University of the Aegean
Department of Financial and Management Engineering

Paratransit Services under Normal and Emergency Conditions using Public Transport Resources

Ph.D. dissertation by
Georgios Dikas

Chios, Greece
December, 2014

University of the Aegean
School of Business
Department of Financial and Management Engineering

Paratransit Services under Normal and Emergency Conditions using Public Transport Resources

Dissertation Submitted by
Georgios Dikas

in Partial Fulfillment of the Requirements for the
Doctorate Degree

Dissertation Committee

Ioannis Minis, Professor (Supervisor)

Department of Financial and Management Engineering, University of the Aegean

Agapios Platis, Associate Professor (Advisory Committee Member)

Department of Financial and Management Engineering, University of the Aegean

Stefan Voß, Professor (Advisory Committee Member)

Department of Business Administration, University of Hamburg

Dimosthenis Drivaliaris, Assistant Professor (Examination Committee Member)

Department of Financial and Management Engineering, University of the Aegean

Seraphim Kapros, Professor (Examination Committee Member)

Department of Shipping, Trade and Transport, University of the Aegean

Efstratios Papadimitriou, Professor (Examination Committee Member)

Department of Maritime Studies, University of Piraeus

Christos Tarantilis, Professor (Examination Committee Member)

Department of Management Science and Technology, Athens University of Economics and Business

Chios, 2014

Acknowledgements

Firstly, I would like to express my sincere appreciation to my advisor Professor Ioannis Minis, who patiently supervised and guided the course of this dissertation, for his substantial contribution towards my research and my growth as a scientist.

I would also like to acknowledge that I am grateful to the members of DeOPSys Lab: Dr. Theodore Athanasopoulos, Georgios Ninikas, Dr. Konstantinos Mamasis, Christina Arampantzi, Lemonia Amygdalou and Dr. Vasilis Zeimpekis for their friendship, support, encouragement, their valuable feedback and for making my everyday life easier throughout the PhD years.

Furthermore, I am sincerely grateful to Professor Agapios Platis, Associate Professor Dimosthenis Drivaliaris and Dr. Vasilis Koutras for their assistance and the accessibility they provided to me all these years. I would also like express my warm thanks to Eugenia Fronimaki, administrator for postgraduate studies, for her valuable assistance and friendship.

Last but not least, my parents, Giannis and Ntina, deserve special thanks for their continued support and encouragement.

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.

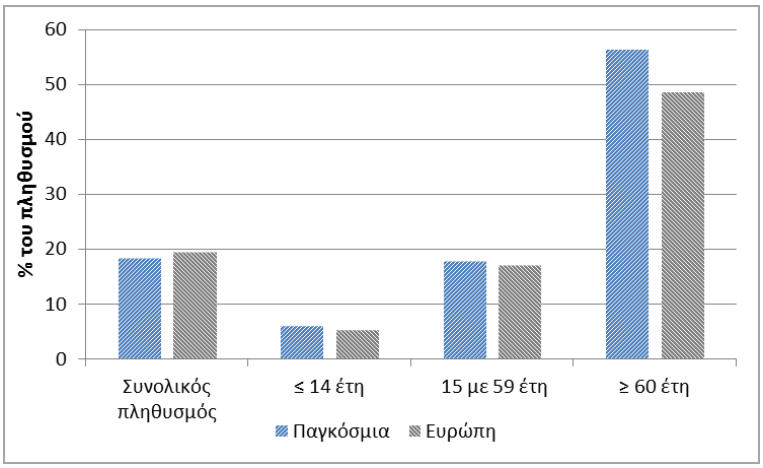
Εκτενής Περίληψη Διατριβής

Extended Summary (in Greek)

Εισαγωγή

Στην παρούσα διδακτορική διατριβή διερευνώνται υπηρεσίες μεταφοράς επιβατών με ειδικές απαιτήσεις με τη χρήση μέσων μαζικής μεταφοράς υπό κανονικές και έκτακτες συνθήκες. Το συγκεκριμένο αντικείμενο παρουσιάζει εξαιρετικό ενδιαφέρον, καθώς η ζήτηση για τις υπηρεσίες αυτές αναμένεται να αυξηθεί σημαντικά στο βραχυπρόθεσμο και μεσοπρόθεσμο μέλλον.

Συγκεκριμένα, σύμφωνα με τις εκτιμήσεις των Ηνωμένων Εθνών ενώ για το 2013 (United Nations, 2013) το ποσοστό των ατόμων άνω των 60 ετών άγγιξε το 11,7% του παγκόσμιου πληθυσμού και το 20% του πληθυσμού των Ευρωπαϊκών κρατών (World Health Organization, 2014), το ποσοστό αυτό αναμένεται να αυξηθεί σημαντικά άνω του 21% του παγκόσμιου πληθυσμού έως το 2050. Επιπλέον, σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization, 2014) η πιθανότητα εμφάνισης κάποιας μορφής αναπηρίας αυξάνεται καθώς αυξάνεται η ηλικία. Συγκεκριμένα, το Σχήμα 1 απεικονίζει τα αποτελέσματα της Έρευνας Παγκόσμιας Υγείας που διεξήχθη σε 59 κράτη από τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization, 2011) σχετικά με την εμφάνιση αναπηρίας στον γενικό πληθυσμό. Είναι εμφανές ότι το ποσοστό αναπηρίας αυξάνεται σημαντικά με την αύξηση της ηλικίας, ενώ στο σύνολο του πληθυσμού το ποσοστό αυτό ανέρχεται σε 18,2%.



