#### FBE/IE-03/2019-3

# A RICH-VRP MODEL FOR ASSESSING THE PERFORMANCE OF ALTERNATIVE LINEAR OBJECTIVE FUNCTIONS IN EFFECTIVENESS-FAIRNESS TRADE-OFF

### GÖKALP ERBEYOĞLU <sup>1</sup> , ÜMİT BİLGE <sup>1</sup>

Mayıs 2019 May 2019

Fen Bilimleri Enstitüsü Institute for Graduate Studies in Science and Engineering

# Boğaziçi University, Bebek, Istanbul, Turkey

Boğaziçi Araştırmaları deneme

niteliğinde olup, bilimsel tar-

tışmaya katkı amacıyla yayınlan-

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dıklarından, yazar(lar)ın yazılı

izni olmaksızın kendilerine

atıfta bulunulamaz.

Boğaziçi Research papers are of a preliminary nature, circulated to promote scientific discussion and not to be quoted without written permission of the author(s).

<sup>&</sup>lt;sup>1</sup> Industrial Eng. Dept., Boğaziçi University, Istanbul, Turkey

# **TABLE OF CONTENTS**





# **1. INTRODUCTION**

In the thesis proposal, we suggested to study the preparedness and response stages of disaster relief operations while maintaining fairness among beneficiaries as a concern. Fairness in the preparedness stage is incorporated as a way to design a network on which the items can be distributed fairly. For the response stage model, we aim to build a model that enables us to distribute relief items in a timely and fairly manner. We have completed the studies related with the preparedness stage by proposing mathematical models, *CLFDIP* and its variant *CLFDIP\_EarlyNotify*. We suggested three heuristic approaches to solve the problem, namely *Cluster-first*, *Iterative*, and *Simulated Annealing*. Since the performance of the heuristic methods were found to be inadequate for large instances, Benders decomposition based exact methods to solve these models are proposed. Numerical studies indicate the success of the exact solution method.

Aligned with our thesis proposal, we moved on to the response stage of disaster relief operations. As the first step, we studied the literature on the fairness concept to come up with a general research direction, and a basic relief item distribution model is proposed. The minimization of inequality between demand points may lead to inferior solutions since it contradicts the maximization of individual utilities of the beneficiaries. Therefore we need to focus on fair optimization which considers equity, effectiveness and efficiency rather than inequality minimization. In this semester we refined the model and the proposed objective functions that can focus on both fairness and effectiveness. Since off-the-shelf solvers are unable to solve even small instances in a reasonable amount of time, exact solution techniques that could tackle the problem are considered first, then a heuristic solution method is implemented. A data generation scheme is defined and the numerical performance of the model and objectives are studied. Also, the proposed objectives are evaluated using different performance measures in the literature.

The rest of this manuscript is organized as follows: Chapter 2 introduces some

of the studies in inequity averse optimization. Chapter 3 presents the base routing model is used in studies and proposed objective functions that considers effectiveness and fairness in distribution of goods. These two chapters are similar to the previous report, however they are kept to have a complete manuscript. Chapter 4 introduces the proposed heuristic solution method. Finally, Chapter 5 provide numerical results and Chapter 6 concludes the report.

# **2. FAIRNESS CONCERN AND OBJECTIVES USED IN DISASTER RESPONSE**

In the delivery of disaster relief items to the affected population, a fair distribution of goods in terms of amount and timeliness is crucial. It is at least as important as the equity provided in other public services since human lives may be at stake. Many disciplines (including economics, political science, and operations research among others) are interested in the equity concept (Leclerc et al., 2012). In economics literature, several studies on the equity and efficiency trade-off and its applications can be found. One can refer to Sen and Foster (1997) and Young (1995) for books on the subject with the economics point of view. Savas (1978) discuss efficiency, effectiveness and equity as three performance measures in public service deliverance. They define efficiency as how much output is obtained compared to the input, effectiveness as the satisfaction provided by the service relative to the needs, and define equity as "fairness, impartiality, or equality of service". Savas (1978) also elaborates alternative approaches to allocate public services equitably as equal payments, equal outputs, equal inputs, and equal satisfaction of demand. Marsh and Schilling (1994) provide a review on the subject and present 20 different equity measures, then build a framework and analyze their characteristics. Mandell (1991), Bertsimas et al. (2011), and Bertsimas et al. (2012) provide studies that investigate the trade-off between efficiency and fairness. Apart from these, Mulligan (1991) and Kalcsics et al. (2015) provide studies on equity in the facility location problem setting.

Balcik et al. (2010), Ogryczak et al. (2014), Karsu and Morton (2015) and Matl et al. (2016) provide reviews on the fair optimization models. Balcik et al. (2010) focus on nonprofit and public sector with a vehicle routing perspective. Ogryczak et al. (2014) review equity on network models. They present "max-min fairness" and "lexicographic maximin" as equity measures. Then they state minimization of inequality measures may lead to inferior solutions since it contradicts maximization of individual utilities of the beneficiaries. Therefore they express the need to focus on fair optimization which considers both equity and efficiency rather than inequality minimization. They provide a review on objectives that can be used for this purpose, and the fair optimization models in communication networks and location and allocation models. Karsu and Morton (2015) provide a systematic review on inequity averse optimization in operational research. They mainly focus on equitability and balance issues. The main difference between equitability and balance is the anonymity among entities, i.e. equitability compares identical individuals whereas balance concerns with entities with different needs and preferences. They review the allocation, location, vehicle routing, scheduling, transportation and network design, and other studies in terms of equitability concerns. Then they categorize Rawlsian (Rawls, 1971), lexicographic, inequality index based, and inequity-averse aggregation function based approaches as different ways to handle fairness. Matl et al. (2016) focus on equity objectives in vehicle routing perspective. They classified the literature in vehicle routing into five groups. The first two are "vehicle routing problem with route balancing" (VRPRB) and its time window extension. VRPRB models are bi-objective where the second objective minimizes range of distance traveled for equity. The third group, "min-max VRP" contains the models that include a single objective to incorporate equity. Then they define other methods, and application papers as the remaining classes. Matl et al. (2016) also provide theoretical and numerical analysis on inequality measures and their effects on the solutions.

#### **2.1. Approaches to Handle Equity Concerns**

Assuming  $y = (y_1, y_2, ..., y_m)$  is a vector of utilities of beneficiaries (Let  $y_i$  be the outcome of a given distribution for beneficiary  $i \in I$ ,  $I = \{1, 2, ..., m\}$ , the utilitarian solution maximizes  $\sum_{i \in I} y_i$  among feasible solutions. It is a natural choice where the system efficiency is measured by the total utilities. Although in most cases the total system utility reduces under a fairness consideration, the nature of the problem may require a fairness scheme. However there is no universally accepted scheme of equity since its interpretation is subjective and problem specific (Bertsimas et al., 2011). The most widely used approaches to handle equity can be summarized below (Karsu and Morton, 2015; Ogryczak et al., 2014; Matl et al., 2016).

#### **2.1.1. The Rawlsian approach**

It is one of the simplest and most commonly used approaches. It is also referred as "max-min fairness" since it maximizes  $\min_{i \in I} y_i$  to provide equity. However this approach cannot differentiate among distributions as long as they have the same utility level for the most unfavorable entity (e.g. utility vectors  $(1,1,9)$ ) and  $(1,5,5)$  are the same for the Rawlsian approach).

Lexicographic approach is an extension to the Rawlsian approach, which aims to alleviate the the problem of not being able to differentiate the differences among the entities except the worst one of the max-min approach. After maximizing the smallest utility level, it considers the next worst-off entity, and moves on for the next ones.

