Package Delivery on "Double 11" Day

Project for Algorithm and Complexity

Xiaofeng Gao, Mingran Peng (Responsible TA)
Department of Computer Science,
Shanghai Jiao Tong University, Shanghai, China

Abstract. This project introduces the concept of package delivery. It presents the issue of delivering a huge number of packages, faced by many online shop managers especially for busy periods like "Double 11" Day. Meanwhile, it lists all the tasks and requirements for this project. Please read this document carefully and complete the tasks.
Keywords: Package Delivery, Hubs

1 Package Delivery Problem

Package delivery is a crucial issue for online shops in terms of reducing transportation costs and winning higher rates of consumers. It turns to be more challenging in some busy periods annually, such as "Black Friday" in U.S. and "Double 11" in China, since the number of packages explodes. In 2017, It was just 12 hours before the number of packages exceeded 100 million during "Double 11" period. Such surge of packages requires online shop managers to refine their delivery schemes, including delivery route for each package, transportation tools, delivery orders for different packages, and so on.

Now suppose that the SF Express asks you to help contrive of delivering schemes for the "Double 11" Day. They have 4 choices of transportation tools: airlines, trains, trucks, and shipping. The SF Express wishes to win high rates from customers, while cuts down the total cost and increases profits as much as possible.

The information you have initially is as follows:

1. There are 656 cities involved in, denoted as 1 to 656 respectively.
2. There are 123 commodities involved in, denoted as 1 to 123 respectively, and they are divided exclusively into 8 categories: Plastic, Metal, Glass, Liquid, Food, Inflammable Products, Electronics, and Big Furniture.
3. Table A provides all the information about the orders during the "Double 11" day. Each row represents the information of one order (totally $247,565$ orders).
4. Table B provides all the information about the commodities.
5. Table C provides the distance between any two cities. A distance matrix is formed in this table, where element (i, j) denotes the distance between city i and city j (Unit: km).
6. Table D provides all the information about the transportation tools. We assume that the schedules for all transportation tools keep unchanged every day.

Note: If the information you need is not provided, you may search online with the sources quoted, or set by yourselves properly, in which case you should clarify why your setting is reasonable.

2 Tasks and Requirements

In this project, there are four tasks as follows, each with connection to others.

1. Suppose that SF Express has its substations on all 656 cities covered in the orders, which means that it can pick up the packages from the sellers and give them to purchasers in the cities they are in directly. Please design an efficient scheme you consider as the best to deliver all these packages, and analyze the delivery performance.
2. Someone in SF Express intends to set some cities to be **hubs** in the whole transport networks. Hubs can gather packages and send them together to the same destinations (not necessarily the ultimate destinations) with lower unit cost. However, the packages gathered together at a hub to one destination can only be sent by one transportation tool. Please design a setting of hubs you consider as the best, and analyze the possible difference it will make to your model, solution, and performance in Problem 1. Note that you may consider the cost of building new hubs by yourselves.
3. In real case, some other constraints should be considered: the hubs may be capacitated; some hubs may not accept some specific packages, such as glass-made or inflammable products; some packages may not be transferred by some transportation tools: liquid and inflammable products cannot be on the plane. There are also other possible constraints you can come up with. Please discuss the possible difference it may happen on your model if these constraints are involved? Does your original scheme need to be changed?
4. In contrast to Problem 1, suppose that the SF Express does not have substations in all cities, but only in big cities. Here we suppose that the big cities are those supporting airline service. This means that the SF Express should first take the packages from sellers to some substations, and when delivering the packages to purchasers, some substations should receive the packages first, and then send them to the city that the purchasers are in. Please analyze what difference it will incur on your model and scheme? Do you need to revise your model?

2.1 Problem Formulation

Please give a clear definition of cost, and quantify the customer's rating of a package, for example how the rating is related to the delivery time. Then please formulate it into a mathematical programming problem. You need to define your variables carefully, clarify all involved parameters clearly, and write down all constraints mathematically. Can you convert your formulation into a linear programming (LP), or even an integer linear programming (ILP)?

2.2 Problem Analysis and Algorithm Design

Firstly, please judge the difficulty of the problems. Whether they are in P, NP, NP-Complete, or NP-Hard? Prove or clarify your conclusion.

Next, please design one or several efficient algorithms to solve the detection problem in **polynomial time** with respect to the input size. You need to describe your design first, introduce the necessary concepts, symbols, definitions, etc., and write the pseudo code of your design.

2.3 Theoretical Analysis, Algotihm Design and Performance Evaluation

For each problem, you need to complete the following tasks:

Theoretical Analysis and Algorithm Design: For this part you are aiming to distinguish the theoretical properties of your problem and algorithm designs, including the following items:

1. Analyze and prove the computational hardness of each problem. (P, NP, NP-Complete, or NP-Hard?)
2. Analyze the time complexity of your algorithms.
3. Show the correctness of your algorithm if you prove your problem to be in P, otherwise analyze the approximation ratio of your algorithm.

Performance Evaluation: it includes the following requirements:

1. For each case, based on the information we have attached (Table A to D), and other possible information you gather from the Internet, give the delivery scheme for each package, analyze the total cost, and calculate the approximated time that each package will arrive at the substation of its purchaser's city.
2. Discuss the relationship between the average rating of the customers and the cost for transportation: Do they have a trade-off relationship that to win higher rating will cost more on transportation? If not, can your delivery scheme achieve both high rating and low cost?

2.4 Report Requirements

You need to submit a report for this project, with the following requirements:

1. Your report should have the title, the author names, IDs, email addresses, the page header, the page numbers, figure for your simulations, tables for discussions and comparisons, with the corresponding figure titles and table titles.
2. Your report is English only, with a clear structure, divided by sections, and may contain organizational architecture like itemizations, definitions, or theorems and proofs.
3. Please include reference section and acknowledgement section. You may also include your feelings, suggestion, and comments in the acknowledgement section.
4. Please define your variables clearly. If needed, a symbol table is strongly recommended to help readers catch your design.
5. Please also include your latex source and simulation codes upon submission.