Σχήμα 1: Εκτίμηση ποσοστού αναπηρίας ανά ηλικιακή ομάδα (World Health Organization, 2011)

Τα στοιχεία του Παγκόσμιου Οργανισμού Υγείας (World Health Organization, 2011) και τα αποτελέσματα της έρευνας των Avendano *et al.* (2009) αναφέρουν ότι στις Ηνωμένες

Πολιτείες, για το έτος 2004, το 59% των ανδρών και των γυναικών ηλικίας μεταξύ 50 και 70 ετών αντιμετωπίζουν κινητικές δυσκολίες, ενώ το αντίστοιχο ποσοστό για τον πληθυσμό της Ευρώπης είναι 43%.

Η δυνατότητα μετακίνησης αποτελεί βασικό παράγοντα για το επίπεδο διαβίωσης, καθώς επηρεάζει σχεδόν όλες τις δραστηριότητες της καθημερινότητας και αποτελεί σημαντική κοινωνική παροχή. Η ανάγκη προσφοράς ειδικών υπηρεσιών μεταφοράς στα άτομα με κινητικές δυσκολίες είναι πιο εμφανής κατά τη διάρκεια έκτακτων συνθηκών, όπως σε περιπτώσεις που απαιτείται η εκκένωση περιοχής λόγω πρόγνωσης επικείμενης φυσικής καταστροφής. Η σημαντική αυτή ανάγκη επιβεβαιώθηκε κατά τη διάρκεια του τυφώνα Κατρίνα που χτύπησε τη Νέα Ορλεάνη το έτος 2005. Από τους 971 θανάτους που αποδόθηκαν στον τυφώνα Κατρίνα, το 47% ήταν άτομα άνω των 75 ετών. Ενώ, λιγότερο του 10% των θυμάτων ήταν νεότεροι των 45 ετών και η μέση ηλικία όλων των θυμάτων ήταν τα 69 έτη (Brunkard *et al.*, 2008). Σε περιπτώσεις εκκένωσης περιοχών υφίσταται επίσης ανάγκη υποστήριξης της μετακίνησης ατόμων που δεν έχουν πρόσβαση σε ιδιωτικά μέσα μεταφοράς (carless population). Χαρακτηριστικά αναφέρεται ότι στη Ν. Ορλεάνη κατά τη διάρκεια του τυφώνα Κατρίνα κατοικούσαν 200.000 άτομα εξαρτώμενα από τα δημόσια μέσα μαζικής μεταφοράς (Wolshon, 2002).

Η παρούσα διατριβή αρχικά επικεντρώνεται στον σχεδιασμό εξατομικευμένων μεταφορικών υπηρεσιών με τη χρήση μέσων μαζικής μεταφοράς, και συγκεκριμένα του δημόσιου συστήματος λεωφορείων. Επιπλέον, εξετάζεται πως οι σχετικές υπηρεσίες δύνανται να βελτιωθούν ενισχύοντας το σύστημα λεωφορείων με μισθωμένα οχήματα (π.χ. ταξί). Οι εν λόγω υπηρεσίες έχουν στόχο την παροχή κινητικότητας υψηλής ποιότητας σε άτομα που χρήζουν ειδικών συνθηκών μεταφοράς, ενώ λαμβάνεται υπόψη και το σχετικό κόστος.

Επίσης, στην παρούσα διατριβή διερευνώνται περιπτώσεις εκκενώσεων περιοχών εν όψει εκτάκτων καταστάσεων. Συνήθως τα σχετικά σχέδια αντιμετώπισης εκτάκτων καταστάσεων βασίζονται σε οχήματα ιδιωτικής χρήσης (ΙΧ). Ωστόσο, αυτή η πρακτική δεν καλύπτει επαρκώς τις ανάγκες των ατόμων που εξαρτώνται από τις δημόσιες συγκοινωνίες (π.χ. των ηλικιωμένων, των ατόμων με κινητικές δυσκολίες ή προβλήματα υγείας). Για την περίπτωση αυτή, διερευνούμε το πρόβλημα του προγραμματισμού εκκένωσης περιοχών με τη χρήση υφιστάμενων δημόσιων συστημάτων λεωφορείων, λαμβάνοντας υπόψη σημαντικούς περιορισμούς που σχετίζονται με τη δυναμικότητα των διαθέσιμων καταφυγίων.

Επιπροσθέτως, διερευνούμε επέκταση του παραπάνω προβλήματος, η οποία αφορά τη μεταφορά τραυματιών σε νοσοκομειακές μονάδες, είτε μόνιμες και μεγάλης δυναμικότητας είτε προσωρινές και μικρής δυναμικότητας. Το νέο αυτό πρόβλημα αφορά καταστάσεις που προκύπτουν μετά από φυσικές καταστροφές ή τρομοκρατικές επιθέσεις. Αν και το νέο αυτό πρόβλημα έχει παρόμοιους περιορισμούς με το πρόβλημα προγραμματισμού της εκκένωσης περιοχών, διαφέρει σημαντικά ως προς τον αντικειμενικό σκοπό και τον αριθμό των εμπλεκόμενων οντοτήτων.

Ερευνητική συνεισφορά διατριβής

Όπως προαναφέρθηκε, η παρούσα διατριβή επικεντρώνεται στις μεταφορικές υπηρεσίες επιβατών με ειδικές απαιτήσεις με χρήση μέσων μαζικής μεταφοράς υπό κανονικές και έκτακτες συνθήκες. Η διδακτορική έρευνα περιλαμβάνει την ανάπτυξη νέων μαθηματικών μοντέλων για την περιγραφή των σχετικών προβλημάτων, την ανάπτυξη καινοτόμων αλγορίθμων για την επίλυση τους, καθώς και τη μελέτη των κατάλληλων παραμέτρων σχεδιασμού των σχετικών συστημάτων. Η ερευνητική συνεισφορά της διδακτορικής διατριβής περιγράφεται συνοπτικά στον Πίνακα 1.