#### **2.1.2. Inequality index based approach**

Inequality indices are functions that assign a scalar value to a given utility vector. They are usually utilized in multiobjective models since they only consider equitability, and ignore efficiency concerns. The most commonly used inequality indices can be given as follows:

- **Range:** The difference between highest and lowest utilities  $(\max_{i \in I} y_i \min_{i \in I} y_i)$ .
- **Deviation:** The deviation from the mean. Total absolute deviation  $(\sum_{i \in I} |y_i \overline{y}|)$ or maximum component-wise deviation  $(\max_{i \in I} |y_i - \overline{y}|)$  can be considered.
- **Variance:** The variance of utilities of entities  $(\sum_{i \in I} (y_i \overline{y})^2 / |I|)$ .
- **Gini coefficient:** It is one of the most widely used index in economics. Gini  $\text{coefficient } \left( \frac{\sum_{i \in I} \sum_{j \in I} |y_i - y_j|}{2^{|I|} \sum_{j \in I} |y_j - y_j|} \right)$  $\frac{\sum_{i}^{\in}I\sum_{j\in I}\sum_{j\in I}y_i}{2|I|\sum_{i\in I}y_i}$  takes values between 0 and 1, and the smaller the gini index is, the lower is the inequality.
- **Sum of pairwise absolute differences:** The sum of pairwise absolute differences among all entities  $(\sum_{i \in I} \sum_{j \in I} |y_i - y_j|)$

#### **2.1.3. Inequity-averse aggregation function based approach**

In order to achieve equity-efficiency trade-off, a function  $U : \mathbb{R}^m \to \mathbb{R}$  that concerns both equity and efficiency can be used. In this case, the original problem is modified as  $\max\{U(y) : y \in Y\}$  where  $Y \in \mathbb{R}^m$  is set of feasible solutions.

A function needs to satisfy the following properties to be inequity-averse (Karsu and Morton, 2015):

- If  $y^1 < y^2$  then  $U(y^1) < U(y^2)$   $\forall y^1, y^2 \in Y$  (Function *U* is strictly increasing with respect to every coordinate).
- $U(y) = U(\Pi^l(y))$ , where  $\Pi^l(y)$  is an arbitrary permutation of the *y* vector (Function *U* is symmetric).
- If  $y_j > y_i$  then  $U(y) < U(y \varepsilon e_j + \varepsilon e_i)$   $\forall y \in \mathbb{R}^m$ , where  $0 < \varepsilon < y_j y_i$ , and  $e_j, e_i$  are the *j*<sup>th</sup> and *i*<sup>th</sup> unit vectors in  $\mathbb{R}^m$  (Function *U* satisfies Pigou-Dalton principle of transfers).

A general fairness scheme, "*α*-fairness" (Atkinson, 1970), defines a parametric class of utility functions as given in Equation (2.1).

$$
U_{\alpha}(y) = \begin{cases} \sum_{i \in I} \frac{y_i^{1-\alpha}}{1-\alpha} & \alpha \ge 0, \alpha \ne 1 \\ \sum_{i \in I} \log(y_i) & \alpha = 1 \end{cases}
$$
 (2.1)

Moreover, proportional fairness scheme (for which a transfer of a resource between entities is favorable only if the percentage increase in the utility of the receiving end is higher than the percentage decrease in the utility of the other) and max-min fairness schemes are captured as the special cases of  $\alpha$ -fairness for  $\alpha = 1$  and  $\alpha \to \infty$ , respectively. For the case  $\alpha = 0$ ,  $\alpha$ -fairness scheme is equivalent to the utilitarian approach. Therefore  $\alpha$  parameter can be used to change the attitude on efficiency-fairness tradeoff and it is called the inequality aversion parameter (Bertsimas et al., 2012).

#### **2.2. Efficieny-Fairness Trade-off**

It is obvious that any fair solution concept will probably cause some deterioration in the total system efficiency. Bertsimas et al. (2011) and Bertsimas et al. (2012) provide means to discuss the relation of efficiency loss and fairness increase. Let  $S(Y)$ be the objective of the utilitarian solution for the set of feasible solutions  $Y \in \mathbb{R}^m$ , and *F*(*Y* ) be the objective of a fair solution under a certain fairness scheme. Then they define "price of fairness" as the the relative efficiency loss under an imposed fairness scheme, given in Equation (2.2). They also derive theoretical bounds for price of fairness under  $\alpha$ -fairness scheme as a function of the number of entities and inequality aversion parameter.

$$
POF(Y) = \frac{S(Y) - F(Y)}{S(Y)}\tag{2.2}
$$

#### **2.3. Objectives Used in Disaster Response**

Various objectives are employed in humanitarian relief distribution models that represent the needs. Trade-offs involve the decision of item types to satisfy and in which order the demand points will receive them, which bring conflicting aims in efficiency, effectiveness and equity (Gralla et al., 2014).

Efficiency objectives mostly consider operational costs of the humanitarian agencies or total travel times. Balcik et al. (2008), Tzeng et al. (2007) and Pérez-Rodríguez and Holguín-Veras (2015) consider minimization of operational costs. Campbell et al. (2008), Huang et al. (2012), Tzeng et al. (2007), Lin et al. (2011), Lin et al. (2012) consider travel times. Huang et al. (2015) define their efficiency objective as " the lifesaving utility", which weights each demand node by a marginal utility of that node.

Effectiveness objectives include maximization of the amount of demand satisfied (Lin et al., 2011, 2012), penalty cost due to unsatisfied demand (Balcik et al., 2008)

or demand weighted arrival times (Huang et al., 2012). Huang et al. (2015) calculate effectiveness as the "delay cost of suffering", which is a sum of unsatisfied demand weighted by a linearly increasing function of time.

Equity objectives are usually related with the arrival times of relief items or the satisfaction rate. Among the ones that consider arrival times Campbell et al. (2008) minimizes the latest arrival time and the sum of arrival times. Huang et al. (2012) minimize time weighted disutility by using a piecewise linearly increasing function to penalize late arrivals more than others. Pérez-Rodríguez and Holguín-Veras (2015) uses social cost concept that increases exponentially as the deprivation time increases. Among the ones that take satisfaction rates into account, Tzeng et al. (2007) maximize the minimum satisfaction rate. On the other hand, Lin et al. (2011) and Lin et al. (2012) minimize the maximum difference in satisfaction rates, and Huang et al. (2015) minimize the variance in satisfaction rates.

# **2.4. Aim**

In this study, we aim to provide linear objective functions that can focus on effectiveness and fairness depending on a parameter, and can be applied to the response stage model that we develop. In our model, we do not assume the demand of items occur at the initial period and demand requirements can occur during the planning horizon. All of the studies mentioned in Section 2.3 assume all demand requirements are determined at time zero, except Lin et al. (2011) and Lin et al. (2012). Moreover, we ensure that demand satisfaction is guaranteed at the end of the planning horizon for all demand points. In the instances, we fix the planning horizon, which inherently considers the time related objectives. Therefore for the objective functions that we will propose, we tried to capture both the timeliness and satisfaction levels in a way that their progression is important, rather than their final values.

It should also be noted that the properties of inequity-averse functions that are stated in Section 2.1.3 deals with allocation of resources in a static environment. However, in relief distribution we deal with a multi-period problem, and the network and problem structure such as number of vehicles and their capacities limit the feasible allocation items and their times. Therefore the *y* utility vector is not the allocation values for each entity, but a vector of satisfaction levels for each time period. Therefore we were unable to come up with a inequity-averse aggregation function that satisfy the given properties. However, the problem properties (guaranteed demand satisfaction, choosing smallest planing horizon) helps us to observe similar properties.

In this respect, the response stage model along with alternative objective functions to study are provided in Chapter 3.

# **3. THE RESPONSE STAGE MODEL TO STUDY ALTERNATIVE OBJECTIVES**

### **3.1. Response Stage Model**

Response stage is the phase when the execution of the relief operations performed given the previous decisions. Operational decisions should be considered in detail while the deliveries should be sufficient, timely and fair.

The properties of the model that will be considered in response stage are as follows:

- A single item exits.
- Requirements at TFs does not have to occur at time zero, demands can arrive during the planning horizon.
- Demand must be satisfied as much as possible in a timely manner, and demand satisfaction levels should differ as small as possible to provide fairness. All demand must be satisfied at the end of the planning horizon.
- Multiperiod setting is required.
- Multiple number of depots with different initial supply amounts exist.
- Heterogeneous and capacitated fleet of vehicles do not have assigned depots, therefore tours are not mandatory and open. They can start operating any time at any node.
- Vehicles can collect items from different depots, inventory can be carried from one PF to another, and split deliveries are allowed.

