Heuristics for Interactive Routing. *Networks* **11**, pp. 125–144.

Desrochers, M., J. Desrosiers and M. Solomon (1992), A New Optimization Algorithm for the Vehicle Routing Problem with Time Windows. *Operations Research* **40**, pp. 342–354.

Federgruen, A. and G. van Ryzin (1992), Probabilistic Analysis of a Generalized Bin Packing Problem with Applications to Vehicle Routing and Scheduling Problems. Working Paper, Columbia Graduate School of Business.

Fisher, M. L. (1992), Vehicle Routing. To appear in *Handbooks in Operations Research and Management Science*, the volume on Networks, M. Ball, T. Magnanti, C. Monma and G. Nemhauser, eds.

Noon, C. (1994), Lower Bounds for the VRP via TSSP+1 Decomposition. Working Paper, The University of Tennessee.

Simchi-Levi, D. (1996), *The Logic of Logistics*, to be published by Springer-Verlag, New York.

Solomon, M. M. (1986), On the Worst-Case Performance of Some Heuristics for the Vehicle Routing and Scheduling Problem with Time Window Constraints. *Networks* **16**, pp. 161–174.

Solomon, M. M. and J. Desrosiers (1988), Time Window Constrained Routing and Scheduling Problems. *Transportation Science* **22**, pp. 1–13.