Πίνακας 1: Ερευνητική συνεισφορά διατριβής ανά περιβάλλον εφαρμογής και περίπτωση

Περιβάλλον εφαρμογής	Περίπτωση	Συνεισφορά
Κανονικές συνθήκες	Υπηρεσίες μεταφοράς ατόμων με κινητικές δυσκολίες με χρήση δημόσιου συστήματος λεωφορείων	Ορισμός νέου προβλήματος δρομολόγησης
		Ανάπτυξη νέας μεθόδου ακριβούς επίλυσης Branch and Price
		Ανάλυση των επιπτώσεων των περιορισμών στο σχεδιασμό του μεταφορικού συστήματος
	Υπηρεσίες μεταφοράς ατόμων με κινητικές δυσκολίες με τη χρήση δημόσιου συστήματος καθώς και μισθωμένων προσβάσιμων οχημάτων	Ορισμός νέου προβλήματος δρομολόγησης
		Ανάπτυξη νέας μεθόδου ακριβούς επίλυσης Branch and Price
		Μελέτη του κόστους του μεταφορικού συστήματος σε σχέση με την παρεχόμενη ποιότητα εξυπηρέτησης

Περιβάλλον εφαρμογής	Περίπτωση	Συνεισφορά
Έκτακτες συνθήκες	Προγραμματισμός εκκένωσης περιοχών με τη χρήση δημοσίου συστήματος λεωφορείων	Ανάπτυξη νέου μαθηματικού μοντέλου με τη χρήση δύο δεικτών
		Ανάπτυξη νέου υβριδικού πλαισίου επίλυσης βασισμένο στη Δυναμική Δημιουργία Μεταβλητών (Column Generation)
		Μελέτη των επιπτώσεων του βαθμού ομαδοποίησης της ζήτησης στην ποιότητα της λύσης
		Επικύρωση της ποιότητας των λύσεων που παράγονται από το προτεινόμενο μοντέλο
	Πρόβλημα Εκκένωσης Τραυματιών	Επέκταση του παραπάνω μαθηματικού μοντέλου δύο δεικτών
		Μελέτη της επίπτωσης δύο αντικειμενικών συναρτήσεων στην ποιότητα επίλυσης
		Επικύρωση της ποιότητας των λύσεων που παράγονται από το προτεινόμενο μοντέλο
Εφαρμογή στον τομέα των logistics	Αποφάσεις επιστροφής οχήματος διανομής για αναπλήρωση αποθέματος σε περίπτωση δύο σημείων αναπλήρωσης και προκαθορισμένης σειράς επίσκεψης πελατών	Μελέτη τριών περιπτώσεων σχετικά με τη δυναμικότητα αναπλήρωσης του αποθέματος
		Ανάπτυξη νέων αλγορίθμων επίλυσης
		Μελέτη της ποιότητας λύσεων των αλγορίθμων

Υπηρεσίες μεταφοράς ατόμων με κινητικές δυσκολίες μέσω δημόσιων λεωφορειακών γραμμών

Στην παρούσα διδακτορική διατριβή διερευνήθηκαν εξατομικευμένες μεταφορικές υπηρεσίες με χρήση των υφιστάμενων λεωφορειακών γραμμών. Στο υπό μελέτη σύστημα, λεωφορείο που εκτελεί προγραμματισμένο δρομολόγιο, ενδέχεται να μεταβάλει την πορεία του ώστε να

παραλάβει επιπλέον επιβάτες με κινητικές δυσκολίες και να τους μεταφέρει στον προορισμό τους.

Για τη μελέτη του σχεδιασμού αυτού του συστήματος ορίστηκε ένα νέο πρόβλημα, το Πρόβλημα Προγραμματισμού Λεωφορείων Ευέλικτης Διαδρομής (ΠΠΛΕΔ), το οποίο διατυπώθηκε ως πρόβλημα μεικτού γραμμικού ακεραίου προγραμματισμού και επιλύθηκε με νέα μέθοδο Κλάδου και Αποτίμησης (ΚΑ, Branch and Price).

Διατύπωση μαθηματικού μοντέλου του ΠΠΛΕΔ

Οι παράμετροι του μαθηματικού μοντέλου είναι οι εξής:

- $G(N, A)$ κατευθυνόμενος γράφος, όπου N είναι το σύνολο των κόμβων και A είναι το σύνολό των ακμών ανάμεσα στου κόμβους
- m το σύνολο των στάσεων της τυπικής διαδρομής του λεωφορείου, και v ο αριθμός των δρομολογίων που εξυπηρετούν αυτές τις στάσεις για συγκεκριμένη χρονική περίοδο, όπου $K = \{1, \dots, v\}$ είναι το σύνολο των δρομολογίων της εν λόγω λεωφορειακής γραμμής
- Ένας κόμβος ορίζεται για κάθε στάση της τυπικής διαδρομής και για κάθε δρομολόγιο. Οπότε, $B = \{1, \dots, m \cdot v\}, B \subset N$ είναι το σύνολο όλων των στάσεων του γράφου. Θεωρούμε, επίσης, ότι η σειρά επίσκεψης στις στάσεις της τυπικής διαδρομής πρέπει να διατηρείται
- $B_k = \{1 + m \cdot (k-1), \dots, m + m \cdot (k-1) | \, k \in K\}$ είναι το σύνολο των στάσεων που εξυπηρετούνται κατά τη διάρκεια του δρομολογίου k
- n ο αριθμός των αιτημάτων των επιβατών με κινητικές δυσκολίες. Για κάθε αίτημα ορίζονται δύο κόμβοι, ο κόμβος παραλαβής και ο κόμβος προορισμού
- $S = \{1 + m \cdot (k-1) | \, k \in K\}$ είναι το σύνολο των αφετηριών κάθε δρομολογίου
- $E = \{m + m \cdot (k-1) | \, k \in K\}$ είναι το σύνολο των τερματικών στάσεων κάθε δρομολογίου
- $C \subset N$ είναι το σύνολο των κόμβων που αφορούν τα αιτήματα μεταφοράς των επιβατών με κινητικές δυσκολίες. Συγκεκριμένα $C = C^+ \cup C^-$, όπου $C^+ = \{m \cdot v + 1, \dots, m \cdot v + n\}$ είναι το σύνολο των κόμβων παραλαβής των επιβατών με κινητικές δυσκολίες, ενώ $C^- = \{m \cdot v + n + 1, \dots, m \cdot v + 2n\}$ είναι το αντίστοιχο σύνολο των κόμβων προορισμού. Επίσης, θεωρούμε ότι σε κάθε αίτημα μεταφοράς επιβάτη με κινητικές δυσκολίες με κόμβο παραλαβής $i \in C^+$ αντιστοιχεί ο κόμβος προορισμού $i + n \in C^-$