In order to have a base model that will allow us to study alternative objective functions, we provide a multi-depot, heterogeneous-vehicle, capacitated, open vehicle routing problem with split deliveries that will satisfy the properties above.

The sets, parameters and the decision variables that are needed for the response stage model are given below and Equations (3.1−3.23) define *ResponseModel*.

# *Sets:*



#### *Parameters:*



### *Decision variables:*



- *Yvnmt* Amount of items carried on vehicle *v* that traveled on arc (*n, m*) arrived to node *m* at time *t*
- *Ijt* Remaining amount of inventory at PF *j* at the beginning of time *t*

 $Z_{vit}$  Amount of items satisfied at TF *i* by vehicle *v* at time *t* 

- $G_{vnt}$  Binary variable indicating if vehicle *v* finishes its service at node *n* at time *t*
- $P_{it}$  The percentage of the satisfied cumulative demand in TF  $i$  at time  $t$
- $Q_{it}$  Binary auxiliary variable indicating if the cumulative amount carried to TF  $i$  up to time  $t$  is larger than the cumulative demand requirement

# *ResponseModel:*

min 
$$
f(X, Y, Z, I, P)
$$
 (3.1)  
\ns.t.  
\n $X_{vnmtt} = 0$   $v \in V, (n, m) \in A, t = 0$  (3.2)  
\n $I_{jt} = s_j$   $j \in J, t = 0$  (3.3)  
\n $B_{vnt} + \sum_{m:(n,n) \in A} X_{vmmt} = \sum_{m:(n,m) \in A} X_{vnm(t+\tau_{nm})} + G_{vnt}$   $v \in V, n \in N, t \in T$  (3.4)  
\n $Y_{vnmt} \le c_v X_{vnmt}$   $v \in V, (n, m) \in A, t \in T$  (3.5)  
\n $I_{jt} + \sum_{v \in V} \sum_{n:(n,j) \in A} Y_{vnjt} = I_{j(t+1)} + \sum_{v \in V} \sum_{n:(j,n) \in A} Y_{vjn(t+\tau_{jn})}$   $j \in J, t \in T$  (3.6)  
\n $\sum_{v \in V} Y_{vmi} = Z_{vit} + \sum_{n:(i,n) \in A} Y_{vin(t+\tau_{in})}$   $v \in V, i \in I, t \in T$  (3.7)  
\n $\sum_{v \in V} \sum_{t \in T} Z_{vit} = \sum_{t \in T} d_{it}$   $i \in I$  (3.8)  
\n $\sum_{v \in V} \sum_{t \in T} G_{vnt} = 1$   $v \in V$  (3.9)  
\n $\sum_{n:(n,i) \in A} Y_{vni} = \sum_{t \in T} X_{vnit}$   $v \in V, i \in I, t \in T$  (3.10)  
\n $\sum_{v \in V} \sum_{t \in T} G_{vnt} = 1$   $v \in V, i \in I, t \in T$  (3.11)  
\n $\sum_{v \in V} \sum_{t \in O} Z_{vit} = \sum_{t \in T} X_{vnit}$   $i \in I, t \in T$  (3.12)  
\n $\sum_{t \in O} \sum_{t \in O} Z_{vit} = \sum_{t \in T} Z_{vit} = 0$   
\n $i \in I, t \in T$  (3.1



In *ResponseModel* only the constraints are presented since alternative objective functions will be presented in the following sections. Constraints (3.2) and (3.3) set the initial values for the variables related to vehicle and inventory levels in the first time segment, which will be implemented as variable bounds instead of explicit constraints. Constraints (3.4) define the vehicle conservation and (3.5) limit the amount transported on a vehicle by its capacity. Equations (3.6) define inventory balance constraints for PFs, where items collected from other PFs can be dropped of to another PF. Equation set (3.7) is the inventory balance constraints for TFs. Constraints (3.8) makes sure that all of the demand is satisfied at the end of the planning horizon for all TFs. Equations (3.9) ensures all vehicles will leave the network, which is required for constraints (3.4) to work properly. Constraints (3.10) added to improve effectiveness by making sure that if a vehicle visits a demand point, it has to satisfy some demand there. Equations  $(3.11)$  and  $(3.12)$  determine the value of auxiliary  $Q_{it}$  variables. Constraints  $(3.13)$  – (3.16) calculate the percentage of satisfied demand in TF *i* at time *t*, such that its value is bounded between 0 and 1 even if the amount of item carried to that TF *i* is higher than the cumulative demand up to time *t*. Lastly, equations  $(3.17) - (3.23)$  define the variables.  $Y_{vnmt}$ ,  $I_{jlt}$ , and  $G_{vnt}$  are left as continuous variables since they will obtain integer values due to the definition of the constraints.

#### **3.1.1. Valid Inequalities**

A set of valid inequalities that considers the earliest entrance time to a given node be a given vehicle is defined to improve the solution quality. The valid inequality set eliminates all *Xvnmt* and *Yvnmt* that cannot occur due to initial required paths of a vehicle, and it is defined in equation 3.24 below, where  $\tilde{\tau}_{nm}$  defines the shortest distance between nodes *n* and *m*.

$$
\sum_{n' \in N} \sum_{m:(m,n') \in A} \sum_{t'=0}^{\bar{t} + \tilde{\tau}_{\bar{m}m} + \tau_{mn'} - 1} Y_{\bar{v}mn't'} + X_{\bar{v}mn't'} = 0 \quad \bar{v} \in V, \bar{n} \in N, \bar{t} \in T : B_{\bar{v}\bar{n}\bar{t}} = 1
$$
\n(3.24)

#### **3.2. Proposed Objective Functions**

The *ResponseModel* ensures all requirements will be satisfied at the end of the planning horizon, assuming the total inventory held in the PFs is sufficient to cover all cumulative demands of TFs and the planning horizon is long enough. Therefore effectiveness consideration is related with the timeliness of the distribution of goods and the fairness consideration is related with the discrepancy among TFs during the horizon.

In Section 2.1, utilities of each TFs,  $y_i$ , are defined to discuss fairness considerations. As mentioned before, satisfaction levels of TFs at the end of the planning horizon cannot be considered in our case since demands satisfaction is guaranteed. Therefore we consider the satisfaction levels with respect to time buckets. Moreover, we cannot use demand weighted arrival times since the demand generation is not completed at time zero, and satisfaction levels of a TF may reduce as new demand requirements occur.

One way to define the *"disutility"* of each TF *i* is  $y_i = \sum_{t \in T} (1 - P_{it})$ , which can be interpreted as the area above the satisfaction level curve for that TF. However it should be noted that this definition could be considered as inadequate since it s hard to define the trade-off between short durations of low satisfaction levels and long durations of higher satisfaction levels.

With these in mind, the proposed objective functions are given in the following subsections.

#### **3.2.1.** *ResponseModelObj*<sup>1</sup>

For the first objective, we can consider the discrepancy among TF points at each time period. One can assert that it is unfair for TF *i* and TF *j* have different disutilities. Using this idea, we can penalize the satisfaction ratio differences between TFs at each time periods. Moreover, we can suggest that the differences that are closer to the end of the planning horizon are more important than the earlier ones. Therefore *ResponseModelObj*<sup>1</sup> minimizes the total sum of percentage differences between satisfaction levels weighted by  $t^{\alpha}$ .

The additional decision variables required and *ResponseModelObj*<sup>1</sup> are given below.

#### *Additional decision variables:*

 $D_{ijt}$  Percentage difference between the satisfaction levels of TF *i* and TF *j* at time *t*

# *ResponseModelObj*1*:*

$$
\min \qquad \sum_{i \in I} \sum_{\substack{j \in I \\ j \neq i}} \sum_{t \in T} t^{\alpha} D_{ijt} \tag{3.25}
$$

s.t.

