1. Background

The **bin packing problem** is a well-studied combinatorial NP-hard problem. In the bin packing problem, a sequence of items are packed into bins, the packed items in a bin do not overlap. The objective is to minimize the number of used bins. The size of each item should be no more than the size of a bin. (Zhang, 2011) The algorithms can be divided into two categories: offline algorithm, requiring that the whole series of items be given at the beginning, and online algorithm which is the opposite. Asymptotic approximation/competitive ratio is the ratio of algorithm result over optimal result, used for evaluation. There are heuristics, such as first-fit (FF: placing into the lowest indexed bin which it will fit), best-fit (BF), first-fit decreasing (FFD) and best-fit decreasing (BFD). (Coffman, 1976) The problem has several variations, and 2-dimensional bin packing (2BPP) is of our interest, as in real-life transportation vehicles have not only weight constraints but limitations on loading space as well.

The **vehicle routing problem** (VRP) is another computationally complex problem concerned with delivering cargo located at a depot to a number of customers on various locations on a graph. The goal is to design routes (on the edges of the graph) for vehicles such that a minimum total distance is traveled, while satisfying all the orders from customers. Again the problem has variations, and we are interested in VRP with LIFO loading, always unloading the last loaded item so that we do not require everything being unloaded and loaded (re-arranged) again at each customer’s place.

This project is a combination of both problems: **two-dimensional loading vehicle routing problem** (2L-CVRP). At the depot, there are \( K \) identical vehicles, each with a two-dimensional loading surface and a maximum weight capacity, and there are \( n \) customers, each with a given demand consisting in a set of rectangular two-dimensional weighted items. The aim is to find a partition of the customers into routes of minimum total cost (total edge weight or distance summed over all vehicles). This problem has obvious utility in trans-
portation and logistics. An exact approach is given (Iori, Salazar-González, & Vigo, 2007), based on branch-and-cut linear programming; tabu search, meta-methodology guided tabu search (GTS) and guided local search (GLS) algorithms were also put forward to address the small scale limitation in the exact approach. A recent paper summarized these approaches and proposed a meta-heuristic algorithm with extended guided local search (EGLS), and achieved satisfactory performance (Leung, Zhou, Zhang, & Zheng, 2011). In spite of good performance in such a hard problem, there is not much research on the theoretical bounds on the worst case performance of the algorithms in the literature. Thus the project is devoted to an analysis on the problem from certain simplifications to harder issues, and hopefully deriving a bound.

2. Defining the problem

The segregation of the two sub-problems – the bin packing and vehicle routing is an obstacle in analyzing the overall performance. In heuristic algorithms, a stochastic initial feasible solution is first produced, and then by swapping we can approximate the optimal. However, this renders the two sub-problems essentially unrelated, except the heuristic in assigning weights to customer orders when generating the initial solution, and a few adjustments. To focus on the integration, I would first consider simplification on both sub-problems, leaving out the core issue.

Basic simplification. Simplify the map to a one-dimensional axis, and consider the weight constraint only (one-dimensional bin packing).

![Figure 1](depot_image.png)

There is no space issue, thus no LIFO constraints. I feel that constraint on number of vehicles is not important, as the algorithm must not use vehicles in a wasteful fashion in order to minimize the total distanced traveled. We also relax the constraint (according to some literatures) that the goods for a customer must be packed in one vehicle, so that we can get rid of the concept of “customers” and consider only two arrays: $w_i \in (0,1]$ is the weight of each item, and $d_i \in \mathbb{Z}^+$ is the distance from the depot to its destination (Figure 1). Each vehicle has weight capacity 1. Suppose the solution has $K$ vehicles $v_1, v_2, ..., v_K$, the travel distance of each of them is determined by the farthest item in the vehicle $D_j = \max \{d_i, \text{item } i \in v_j \}$. In reality this is round-trip, but since it is equivalent in minimization, using one-way distance is sufficient. The goal is to minimize $\sum_{v_j} D_j$. 

2
The paragraph above gives a well-defined problem, keeping only the essential integration of the two problems. I would like to further explore these variants:

**Introducing two-dimensional bin packing.** I may either add the space constraint that each vehicle and each item has a width and a length, and items are put in the vehicle in axis-oriented direction, not exceeding the loading space of the vehicle; or simply replace the weight constraint by the space constraint. LIFO constraints may also be added. This step can be difficult as I need to devise certain heuristic and/or randomized approaches that are easy to analyze.