- Οπότε το σύνολο όλων των κόμβων του γράφου ορίζεται ως, $N = \{1,2,\dots,m \cdot v + 2n\} = B \cup C$
- $\bar{N} = \bigcup_{k \in K}\{(i,k)|i \in C \cup B_k\}$ είναι το σύνολο ζευγών κόμβων του συνόλου N και δρομολογίων του συνόλου K. Ο ορισμός των ζευγών αυτών είναι απαραίτητος για να οριστεί αν ένας κόμβος i επιπλέον αιτήματος ενδέχεται να εξυπηρετηθεί κατά τη διάρκεια του δρομολογίου k
- $\bar{A}$ είναι το σύνολο των τριάδων (i,j,k) αποτελούμενων από μία ακμή (i,j) και ένα δρομολόγιο k. Συγκεκριμένα, $\bar{A} = \bigcup_{k \in K}\{(i,j,k)|i \in (S \cap B_k), j \in \{i+1\} \cup C^+\} \cup \bigcup_{k \in K}\{(i,j,k)|i \in B_k \backslash (E \cup S), j \in \{i+1\} \cup C\} \cup \bigcup_{k \in K}\{(i,j,k)|i \in C^+, j \in (C \cup B_k)\backslash(S \cup E \cup \{i\})\} \cup \bigcup_{k \in K}\{(i,j,k)|i \in C^-, j \in (C \cup B_k)\backslash(S \cup \{i-n,i\})\}$
- $[a_i, b_i], i \in N$ είναι το χρονικό διάστημα εντός του οποίου πρέπει να εξυπηρετηθεί ο κόμβος i
- $t_{ij}, i,j \in N$ είναι ο χρόνος ταξιδίου από τον κόμβο i στον κόμβο j. Θεωρούμε ότι οι χρόνοι ταξιδίου ικανοποιούν την τριγωνική ανισότητα
- $d_i, i \in N$ ο αριθμός των επιβατών με κινητικές δυσκολίες που αντιστοιχούν σε κάθε κόμβο. Για τους κόμβους που αντιστοιχούν σε στάσεις θεωρούμε ότι $d_i = 0, i \in B$, για τους κόμβους παραλαβής θεωρούμε ότι $d_i = 1, i \in C^+$, και για τους κόμβους προορισμού θεωρούμε ότι $d_i = -1, i \in C^-$
- Q είναι ο δεσμευμένος αριθμός καθισμάτων σε κάθε λεωφορείο για τη μεταφορά των επιβατών με κινητικές δυσκολίες
- $x_{ijk}, (i,j,k) \in \bar{A}$ λαμβάνει την τιμή 1 αν η ακμή από τον κόμβο i προς τον κόμβο j διασχίζεται κατά τη διάρκεια του δρομολογίου k, αλλιώς λαμβάνει την τιμή 0
- $s_{ik}, (i,k) \in \bar{N}$ είναι ο χρόνος που εξυπηρετείται ο κόμβος i κατά τη διάρκεια του δρομολογίου k
- $q_{ik}, (i,k) \in \bar{N}$ είναι ο αριθμός των επιβατών με κινητικές δυσκολίες επί του λεωφορείου κατά το δρομολόγιο $k \in K$ αμέσως μετά την εξυπηρέτηση του κόμβου i
- $RT_{ik}, i \in C^+, k \in K$ είναι το χρονικό διάστημα που ο επιβάτης που σχετίζεται με το αίτημα μεταφοράς $i \in C^+$ εξυπηρετείται από το δρομολόγιο k
- $w_i, i \in N$ είναι δυαδική μεταβλητή που χρησιμοποιείται για τη γραμμικοποίηση του μαθηματικού μοντέλου.

Η αντικειμενική συνάρτηση του προβλήματος περιλαμβάνει δύο μέρη λεξικογραφικά ταξινομημένα. Το πρώτο μέρος αφορά τη μεγιστοποίηση του αριθμού των επιβατών με κινητικές δυσκολίες που θα εξυπηρετηθούν, ενώ το δεύτερο μέρος αφορά την

ελαχιστοποίηση του συνολικού χρόνου λειτουργίας του λεωφορειακού συστήματος. Η αντικειμενική συνάρτηση ορίζεται ως εξής:

$$TP = max \sum_{k\in K} \sum_{i\in C^+} \sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} - P \sum_{k\in K} \sum_{i\in C^+} RT_{ik} \quad (1)$$

- Η τιμή της ποινής P ορίζεται με βάση τον μέγιστο συνολικό χρόνο επιβίβασης των επιβατών με κινητικές δυσκολίες, ως $P = 1/(n \cdot T_{max})$, όπου $T_{max} = b_m - a_1 = b_{2\cdot m} - a_{m+1} = \cdots = b_{v\cdot m} - a_{(v-1)\cdot m+1}$ είναι ο μέγιστος χρόνος επιβίβασης κατά τη διάρκεια ενός δρομολογίου. Επιπλέον, οι τιμή RT_{ik} για τα αιτήματα μεταφοράς που δεν εξυπηρετούνται ισούται με 0 όπως ορίζεται παρακάτω.