 $(3.2)$ 

$$
-3.23)
$$

 $P_{it} - P_{jt} \le D_{ijt}$  *i* ∈ *I, j* ∈ *I, j* ≠ *i, t* ∈ *T* (3.26)

$$
D_{ijt} \ge 0 \qquad i \in I, j \in I, j \ne i, t \in T \qquad (3.27)
$$

Intuitively, one would expect that the discrepancies between TFs should be smaller in the last periods as the value of  $\alpha$  increases, which indicates that the decision maker accepts early differences more than the later ones.

Also it should be noted that this objective is equivalent to "sum of pairwise absolute differences" when  $\alpha = 0$ .

#### **3.2.2.** *ResponseModelObj*<sup>2</sup>

The second objective function is based on multi-objective optimization where effectiveness can be defined by the total unsatisfied demand percentages throughout the planning horizon, and fairness metric can be given as the total difference between the satisfaction levels between two TFs. It is actually an extension of the first objective by also considering the utility term. The corresponding *ResponseModelObj*<sup>2</sup> model is given below where  $f(\alpha, t)$  is the fairness-efficiency trade-off parameter which is a function of time.

#### *Additional decision variables:*

 $D_{ijt}$  Percentage difference between the satisfaction levels of TF *i* and TF *j* at time *t*

# *ResponseModelObj*2*:*

min 
$$
\sum_{i \in I} \sum_{t \in T} (1 - P_{it}) + \sum_{i \in I} \sum_{\substack{j \in I \\ j \neq i}} \sum_{t \in T} f(\alpha, t) D_{ijt}
$$
 (3.28)

s.t.

(3*.*2 − 3*.*23)

$$
P_{it} - P_{jt} \le D_{ijt} \qquad i \in I, j \neq i, t \in T \qquad (3.29)
$$

$$
D_{ijt} \ge 0 \qquad \qquad i \in I, j \in I, j \ne i, t \in T \qquad (3.30)
$$

Unlike *Obj*1, this objective includes the utility term as well. Also, we can define different  $f(\alpha, t)$  functions that can perform differently. We can have  $f_1(\alpha, t) = t^{\alpha}$  to define  $Obj_{2,1}$ , or it can linearly increase from zero to  $2\alpha$ , i.e.  $f_2(\alpha, t) = \frac{2\alpha t}{T}$  to define

 $Obj_{2,2}$ . Moreover, by defining the objective function as in Equation  $(3.31)$ , we obtain *Obj*2*.*<sup>3</sup> which can emphasize effectiveness in earlier periods, and equality in the latter periods.

$$
\sum_{i \in I} \sum_{t \in T} (2\alpha - f_2(\alpha, t))(1 - P_{it}) + \sum_{i \in I} \sum_{\substack{j \in I \\ j \neq i}} \sum_{t \in T} f_2(\alpha, t) D_{ijt}
$$
(3.31)

Intuitively, one would expect that the discrepancies between TFs should be smaller in the last periods as the value of  $\alpha$  increases. Also, one would expect the model focus more on effectiveness as the value of  $\alpha(t)$  reduces. It should be noted that objective  $Obj_{2,2}$  is equivalent to the "utilitarian" objective with  $\alpha = 0$ .

# **3.2.3.** *ResponseModelObj*<sup>3</sup>

The third objective function is inspired by the Rawlsian approach. However, unlike Rawlsian approach, we consider the minimum satisfaction level among all TFs at each time period. The corresponding  $ResponseModelObj_3$  is given below where  $t^{\alpha}$ term tries to increase effectiveness by pushing the model to satisfy the all demands as early as possible. Also note that the  $Q_{it}$  binary variables are not needed to define  $P_{it}$ variables correctly in this model.

#### *Additional decision variables:*

 $D_t$  The minimum satisfaction level among all TFs at time  $t$ 

#### *ResponseModelObj*3*:*

$$
\min \qquad \sum_{t \in T} (1 - D_t) t^{\alpha} \tag{3.32}
$$

s.t.

$$
(3.2-3.10), (3.17-3.21), (3.23)
$$
\n
$$
\sum_{x \in V} \sum_{t'=0}^{t} Z_{vit'}
$$

$$
P_{it} \ge \frac{\sum_{v \in V} \sum_{t'=0}^{t} Z_{vit'}}{\sum_{t'=0}^{t} d_{it'}} \qquad i \in I, t \in T \qquad (3.33)
$$

$$
P_{it} \ge D_t \qquad i \in I, t \in T \qquad (3.34)
$$

$$
D_t \ge 0 \qquad \qquad t \in T \qquad (3.35)
$$

Intuitively, one would expect that all of the requirements to be satisfied as much as possible in the last periods if the value of  $\alpha$  increases. Although the differences among TFs are ignored, the we expect all of the TFs to be improved by considering the worst one.

#### **3.3. Objective Functions in the Literature**

In this section, we adapt four objective functions in the literature to use in our *ResponseModel*. One of them, namely *ObjHSB* considers the satisfaction percentages of TF points at each time period, whereas the rest uses a single utility definition for each demand point.

### **3.3.1.** *ResponseModelObjHSB*

Huang et al. (2012) solve a vehicle routing problem for the last-mile distribution from a warehouse to demand points. For this model, they define an equity objective, which calculates the fraction of the total unsatisfied demand at each time period for demand nodes, and converts it into disutility using a piecewise linear convex function as demonstrated in Figure 3.1 below. In this way, they aim to prioritize the worst-off demand points before others.



Figure 3.1: An example piecewise linear disutility function (Huang et al., 2012).

The additional parameters and variables, and the corresponding *ResponseModelObjHSB* model is given below.

#### *Additional parameters:*

- $a_p$  Parameter value that defines the slope of the linear constraint in  $p^{th}$  segment. *a<sup>p</sup>* = {4*/*13*,* 8*/*13*,* 16*/*13*,* 24*/*13}
- *b<sup>p</sup>* Parameter value that defines the intercept of the linear constraint in *p th* segment. *b<sup>p</sup>* = {0*,* −1*/*13*,* −5*/*13*,* −11*/*13}

#### *Additional decision variables:*

 $F_{it}$  The disutility of TF *i* in time *t* 

#### *ResponseModelObjHSB:*

min  $\sum$ *i*∈*I*  $\sum$ *t*∈*T*  $F_{it}$  (3.36)

s.t.

$$
(3.2 - 3.23)
$$
  
\n
$$
F_{it} \ge a_p(1 - P_{it}) + b_p
$$
  
\n
$$
i \in I, t \in T, p \in P
$$
  
\n
$$
i \in I, t \in T
$$
  
\n
$$
(3.37)
$$
  
\n
$$
i \in I, t \in T
$$
  
\n
$$
(3.38)
$$

#### **3.3.2.** *ResponseModelP iecewiseP ropF air*

In Section 2.1.3, proportional fairness was introduced as a special case of *α*-fairness. Since it maximizes a nonlinear function  $(\sum_{i \in I} U_i = \sum_{i \in I} log(y_i))$ , we defined it using a piecewise linear approximation, where the utility  $y_i$  is defined as  $y_i = \frac{\sum_{t \in T} P_{it}}{|T|}$  $\frac{q(T^{\perp}t)}{|T|}$  (i.e. normalized area under the satisfaction curve of  $TF$  *i*). The piecewise linear function that we used is given in Figure 3.2 below, where the horizontal axis represent  $y_i$  and the vertical axis represent  $U_i$ .



Figure 3.2: Piecewise linear approximation to the proportional fairness.

Then the corresponding  $ResponseModel PiecewisePropFair$  is given below.

#### *Additional parameters:*

- $a_p$  Parameter value that defines the slope of the linear constraint in  $p^{th}$  segment. *a<sup>p</sup>* = {23*,* 92*/*15*,* 69*/*25*,* 41*/*25*,* 28*/*25}
- $b_p$  Parameter value that defines the intercept of the linear constraint in  $p^{th}$ segment. *b<sup>p</sup>* = {−23*/*5*,* −437*/*150*,* −207*/*100*,* −151*/*100*,* −28*/*25}

#### *Additional decision variables:*

 $U_i$  Approximated logarithm of the utility of TF *i* 

#### *ResponseModelPiecewisePropFair:*

$$
\min \qquad -\sum_{i \in I} U_i \tag{3.39}
$$

s.t.

$$
(3.2 - 3.23)
$$
  

$$
U_i \le a_p \frac{\sum_{t \in T} P_{it}}{|T|} + b_p
$$
  
 $i \in I, p \in P$  (3.40)

$$
U_i \ge 0 \qquad \qquad i \in I \tag{3.41}
$$

#### **3.3.3.** *ResponseModelRange*

The range objective is a pure equality objective where the difference in the maximum and minimum disutilities are minimized. The corresponding *ResponseModelRange* is given below.