**Grid as intermediate step.** In urban areas, the customers are actually densely connected by horizontal and vertical roads on a map, forming a grid to travel on. Consider a plane with $xOy$ coordinates, where the depot is placed at the origin and the position of a customer can be described by $P(x, y), x, y \in \mathbb{Z}$. The edge cost between two points in a graph is replaced by the sum of differences in the two coordinates. Thus it serves as a simplification of graph where essentially every two vertices are connected by an edge. It is one step further than the basic simplification, since intuitively we are not only dealing with absolute distance, but directions as well – as if there were many axes pointing towards different directions.

**Putting together.** The whole problem may be transplanted back to a graph-based routing, or using the grid is good enough.

### 3. Methodology

Bin packing problems are usually bounded by mathematical proofs which use numerical worst case examples. For instance, in proving the $17/10$ bound of first fit and best fit for one-dimensional bin packing (Coffman, 1976), they made a series of items composed of three equal regions, and the items in each of the regions are at the size around $1/6, 1/3$ and $1/2$ respectively. Then they argued that this construction would be the worst possible for the algorithm, and the optimal solution was readily there for comparison. For 2L-CVRP, linear programming is widely used as exact approaches.

There are quite a few difficulties in applying the same methods to this problem. First, so far I have no idea how to construct such an example. Second, in our problem not only the weight is considered, but intuitively we should pack items of similar distances together in order to reduce the number of vehicles traveling to far places, giving rise to a “categorization” of the items. Taking steps further we would involve more constraints that can affect the optimal packing decision.
Consider a greedy algorithm for the basic case that scans from the farthest customer to the nearest one, and groups items so that total weight is just more than 1 (the capacity of the vehicle) into groups \( g_1, g_2, ..., g_k \), where \( g_1 \) is the farthest group and the total weight of \( g_k \) is less than or equal to 1 (the collection of remaining items). Then we can use two vehicles for each group, using \( 2k \) vehicles in total. On the other hand, the total weight of all items requires at least \( k \) vehicles. A coarse upper bound of 2 is determined in terms of the number of vehicles (not our optimization target though).

Since the cost of a vehicle is determined by the farthest item, the other items on the vehicle are indeed transported “for free”. A naïve heuristic for the basic simplification would sort the items in descending order of their distances, and do a first fit, thus reduces the number of vehicles that travels long distances. That is, we categorize the items by their distances. In the grid case, we group the items not only by distances, but by directions as well, resembling the drivers’ common sense of “in one’s way” (Figure 2).

Notice that Figure 2 resembles clustering in spatial data mining, which I learned in the CSIS 323 course. Several clustering algorithms (Ng & Han, 1994) are based on picking “medoids” as bases of clusters, computing dissimilarities and swapping medoids – exploring neighboring graphs. Similar techniques can be employed in initial grouping of our items (reducing from multi-dimensions to two-dimensional space), but although there are performance reports, there is not a worst case bound of this method.

Therefore in the end it seems that I should try some greedy algorithms for the basic case. Then, when moving to multiple constraints, like loading space, LIFO, routing on grid or even routing on graphs, I will need some randomized algorithm techniques, e.g. with the distance of each vehicle as a random variable, and possibilities assigned to items entering a certain vehicle (the hard part), measure concentration can be applied to analyze the (expected) sum of travel distances. When introducing LIFO, the method of computing reshuffling time (Ladany & Mehrez, 1984) can be referred.

4. Project schedule

Sep 2011 – Nov 2011 (Preparation). In this stage I will take courses on randomized algorithm (graduate-level course) and linear programming (from the math department), as mathematical foundations for the project. I will also look into literatures about different
approaches. But as I will be extremely busy for course work, an “entrepreneur” project, and application for graduate schools, the emphasis is on the second semester.

Dec 2011 – Jan 2012 (Foundation). Before the second semester I should have a preliminary report and be prepared for the first presentation. Specifically, I should finish analyzing the basic case in this stage.

Feb 2012 – Apr 2012 (Extensions). In this stage, I should focus on exploring all the variations I have mentioned above. As course workload is expected to be reduced (as compared to 2011), I should try out as many approaches as possible. Due to the nature of this project, I do not allocate time for report writing, and I will note down my discoveries as research goes on, and maintain a finalized report at the same time. I should be prepared for project exhibition.

5. Reference