Η μεγιστοποίηση της τιμής της αντικειμενική συνάρτησης υπόκειται στους εξής περιορισμούς:

$$\sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \quad i \in B_k \backslash E, k \in K \quad (2)$$

$$\sum_{\{i|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \quad j \in (E \cap B_k), k \in K \quad (3)$$

$$\sum_{\{(j,k)|(i,j,k)\in \bar{A}\}} x_{ijk} \leq 1, \quad i \in C^+ \quad (4)$$

$$\sum_{\{i|(i,h,k)\in \bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k)\in \bar{A}\}} x_{hjk} = 0, \quad h \in C \cup (B_k \backslash S \cup E), k \in K \quad (5)$$

$$\sum_{\{i|(i,h,k)\in \bar{A}\}} x_{ihk} - \sum_{\{j|(j,h+n,k)\in \bar{A}\}} x_{j,h+n,k} = 0, \quad h \in C^+, k \in K \quad (6)$$

$$s_{i-1,k} + t_{i-1,i} \leq s_{ik}, \quad i \in B_k \backslash S, k \in K \quad (7)$$

$$s_{ik} + t_{i,i+n} \sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} \leq s_{i+n,k}, \quad i \in C^+, k \in K \quad (8)$$

$$s_{i+n,k} - s_{ik} = RT_{ik}, \quad i \in C^+, k \in K \quad (9)$$

$$s_{ik} + t_{ij} - M(1 - x_{ijk}) \leq s_{jk}, \quad (i,j,k) \in \bar{A} \quad (10)$$

$$s_{jk} \leq a_j + (b_j - a_j)w_j + M(1 - x_{ijk}), \quad (i,j,k) \in \bar{A} \quad (11)$$

$$s_{jk} \leq s_{ik} + t_{ij} + a_j(1 - w_j) + M(1 - x_{ijk}), \quad (i,j,k) \in \bar{A} \quad (12)$$

$$a_i \sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} \leq s_{ik} \leq b_i \sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk}, \quad i \in C, k \in K \quad (13)$$

$$a_i \leq s_{ik} \leq b_i, \quad i \in B_k, k \in K \quad (14)$$

$$q_{ik} + d_j - M(1 - x_{ijk}) \leq q_{jk}, \ \ (i,j,k) \in \bar{A} \tag{15}$$

$$0 \leq q_{ik} \leq Q \sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk}, \ \ (i,k) \in \bar{N} \tag{16}$$

$$\sum_{i\in S\cup E} \sum_{\{k|(i,k)\in\bar{N}\}} q_{ik} = 0 \tag{17}$$

$$x_{ijk} \in \{0,1\}, \ \ (i,j,k) \in \bar{A} \tag{18}$$

$$w_i \in \{0,1\}, \ \ i \in N \tag{19}$$

$$q_{ik} \in \mathbb{N}^0, \ \ (i,k) \in \bar{N} \tag{20}$$

$$s_{ik} \geq 0, \ \ (i,k) \in \bar{N} \tag{21}$$

$$RT_{ik} \geq 0, \ \ i \in C^+, k \in K \tag{22}$$

Οι περιορισμοί (2) και (3) ορίζουν ότι οι στάσεις της τυπικής διαδρομής πρέπει να εξυπηρετηθούν μια φορά ανά δρομολόγιο. Ο περιορισμός (4) καθορίζει ότι οι κόμβοι παραλαβής των αιτημάτων των επιβατών με κινητικές δυσκολίες, και κατά συνέπεια και οι αντίστοιχοι κόμβοι τελικού προορισμού, θα εξυπηρετηθούν μια φορά ή λιγότερο. Μέσω του περιορισμού (5) διασφαλίζεται η συνέχεια μεταξύ των κόμβων, εκτός των αφετηριών και των τερματικών στάσεων. Ο περιορισμός (6) ορίζει ότι αν ένα αίτημα επιβάτη με κινητικές δυσκολίες μεταφοράς εξυπηρετηθεί, τότε οι αντίστοιχοι κόμβοι παραλαβής και προορισμού θα εξυπηρετηθούν κατά τη διάρκεια του ίδιου δρομολογίου. Ο περιορισμός (7) διασφαλίζει ότι οι στάσεις της τυπικής διαδρομής θα εξυπηρετηθούν με την προκαθορισμένη σειρά. Ο περιορισμός (8) ορίζει ότι οι κόμβοι προορισμού των αιτημάτων των επιβατών με κινητικές δυσκολίες θα εξυπηρετηθούν μετά τους αντίστοιχους κόμβους παραλαβής. Ο περιορισμός (9) ορίζει τη χρονική διάρκεια μετακίνησης του επιβάτη με κινητικές δυσκολίες που αντιστοιχεί στο επιπλέον αίτημα $i \in C^+$. Σημειώνεται ότι αν το αίτημα i δεν εξυπηρετηθεί, τότε $RT_{ik} = 0$ μέσω του περιορισμού (13), οπότε δεν επηρεάζει την τιμή της αντικειμενικής συνάρτησης. Μέσω του περιορισμού (10) ορίζεται η ροή του χρόνου, θεωρώντας ότι $M \gg 1$.