# *Additional decision variables:*

- *R*<sup>1</sup> The maximum disutility among TFs
- *R*<sup>2</sup> The minimum disutility among TFs

# *ResponseModelRange:*



# **3.3.4.** *ResponseModelRawlsian*

The Rawlsian objective is another pure equality objective where the maximum disutility is minimized. The corresponding *ResponseModelRawlsian* is given below.

# *Additional decision variables:*

*R* The maximum disutility among TFs

# *ResponseModelRawlsian:*



# **4. SOLUTION METHOD**

As mentioned in Section 3.1 is a hard problem. Also in our preliminary experimentations, we observed that the CPLEX solver fails to find the solution in a reasonable amount of time even for very small number of facilities, as expected. Although we did not expect to solve large instances efficiently, we first studied decomposition techniques, aiming to solve medium-sized problems.

The first idea was to use *Logic-based-Benders-Decomposition (LBBD)* (Hooker, 2000; Hooker and Ottosson, 2003) to decompose the problem. The idea was built on the relationship between the vehicular decisions  $(X_{v n m t})$  and the amount of items carried on them (*Yvnmt*). By relaxing the vehicle balance constraints 3.4 and leaving the *Xvnmt* variables in the subproblem, the remaining master problem decides on the amount of items moved from node to node while satisfying the inventory balance equations. However, the solution of the relaxed master problem may yield solutions where a vehicle can perform multiple tasks simultaneously. Then, depending on the type of infeasibility the logic-based feasibility cuts would be used to eliminate these infeasible solutions. However for some of the cases, we were unable to find feasibility cuts that are proven to be valid. Moreover, when the cuts that are proven to be valid are implemented, the run times are worsened. Therefore our attempt to decompose the model using LBBD is failed.

In the VRP literature, the most common decomposition method to tackle vehicle routing problems is Branch & Price. We can reformulate our model, by enumerating all possible routes of a vehicle. A route for a vehicle corresponds to a sequence of nodes, starting at its initial node and visiting other nodes without delay until the end of the planning horizon. Then we obtain a relaxed master problem by generating a subset of these routes initially, which can be solved using column generation by generating promising columns using the pricing problem. However we postponed the effort on implementing a Branch & Price algorithm since we believe it is not a very promising direction. First of all, the requirement to satisfy all demand makes it harder to come up with feasible routes. Other than that, the pricing problem contains almost all constraints that decides on the vehicle routing. Moreover, in our problem depots are shared by vehicles and inventory relocation is allowed, therefore the pricing problems are not separable over vehicles. Therefore we decided to focus on a heuristic method, namely *adaptive large neighborhood search* (ALNS) for the time being.

#### **4.1. Adaptive Large Neighborhood Search**

Adaptive large neighborhood search (ALNS) is introduced by Ropke and Pisinger (2006a) and further improved in Ropke and Pisinger (2006b), and Pisinger and Ropke (2007) for rich pickup and delivery models. It is an extension to the *large neighborhood search* (LNS) framework given by Shaw (1998) by incorporating an adaptive layer. ALNS provides a general out-of-the-box framework that can be applied to many vehicle routing problem types and allows to mix various VRP variants.

ALNS is a local search framework, in which destroy and repair heuristics compete to alter the current solution vector. It requires the definition of a set of *destroy neighborhoods* and a set of *repair neighborhoods*. At each iteration, a destroy heuristic is selected, which removes at most *q* many variables from the solution vector. This is followed by assigning feasible variables to those variables using the selected repair heuristic. This framework is defined within a local search framework (e.g. simulated annealing, tabu search, or guided local search), and the generated candidate solution is accepted if it satisfies the acceptance criteria of the local search framework at the master level. In each iteration a destroy and a repair heuristic is selected separately using roulette wheel with probabilities determined by the past performance. The adaptive layer of the algorithm updates the performance (scores) of the neighborhoods, and the heuristics that perform better obtains a higher chance to be selected. Pisinger and Ropke (2007) state that neighborhoods are usually searched with fast heuristics, and it could be advantageous to use neighborhoods that can provide diversification.

#### **4.2. ALNS Applied to the** *ResponseModel*

For our *ResponseModel*, we build the ALNS algorithm within Simulated Annealing (SA) framework. Solution representation, destroy and repair heuristics, improvement step that we incorporated, adjustments of the adaptive scores of these heuristics, initial solution generation and other implementation details are given in the following subsections.

#### **4.2.1. Solution Representation**

The solution vector of the proposed SA-ALNS heuristic is composed of |V| many list of nodes. For each vehicle, there is a list of node-time pair that represents the sequence of the nodes that will be visited by that vehicle, and arrival times. Obviously, the first node in each list is the node where the corresponding vehicle starts its mission. Given a list of nodes to visit and the starting time of a vehicle (i.e.  $B_{vnt}$ ), and given the fact that vehicles cannot wait in nodes, the sequence can be used to come up with the schedule of the vehicle. A feasible solution vector contains lists of nodes that do not violate the planning horizon imposed by the decision maker. Therefore we can obtain the values of all *Xvnmt* variables from the solution vector. Then the fitness value of a solution vector can be calculated by solving the remaining mixed-integerprogram (MIP) after fixing the corresponding *Xvnmt* variables. However, since we cannot guarantee the feasibility of a solution vector, constraint (3.8) is replaced with the relaxed equation  $(4.1)$  below. In constraint  $(4.1)$ ,  $H_i$  variables indicate the amount of unsatisfied items for TF  $i$  at the end of the planning horizon. Then  $H_i$  variables are penalized in the selected objective function with appropriate penalty values. Although the remaining fitness calculation problem is a MIP, it takes relatively short time to solve the problem.

$$
\sum_{v \in V} \sum_{t \in T} Z_{vit} + H_i = \sum_{t \in T} d_{it} \qquad i \in I \tag{4.1}
$$

#### **4.2.2. Destroy Heuristics**

Four different destroy heuristics are defined for the proposed SA-ALNS. Each of the destroy heuristics use *destroyProb* as a parameter that determines neighborhood length.

- *RandomRemoval***:** It is the simplest destroy heuristic that takes solution vector as an input, and nullifies each node with *destroyProb* probability. This helps us to improve diversification in the search.
- *LongestArcRemoval***:** Given the solution vector, nodes that it takes longest to travel to are nullified. The number of removals equals to the *destroyProb* fraction of all visited nodes. In this way we aim to eliminate the long travel times, and to be able to increase the number of nodes that can be visited within the planning horizon.
- *MaxFilledRemoval***:** Since the demand occurs any time in the planning horizon, some of the demand nodes may receive more items than the generated demand at that time. Since nodes with  $P_{it} > 1$  are corrected to  $P_{it} = 1$ , this indicate potential improvement on TF points with smaller fill rates without worsening such nodes. Therefore given the solution vector, *destroyProb* fraction of nodes with highest fill rates are nullified.
- *RandomRemovalWithoutInfeasibleNodes***:** It is the same as the *Random-Removal* heuristic, except we do not nullify TF nodes that receive  $H_i > 0$  in the current solution. In this way we aim to increase diversity of the solution vector while trying to remain/become feasible.

#### **4.2.3. Repair Heuristics**

After a destroy heuristic is applied, the solution vector still remain as a list of nodes for each vehicle. However, consecutive node-time pairs do not constitute a feasible vehicle path. Then, four different repair heuristics are defined for the remaining incomplete solution vector.

• *SolverRelaxedPath***:** The remaining incomplete solution vector indicate at some specific time periods, a vehicle should enter a node, without indicating the originating node. Using this idea, instead of fixing all  $X_{v n m t}$  variables, we can fix the remaining ones as hard constraints, and solve the *SolverRelaxedPath* MIP model. Since this may yield only one of few alternative solutions then the previous one, we relax the entrance times to nodes using a *maxTimeRelax* parameter. We generate a time relaxation value *tr*, between 1 and *maxTimeRelax* randomly, and include the following constraints (4.2) in the model. In equation  $(4.2)$ ,  $\bar{n}$  and  $\bar{t}$  indicate a node-time pair in the incomplete solution vector for vehicle  $\bar{v}$ . In order to avoid spending too much time in solving *SolverRelaxedPath*, 60 seconds of time limit is imposed.

$$
\sum_{t=\bar{t}-tr}^{\bar{t}+tr} \sum_{m:(m,\bar{n})\in A} X_{\bar{v}m\bar{n}t} = 1
$$
\n(4.2)

• *AppearanceLikelihoodGreedy***:** In this repair heuristic, we ignore the time stamps on node-time pairs in the solution vector. For each TF and PF in the instance, we calculate a so-called *appearance likelihood*. For TFs, *appearance likelihood* means the expected number of appearance for a TF in a feasible solution vector, and it is calculated by dividing its total demand to average vehicle capacity. For PFs, we expect a PF to appear in a feasible solution vector *T otalOverallDemand/AverageV ehicleCapacity* many times. By multiplying this value with the fraction of supply by a PF among all, we calculate its *appearance likelihood*. In each iteration, the number of appearance of a node in the incomplete solution vector are subtracted from its *appearance likelihood* value to use as a selection parameter.

Apart from that, we calculate *expectedTFrun*, which indicates expected number of consecutive TF nodes in a solution vector of a vehicle before visiting a PF node as (*AverageV ehicleCapacity* ∗ *T*)*/T otalOverallDemand*. This value is rounded up or down stochastically.

Then *AppearanceLikelihoodGreedy* heuristic start to repair the incomplete solution vector by considering the node with the highest *appearance likelihood*. Depending on the type of the currently considered node (TF or PF), and considering the *expectedTFrun* length, we insert the current node into the solution vector position which cause the shortest detour length among all possibilities, while obeying planning horizon constraints. After each insertion *appearance likelihood* of that node is reduced by one, and procedure continues until all nodes have negative *appearance likelihood* values, or there are no more possible insertions due to planning horizon limit. As the last step, time-stamps in the node-time pairs in the solution vector are corrected.

- *AppearanceLikelihoodRandom***:** Performs in the similar fashion with *AppearanceLikelihoodGreedy* heuristic, except we insert the node with the highest *appearance likelihood* to a random position. In this way we aim to improve diversification.
- **RandomInsertion:** This heuristic randomly selects an available vehicle, inserts a random node into a random position. A vehicle becomes unavailable whenever the required time to visit all nodes in its list exceeds planning horizon.

#### **4.2.4. Improvement Step**

We incorporate an improvement step in our SA-ALNS approach which is not originally present in the ALNS framework. Constraints (3.10) ensure that if a vehicle enters a TF node, it has to satisfy at lest one unit of demand there, to improve efficiency in the original formulation. If a repaired solution yields  $Z_{vit} = 1$  for a node *i* and vehicle *v*, this indicates that the solution may be improved by visiting that node. Then at the end of repair operations, we eliminate nodes with  $Z_{vit} = 1$  from the solution vector as an improvement step.

#### **4.2.5. Adjustment of Adaptive Weights**

The roulette wheel selection mechanism of the destroy and repair heuristics depend on the score of each heuristic. The adaptive layer in the ALNS framework enables better performing heuristics to be selected with a higher probability. The scores are collected within iteration segments, which is defined by *scoreUpdateInterval* parameter. At the beginning of each interval, interval scores are set to zero. Then within the interval, scores of the selected heuristics are incremented by *scoreG*, *scoreI* and *scoreW* for global best solution, improving solution, and accepted worsening solution, respectively. The total interval score of a heuristic is divided by its number selection to find its interval score at the end of the time segment. Then the previous scores are smoothed using a smoothing parameter *scoreSmooth* to incorporate the previous scores in the roulette wheel selection.

#### **4.2.6. Initial Solution Generation**

We construct an initial solution greedily, and considering the vehicles and nodes dynamically. We assume all vehicles will collect items from their starting PF node as much as their capacities, and construct a solution starting there. For each vehicle, the nodes that we can reach to are the candidates, and we calculate a *nodeWeight* for those. If the candidate node is a PF node, we can increase the amount carried on the vehicle up to our capacity if there are sufficient items left. Then the *nodeWeight* for that PF is the ratio of the amount that can be taken from that PF to the travel distance to that PF. On the other hand if the candidate node is a TF, by going to that TF we can satisfy some of the generated demand up to that time. Then the *nodeWeight* for that TF is calculated dividing the average of the current load on the vehicle and the unsatisfied remaining demand of that TF by the travel distance. Among all candidates, we chose the one with the highest *nodeWeight*, and we add the selected candidate node to the visiting sequence of the corresponding vehicle. If the selected node is a TF, we assume the whole demand generated up to that point is satisfied using the items on the vehicle, if there are enough of them. Otherwise, all items carried on the vehicle are used to satisfy a portion of the generated demand. On the other hand, if the selected node is a PF, we assume the vehicle will be filled up to its capacity as long as there are sufficiently many items remain in the PF. In each case, the amount of items on the vehicle and PFs, and the total amount of items that are satisfied in TFs are updated and the algorithm terminates whenever no other nodes can be visited by a vehicle due to imposed planning horizon.

# **4.2.7. Other Implementation Details**

As mentioned before, ALNS is implemented within SA framework. The SA procedure is run for *numIter* many iterations. A temperature value *T* is initialized, and it is decreased in each iteration by geometric cooling scheme, given in  $(4.3)$  where  $\alpha$  is the cooling coefficient. At each temperature level, a neighboring solution  $S'$  is generated from the current solution *S*. If the objective value of the neighboring solution,  $F(S')$ , is an improved value over the current objective value  $F(S)$ , then the move is always accepted. Moreover, the worsening moves are accepted with the probability function given in (4.4), for the minimization problem case. At high temperatures probability of accepting a worsening move is more than the lower temperature case, and this results in diversification at the beginning where the temperature is high, and intensification at the end of the procedure where temperature is lower.

$$
T_{iter+1} = \alpha T_{iter} \tag{4.3}
$$

$$
P(acceptance) = \exp\left(\frac{F(S) - F(S')}{T}\right)
$$
\n(4.4)

The values for the parameters used in SA-ALNS procedure are given as follows:

- $number = 300$
- $\alpha = 0.97$
- $T = 1000000$
- $destroyProb = 0.3$
- $maxTimeRelax = 2$
- $scoreUpdateInterval = 30$
- $scoreSmooth = 0.25$
- $scoreG = 4$
- $scoreI = 2$
- $scoreW = 1$

Then the algorithmic steps of the SA-ALNS heuristic is given in Algorithm 1.



# **5. NUMERICAL RESULTS**

In this chapter we analyze the performance of the *ResponseModel* with different objective functions using a commercial solver and using the SA-ALNS heuristic. Then the properties and the behavior of the alternative objective functions are evaluated using nine metrics that range from utilitarian to egalitarian approaches. For this purpose, a set of test problems with different sizes are generated.

#### **5.1. Data Generation and Test Instances**

The data generation scheme we use is parameterized and enables the decision maker to generate instances with different characteristics. As parameters, the *seed* value to initialize a random stream, the planning horizon *T*, the *number of PFs*, the *number of TFs*, and the *number of vehicles* are given by the decision maker. Since we generate a graph, where all  $J \times I \cup I \times I$  arcs are present, we use the distance matrix of the 81 cities in Turkey. We randomly select the PFs and TFs among the cities. Then the distances between these cities are used to generate the arc distances, with respect to the parameter that determines the *distance in kilometers that correspond to a time bucket*. We assume a heterogeneous fleet of vehicles, so we provide a set of *vehicle capacities*, which are randomly assigned to the vehicles by the given *vehicle size probabilities*. If the *number of vehicles* is equal to the *number of PFs*, each vehicle starts from a different random PF node. If there are more vehicles than PFs, the excess vehicles are randomly assigned to PFs. If there are less number of vehicles than PFs, then the vehicles are assigned to PF nodes randomly.