Οι περιορισμοί (11) και (12) διασφαλίσουν ότι όταν ένας κόμβος εξυπηρετείται, τότε αυτό συμβαίνει τον νωρίτερο δυνατό χρόνο. Συγκεκριμένα, οι περιορισμοί (10) έως (14) γραμμικοποιούν τη σχέση (23). Μέσω αυτής της σχέσης ο χρόνος εξυπηρέτησης ισούται με το μέγιστο της έναρξης του επιτρεπόμενου χρόνου εξυπηρέτησης του κόμβου ή του χρόνου άφιξης του οχήματος.

$$x_{ijk} = 1 \implies s_{jk} = \max\{a_j, s_{ik} + t_{ij}\}, (i,j,k) \in \bar{A} \tag{23}$$

Οι περιορισμοί (13) και (14) διασφαλίζουν ότι εάν ένα κόμβος εξυπηρετηθεί, αυτό θα συμβεί εντός του προκαθορισμένου διαστήματος $[a_i, b_i]$. Ο περιορισμός (15) ορίζει τη μεταβολή του αριθμού των επιβατών με κινητικές δυσκολίες σε κάθε δρομολόγιο. Ο περιορισμός (16) διασφαλίζει ότι ο αριθμός των επιβατών με κινητικές δυσκολίες που επιβαίνουν στο λεωφορείο είναι θετικός ή μηδέν και μικρότερος ή ίσος με τη χωρητικότητα του οχήματος Q. Ο περιορισμός (17) ορίζει ότι σε κάθε δρομολόγιο το όχημα εκκινεί και τερματίζει χωρίς κανένα επιβάτη με κινητικές δυσκολίες. Τέλος, μέσω των περιορισμών (18) και (19) ορίζεται η δυαδική φύση των μεταβλητών απόφασης x_{ijk}και w_i, μέσω του περιορισμού (20) διασφαλίζεται ότι οι μεταβλητές απόφασης q_{ik} είναι ακέραιοι αριθμοί μεγαλύτεροι ή ίσοι με το μηδέν, και μέσω των περιορισμών (21) και (22) ορίζεται ότι οι μεταβλητές απόφασης s_{ik} και RT_{ik} λαμβάνουν τιμές μεγαλύτερες ή ίσες του μηδενός.

Προτεινόμενη μέθοδος επίλυσης

Για τη βέλτιστη επίλυση του ΠΠΛΕΔ υλοποιήθηκε πλαίσιο Κλάδου και Αποτίμησης (ΚΑ, Branch and Price). Στο πλαίσιο ΚΑ ένα πρόβλημα μαθηματικού προγραμματισμού επιλύεται ενσωματώνοντας τη μέθοδο Δυναμικής Δημιουργίας Μεταβλητών (ΔΔΜ, Column Generation) σε ένα δένδρο της μεθόδου Κλάδου και Φραγής (ΚΦ, Branch and Bound, Desaulniers *et al*., 2005).

Η μέθοδος ΔΔΜ είναι τεχνική επίλυσης προβλημάτων γραμμικού προγραμματισμού, κατά την οποία μόνο ένα μέρος των μεταβλητών του προβλήματος διερευνάται. Η ΔΔΜ αποδομεί το πρόβλημα σε δύο υποπροβλήματα: το κυρίως πρόβλημα και το πρόβλημα αποτίμησης (pricing problem), τα οποία επιλύονται διαδοχικά. Το κυρίως πρόβλημα περιλαμβάνει υποσύνολο των πιθανών μεταβλητών και το πρόβλημα αποτίμησης αναζητεί κατάλληλες μεταβλητές που θα προστεθούν στο παραπάνω υποσύνολο. Για την αναγνώριση αυτών των μεταβλητών χρησιμοποιούνται οι δυικές τιμές της λύσης του κύριου προβλήματος.

Μέσω της μεθόδου ΚΦ εξασφαλίζεται ότι όσες μεταβλητές απαιτούν ακέραιες τιμές θα τις λάβουν στην τελική λύση του προβλήματος. Η ΚΦ χωρίζει τον χώρο εφικτών λύσεων σε υποχώρους θέτοντας νέους περιορισμούς (φραγμούς) με βάση τη λύση του χαλαρωμένου αρχικού προβλήματος. Κάθε φορά που τίθεται ένας νέος περιορισμός, δημιουργείται νέο πρόβλημα (κλάδος), από το οποίο ενδέχεται να προκύψουν νέοι κλάδοι. Στην περίπτωση του ΚΑ κάθε κλάδος της ΚΦ επιλύεται χρησιμοποιώντας τη μέθοδο ΔΔΜ.

Συγκεκριμένα για το ΠΠΛΕΔ, το κυρίως πρόβλημα είναι πρόβλημα δυαδικού προγραμματισμού, στο οποίο οι μεταβλητές απόφασης σχετίζονται με τη χρήση μιας

διαδρομής ή όχι. Για τη μαθηματική διατύπωση του προβλήματος θεωρούμε ότι το δ_k περιέχει όλες τις εφικτές διαδρομές του δρομολογίου k, όπου όλες οι πιθανές διαδρομές ανήκουν στο σύνολο $\delta = \bigcup_{k \in K} \delta_k$.

Το πρόβλημα αποτίμησης στοχεύει στην εύρεση της καλύτερης τιμής του μειωμένου κόστους (reduced cost), το οποίο υπολογίζεται με βάση τις δυικές τιμές της λύσης του κύριου προβλήματος και επιλύθηκε με Αλγόριθμο Ανάθεσης Ετικετών (ΑΑΕ, Labeling Algorithm). Ο ΑΕΕ υπολογίζει και αναθέτει ετικέτες σε κάθε επιμέρους μονοπάτι του σχετικού γράφου, οι οποίες περιέχουν πληροφορίες σχετικά με:

- Ποιες στάσεις έχουν εξυπηρετηθεί από το επιμέρους μονοπάτι
- Τον χρόνο άφιξης στον τελευταίο κόμβο του μονοπατιού
- Τον συνολικό χρόνο μετακίνησης των επιβατών με κινητικές δυσκολίες
- Τον αριθμό των επιβατών με κινητικές δυσκολίες οι οποίοι επιβαίνουν στο όχημα
- Το σύνολο των κόμβων παραλαβής των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες που έχουν εξυπηρετηθεί από το επιμέρους μονοπάτι
- Το σύνολο των κόμβων προορισμού των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες που έχουν εξυπηρετηθεί από το επιμέρους μονοπάτι.