In order to generate demand amounts, decision maker provides a *total expected number of vehicle runs* parameter, which is multiplied by the total vehicle capacity to obtain the total demand over all TFs. A weight is assigned randomly to each TF between 1 and *maximum base demand multiplier for TFs* parameter, and the total demand is distributed over TF points with respect to these weights. In order to distribute

the total demand of a TF over time periods, a random number of arrivals between *min number of arrivals* and *max number of arrivals* are generated.

The last step is to determine the supply amounts in the PF nodes. We multiply the total demand by a random value generated between *min supply demand ratio* and *max supply demand ratio* to obtain the total supply. This total supply is distributed over PF nodes similar to TFs, a weight is assigned randomly to each PF between 1 and *maximum base supply multiplier for PFs* parameter. Then the total supply is assigned to PFs with respect to these weights.

Since we fix the planning horizon *T* beforehand, there is a chance of having an infeasible instance. In this case, we gradually reduce the *total expected number of vehicle runs* until we get to a feasible solution.

To test numerical performance, twenty test problems with different sizes are generated randomly. Among these, instances with 2 or 3 PFs; 4, 6, 7, or 8 TFs; 3 or 4 vehicles; and 16, 20 or 23 time segments exist. Instances are named as "P-*j*-*i*-*v*-*t*", where *j*, *i*, *v*, and *t* represents the number of PFs, TFs, vehicles and time segments, respectively.

#### **5.2. Results**

The models are solved for the given test instances by using IBM ILOG CPLEX 12.8. In all models, *Xvnmt* variables have the highest branching priority. Run times are limited to 8 hours as a stopping condition and the number of threads that the solver can utilize is set to 8. Also %1 relative optimality gap is imposed as a stopping condition. Finally, for *ObjRange*, a 0.01 absolute optimality gap is defined to avoid numerical problems. All experiments are carried out on a PC with 3.30 GHz CPU and 16 GB RAM, running under 64-bit Windows 7 operating system.

Tables 5.1, 5.2, and 5.3 provides summaries of the results of the CPLEX solver















Table 5.4: Results for the objectives in the literature. Table 5.4: Results for the objectives in the literature.

and SA-ALNS for the proposed objectives with  $\alpha = 0$ ,  $\alpha = 1$ ,  $\alpha = 2$ . Similarly Table 5.4 provides the results for the objectives in the literature. In these tables, the first column indicate the problem instance, and the subsequent columns provide relative optimality gaps and run times for each objective, for solver and SA-ALNS. In Tables 5.1, 5.2, 5.3, and 5.4, solver results are given in bold face if a run is terminated before run time (i.e. stopping criteria for the optimality gap is attained). For the SA-ALNS results, gap percentages are calculated with respect to the upper bound of the corresponding solver solution. Therefore a negative value means that SA-ALNS found a better feasible solution then the solver. If SA-ALNS fails to find a feasible solution, then the gap value is given by a dash. Among the SA-ALNS results, the solutions that have gap percentage smaller than  $\%10$  are presented in bold face.

Tables 5.1, 5.2, 5.3, and 5.4 indicate that even small instances are hard to solve to optimality by CPLEX in a reasonable run time. In general, we can observe up to %100 percent optimality gaps, which is caused by poor lower bounds. Also we observe that the proposed objectives are harder to solve compared to the ones in the literature. For SA-ALNS, although there are cases without a feasible solution, an optimal or near optimal solution is also found in some cases. Moreover, SA-ALNS yield better upper bounds compared to the solver in some cases as well. Also the run times are usually much shorter than the other. It should be noted that SA-ALNS procedure is still being improved, and parameter search is not performed yet. Therefore, by incorporating better destroy and repair heuristics, and searching for better parameter values, SA-ALNS procedure appears to be a promising method for our problem.

#### **5.3. Objective Function Evaluation**

As given in Section 5.2, we solve the instances with  $\alpha = \{0, 1, 2\}$  for the proposed objective functions. We first start by analyzing the effect of  $\alpha$  for each of these objective functions. Then we select an alpha value for these proposed objectives and compare them with the objective functions in the literature. In these analyses, we only used the instances where the optimal solution is found for all objective function types. Each solution of a given objective function with a given  $\alpha$  is evaluated with nine different metrics in the literature. These are utility,  $\alpha$ -fairness with  $\alpha = \{1, 2\}$ , range, rawlsian, sum of pairwise absolute differences, total absolute deviation, maximum deviation, and gini coefficient. In order to make the comparison easier, the calculated performance metrics are scaled between 0 and 1 within each performance metric. The comparisons are done on a radar chart, where values closer to zero are the best ones since all objectives are of minimization type. In each axis, the alternatives can be compared within each other, however it should be noted that comparison between axes is not possible. Since the performance metrics range from purely utilitarian to purely equality metrics, an objective alternative where all metrics are close to zero would be an ideal one. However, since utility and equality are conflicting measures, we will look for the best compromise which perform relatively well among all.

Figures 5.1 to 5.5 present the effect of different  $\alpha$  values over the proposed objectives, respectively. Note that  $\alpha = 0$  is not defined for  $Obj_{2,3}$ . As mentioned before, a tight and balanced loop indicates the given objective performs well in all performance metrics. For  $Obj_1$ , we can conclude that  $\alpha = 0$  performs usually well on equality metrices, whereas the performance is worsened for the utilitarian terms. On the other hand, with  $\alpha = 1$  and  $\alpha = 2$  the exact opposite is true. Although there is no dominating  $\alpha$  value among them, we choose  $\alpha = 0$  that focus more on equality. The discussion on *Obj*2*.*1 is in parallel with *Obj*1, however the utilitarian metrics perform better. Therefore we choose  $\alpha = 0$  for  $Obj_{2,1}$  as well. For  $Obj_{2,2}$ ,  $\alpha = 0$  performs well in utilitarian metrics (since it is equivalent to the utilitarian objective) and performs poorly on equality metrics, and the opposite is true for  $\alpha = 2$ . However, we observe that  $\alpha = 1$  provides a nice trade-off among utilitarian and equality metrics, therefore we choose that. For  $Obj_{2,3}$ , we choose  $\alpha = 1$  since it performs better on more metrics, and finally we choose  $\alpha = 1$  for  $Obj_3$  since it provides a compromise between utilitarian and equality metrics.

The next step in the analysis is to compare all proposed objectives with the selected  $\alpha$  values to the objective functions in the literature. Similarly, the metrics are calculated for the results of each objective, and scaled. Figure 5.6 provide the comparison among all objectives. Figure 5.6 reflects that, among the proposed objectives  $Obj_1$  focuses more on equality, whereas  $Obj_2$  versions provide more balanced results compared to  $Obj_1$ .  $Obj_3$  performed better on equality metrics compared to others, which can be explained by considering only the worst-off TFs at each time segments, and one can expect this deterioration to increase as the number of TFs increases further. An important observation is that  $ObjHSB$  and  $ObjPiecewisePropFair$  yield very good results for utilitarian metrics, whereas they perform poorly on equality metrics. *ObjRange* performed well on equality metrics and performed poorly on others, and the opposite is valid for *ObjU tilitarian* as expected. Finally, *ObjRawlsian* yields a good compromise between equality and utility metrics. However, one could expect this performance to deteriorate as the number of TFs is increased. Figure 5.7 emphasizes the performance of the selected best three alternatives, namely  $Obj_{2.1}$ ,  $Obj_{2.2}$ , and *ObjRawlsian*, that provide a good trade-off between utilitarian and egalitarian metrics.

As the final step of the analysis, we calculate the "price of fairness" (POF) as introduced in Section 2.2. The POF for each objective alternative for all considered problems are given in Table 5.5 along with their average.

Problem							$[Obj1]Obj2.1]Obj2.2]Obj2.3]Obj3]ObjHSB]ObjPcwsPrFair[ObjRange]ObjRawls]$		
$P-3-4-3-16$ -0.36		$-0.18$	$-0,18$	$-0.18$	$-0.15$	$-0.04$	$-0.03$	$-0.33$	$-0,15$
$ P-3-4-4-16  - 0, 20 $		$-0,20$	$-0,14$	$-0.14$	$-0.07$	$-0.02$	$-0.02$	$-0,26$	$-0,11$
$\textbf{P-}3\textbf{-}4\textbf{-}3\textbf{-}20$ - 0.33		$-0.31$	$-0.10$	$-0.12$	$-0.06$	$-0.06$	$-0.01$	$-0.36$	$-0.15$
$ P-3-6-3-16  -0.39 $		$-0.06$	$-0.01$	$-0.03$	$-0.07$	0,00	0.00	$-0.37$	$-0.17$
$P-2-6-2-20$ -0.11		$-0.02$	$-0.03$	$-0.03$	$-0.09$	$-0.03$	$-0.01$	$-0,55$	$-0,12$
$P-2-8-2-20$ -0.26		$-0.22$	0.00	0.00	$-0.11$	0,00	0.00	$-0.26$	$-0.08$
$ P-2-8-2-23 $ -0.13		$-0.13$	$-0.13$	$-0.03$	$-0.14$	$-0.03$	$-0.03$	$-0.14$	$-0,14$
Average	$-0,25$	$-0.16$	$-0.08$	$-0.08$	$-0.10$	$-0.02$	$-0.01$	$-0,32$	$-0,13$

Table 5.5: Results for the objectives in the literature.



Figure 5.1: Comparison of  $Obj_1$  with different  $\alpha$  values.



Figure 5.2: Comparison of  $Obj_{2.1}$  with different  $\alpha$  values.



Figure 5.3: Comparison of  $Obj_{2.2}$  with different  $\alpha$  values.



Figure 5.4: Comparison of  $Obj_{2.3}$  with different  $\alpha$  values.



Figure 5.5: Comparison of  $Obj_3$  with different  $\alpha$  values.



Figure 5.6: Comparison of all objectives.



Figure 5.7: Comparison of selected best objectives.

# **6. SUMMARY AND FUTURE WORK**

During the semester the base model is finalized, that is used to study alternative objectives to provide effective and fair solutions. We analyze the performance of alternative objectives over a set of small instances that are generated, and we can conclude that the proposed objectives can be used to obtain a trade-off between effectiveness and fairness. Since the problem is hard, we propose a ALNS within SA framework. Although parameter optimization is not performed and the heuristic is still being improved, the preliminary results are promising to solve the problem in a reasonable time.

As the future work, the following tasks will be performed in the thesis progression:

- Parameter search will be performed for SA-ALNS.
- Better and faster destroy and repair heuristics will be studied to improve solution quality.
- Repair heuristics that can guarantee or increase the chance to obtain a feasible solution will be studied.
- Objective function specific destroy and repair heuristics will be considered.
- Parallel computing will be implemented for SA-ALNS.
- Branch & Price method could be considered if it is suggested.
- Problem setting where the prepositioned items are insufficient to cover all demand will be studied.
- The value of fairness consideration in the preparedness stage will be analyzed using the routing model that will be developed.

# **REFERENCES**

- A. B. Atkinson. On the measurement of inequality. *Journal of economic theory*, 2(3): 244–263, 1970.
- B. Balcik, B. M. Beamon, and K. Smilowitz. Last mile distribution in humanitarian relief. *Journal of Intelligent Transportation Systems*, 12(2):51–63, 2008.
- B. Balcik, S. M. Iravani, and K. Smilowitz. A review of equity in nonprofit and public sector: a vehicle routing perspective. *Wiley encyclopedia of operations research and management science*, 2010.
- D. Bertsimas, V. F. Farias, and N. Trichakis. The price of fairness. *Operations research*, 59(1):17–31, 2011.
- D. Bertsimas, V. F. Farias, and N. Trichakis. On the efficiency-fairness trade-off. *Management Science*, 58(12):2234–2250, 2012.
- A. M. Campbell, D. Vandenbussche, and W. Hermann. Routing for relief efforts. *Transportation Science*, 42(2):127–145, 2008.
- E. Gralla, J. Goentzel, and C. Fine. Assessing trade-offs among multiple objectives for humanitarian aid delivery using expert preferences. *Production and Operations Management*, 23(6):978–989, 2014.
- J. Hooker. Logic-based methods for optimization combining optimization and constraint satisfaction. 2000.
- J. N. Hooker and G. Ottosson. Logic-based benders decomposition. *Mathematical Programming*, 96(1):33–60, 2003.
- K. Huang, Y. Jiang, Y. Yuan, and L. Zhao. Modeling multiple humanitarian objectives in emergency response to large-scale disasters. *Transportation Research Part E: Logistics and Transportation Review*, 75:1–17, 2015.
- M. Huang, K. Smilowitz, and B. Balcik. Models for relief routing: Equity, efficiency and efficacy. *Transportation research part E: logistics and transportation review*, 48 (1):2–18, 2012.
- J. Kalcsics, S. Nickel, J. Puerto, and A. M. Rodríguez-Chía. Several 2-facility location problems on networks with equity objectives. *Networks*, 65(1):1–9, 2015.
- Ö. Karsu and A. Morton. Inequity averse optimization in operational research. *European Journal of Operational Research*, 245(2):343–359, 2015.
- P. D. Leclerc, L. A. McLay, and M. E. Mayorga. Modeling equity for allocating public resources. In *Community-based operations research*, pages 97–118. Springer, 2012.
- Y.-H. Lin, R. Batta, P. A. Rogerson, A. Blatt, and M. Flanigan. A logistics model for emergency supply of critical items in the aftermath of a disaster. *Socio-Economic Planning Sciences*, 45(4):132–145, 2011.
- Y.-H. Lin, R. Batta, P. A. Rogerson, A. Blatt, and M. Flanigan. Location of temporary depots to facilitate relief operations after an earthquake. *Socio-Economic Planning Sciences*, 46(2):112–123, 2012.
- M. B. Mandell. Modelling effectiveness-equity trade-offs in public service delivery systems. *Management Science*, 37(4):467–482, 1991.
- M. T. Marsh and D. A. Schilling. Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research*, 74(1):1–17, 1994.
- P. Matl, R. F. Hartl, and T. Vidal. Equity objectives in vehicle routing: A survey and analysis. *arXiv preprint arXiv:1605.08565*, 2016.
- G. F. Mulligan. Equality measures and facility location. *Papers in Regional Science*, 70(4):345–365, 1991.
- W. Ogryczak, H. Luss, M. Pióro, D. Nace, and A. Tomaszewski. Fair optimization and networks: A survey. *Journal of Applied Mathematics*, 2014, 2014.
- N. Pérez-Rodríguez and J. Holguín-Veras. Inventory-allocation distribution models for postdisaster humanitarian logistics with explicit consideration of deprivation costs. *Transportation Science*, 2015.
- D. Pisinger and S. Ropke. A general heuristic for vehicle routing problems. *Computers & operations research*, 34(8):2403–2435, 2007.
- J. Rawls. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- S. Ropke and D. Pisinger. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation science*, 40(4):455– 472, 2006a.
- S. Ropke and D. Pisinger. A unified heuristic for a large class of vehicle routing problems with backhauls. *European Journal of Operational Research*, 171(3):750–775, 2006b.
- E. S. Savas. On equity in providing public services. *Management Science*, 24(8): 800–808, 1978.
- A. Sen and J. Foster. *On economic inequality*. Oxford: Clarendon Press, 1997.
- P. Shaw. Using constraint programming and local search methods to solve vehicle routing problems. In *International conference on principles and practice of constraint programming*, pages 417–431. Springer, 1998.
- G.-H. Tzeng, H.-J. Cheng, and T. D. Huang. Multi-objective optimal planning for designing relief delivery systems. *Transportation Research Part E: Logistics and Transportation Review*, 43(6):673–686, 2007.
- H. P. Young. *Equity: in theory and practice*. Princeton University Press, 1995.