Κάθε επιμέρους μονοπάτι επεκτείνεται προς όλους τους πιθανούς κόμβους, ενώ απορρίπτεται όταν παραβιαστεί κάποιος από τους περιορισμούς του προβλήματος ή υφίσταται άλλο μονοπάτι που θεωρείται κυρίαρχο. Για να θεωρηθεί ένα μονοπάτι κυρίαρχο θα πρέπει να ισχύουν οι κανόνες κυριαρχίας που αφορούν την κατανάλωση των πόρων και σχετίζονται με τις πληροφορίες των ετικετών και την τιμή του μειωμένου κόστους του μονοπατιού. Επιπροσθέτως, η δομή του ΑΕΕ που αναπτύχθηκε τροποποιήθηκε κατάλληλα ώστε να εκμεταλλεύεται ότι: (α) κάθε επιμέρους μονοπάτι είναι εφικτό μόνο όταν διατηρείται η προκαθορισμένη σειρά επίσκεψης των στάσεων της διαδρομής των λεωφορείων, και (β) οι κανόνες κυριαρχίας εφαρμόζονται μεταξύ επιμέρους μονοπατιών που έχουν αφιχθεί στην ίδια στάση της διαδρομής των λεωφορείων.

Ενδεικτικά αποτελέσματα και συμπεράσματα

Για τη μελέτη σχεδιασμού του παραπάνω μεταφορικού συστήματος, το οποίο επιτρέπει σε λεωφορεία που εκτελούν προγραμματισμένα δρομολόγια να μεταβάλλουν την τυπική διαδρομή τους, ώστε να εξυπηρετήσουν επιβάτες, διεξήχθη εκτεταμένη πειραματική διερεύνηση με τη βοήθεια του παραπάνω μοντέλου και αλγορίθμου επίλυσης ΠΠΛΕΔ.

Μέσω του προτεινόμενου νέου πλαισίου επίλυσης, ο απαιτούμενος υπολογιστικός χρόνος βελτιώθηκε κατά 50% σε σχέση με παρόμοιες μεθόδους της βιβλιογραφίας.

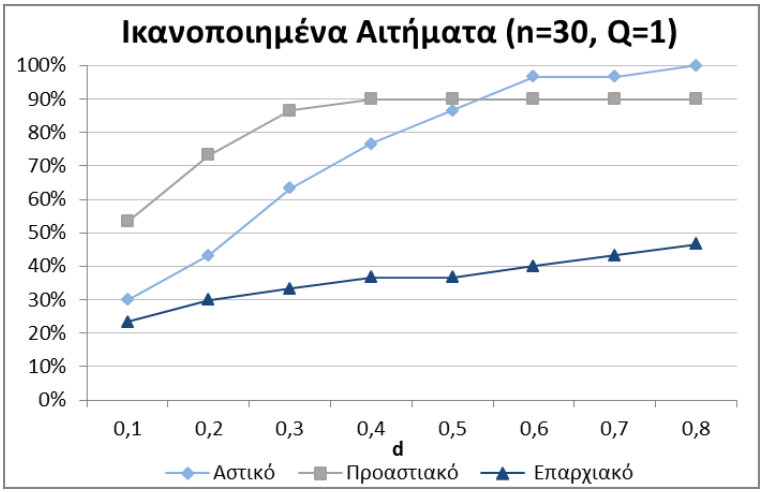
Επίσης, με βάση την πειραματική διερεύνηση, μελετήθηκαν τα εξής: (α) η επίδραση των βασικών παραμέτρων του συστήματος στην ποιότητα εξυπηρέτησης, και (β) ο σχεδιασμός και η εφαρμογή του συστήματος σε τυπικά περιβάλλοντα μεταφορών.

Η μελέτη των βασικών παραμέτρων του συστήματος ανέδειξε την ανάγκη για τον κατάλληλο προσδιορισμό των τιμών των εξής παραμέτρων:

(a) του επιτρεπόμενου χρόνου παράκαμψης των οχημάτων από την τυπική διαδρομή, που ορίζεται με βάση των συντελεστή d,
(b) τον λόγο $n/(Q \cdot k)$ του αριθμό των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες (n) προς αριθμό των δεσμευμένων θέσεων του συνόλου των δρομολογίων $(Q \cdot k)$, και
(c) τον λόγο ξ/f του επιτρεπόμενου χρόνου εξυπηρέτησης των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες (ξ) προς τη χρονοαπόσταση των δρομολογίων (f).

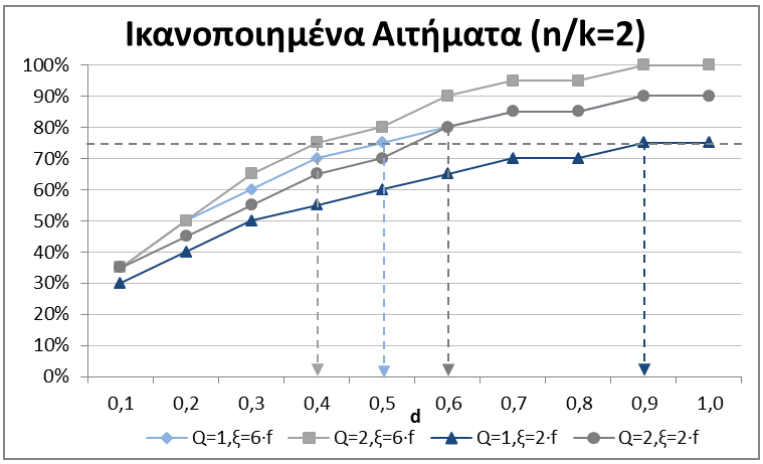
Η πειραματική διερεύνηση ανέδειξε ότι οι τιμές αυτών των παραμέτρων πρέπει να είναι κατάλληλα υψηλές ώστε να εξασφαλίζεται ένα αποδεκτό επίπεδο ποιότητας εξυπηρέτησης. Ωστόσο δεν πρέπει να ξεπερνούν ορισμένα όρια, ώστε να μην επιφέρονται σημαντικές αναταραχές στα τυπικά δρομολόγια. Επιπροσθέτως, η μελέτη διαφορετικών περιπτώσεων εφαρμογής του συστήματος (σε αστικό, προαστιακό και επαρχιακό περιβάλλον) έδειξε ότι οι τιμές των παραμέτρων θα πρέπει να διαφοροποιούνται ανάλογα με το περιβάλλον εφαρμογής.

Ενδεικτικά το Σχήμα 2 απεικονίζει την απόδοση του συστήματος για τις τρεις περιπτώσεις περιβάλλοντος εφαρμογής. Για παράδειγμα, από το Σχήμα 2 τεκμαίρεται ότι το σύστημα μπορεί να εφαρμοστεί επιτυχώς σε αστικό περιβάλλον, το οποίο χαρακτηρίζεται από υψηλές τιμές $Q/(n/k)$ και ξ/f, ακόμα και για χαμηλό επίπεδο διαταραχής της τυπικής διαδρομής (χαμηλές τιμές d).



Σχήμα 2: Επίπεδο εξυπηρέτησης για $n = 30$ και $Q = 1$ για τα τρία περιβάλλοντα εφαρμογής

Μια ενδεικτική περίπτωση αυτής της επιλογής παρουσιάζεται στο Σχήμα 3, όπου έχει οριστεί ότι το ελάχιστο αποδεκτό ποσοστό ικανοποιημένων αιτημάτων είναι 75%. Σε αυτήν την περίπτωση, ο συνδυασμός των παραμέτρων με τιμές $\xi = 2 \cdot f$, $Q = 2$ και $d = 0.6$ επιτυγχάνει ικανοποιητικό ποσοστό εξυπηρέτησης και εξασφαλίζει αποδεκτό επίπεδο ποιότητας εξυπηρέτησης και για τους δύο τύπους επιβατών.



Σχήμα 3: Επίπεδο εξυπηρέτησης από το ΠΠΛΕΔ για διαφορετικούς συνδυασμούς παραμέτρων για την περίπτωση $n/k = 2$

Υπηρεσίες μεταφοράς ατόμων με κινητικές δυσκολίες με τη χρήση δημόσιων λεωφορειακών γραμμών σε συνδυασμό με μισθωμένα προσβάσιμα οχήματα

Διερευνήθηκε επίσης η ενίσχυση του προαναφερθέντος συστήματος με μισθωμένα οχήματα, π.χ. ταξί, ώστε το ολοκληρωμένο σύστημα μεταφοράς να είναι ικανό να εξυπηρετήσει το σύνολο των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες. Σκοπός του ενισχυμένου συστήματος είναι η ελαχιστοποίηση του συνολικού κόστους μεταφοράς,

δεδομένου των περιορισμών που διασφαλίζουν την ποιότητα της εξυπηρέτησης και στα δυο μεταφορικά μέσα.

Όπως και προηγουμένως, τα μεταφορικά αιτήματα είναι γνωστά προ της έναρξης της περιόδου εξυπηρέτησης και κάθε αίτημα περιλαμβάνει: (α) τις επιθυμητές τοποθεσίες παραλαβής και προορισμού, και (β) τον αργότερο χρόνο άφιξης στον προορισμό (ή τον νωρίτερο χρόνο παραλαβής). Σημειώνεται, ότι το μεταφορικό μέσο δεν είναι γνωστό τη στιγμή που εγείρεται το αίτημα. Επιπλέον, ορίζονται κατάλληλοι χρονικοί περιορισμοί για να διασφαλιστεί ότι στους επιβάτες θα προσφερθεί η ίδια ποιότητα εξυπηρέτησης ανεξαρτήτως του μέσου μεταφοράς.

Για τη μελέτη του σχεδιασμού του ολοκληρωμένου συστήματος ορίστηκε το Πρόβλημα Προγραμματισμού Λεωφορείων Ευέλικτης Διαδρομής σε Συνδυασμό με Μισθωμένα Οχήματα (για λόγους απλότητας αναφερόμαστε σε αυτό το πρόβλημα ως ΠΠΛΕΔ-Τ), το οποίο διατυπώθηκε ως πρόβλημα μεικτού γραμμικού ακεραίου προγραμματισμού και επιλύθηκε με νέα μέθοδο Κλάδου και Αποτίμησης για τη βέλτιστη επίλυση του.

Διατύπωση μαθηματικού μοντέλου ΠΠΛΕΔ -Τ

Οι παράμετροι του μαθηματικού μοντέλου είναι οι εξής:

- $G(N, A)$ είναι κατευθυνόμενος γράφος, όπου N είναι το σύνολο όλων των κόμβων και A είναι το σύνολο των ακμών που ενώνουν τους κόμβους
- B είναι το σύνολο των κόμβων που σχετίζονται με τις στάσεις της τυπικής διαδρομής και αποτελείται από τα υποσύνολα κόμβων B_k, καθένα από τα οποία αφορά τις στάσεις κάθε δρομολογίου $k \in K$, συνεπώς $B = \bigcup_{k \in K} B_k$, όπου K είναι το σύνολο των δρομολογίων των λεωφορείων. Επιπλέον, όπως και προηγουμένως, το σύνολο $S \subseteq B$ περιέχει όλες τις αφετηρίες των δρομολογίων, και το σύνολο $E \subseteq B$ το περιέχει τις αντίστοιχες τερματικές στάσεις
- $C = C^+ \cup C^-$ είναι, όπως και προηγουμένως, το σύνολο των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες, όπου το C^+ περιέχει τους κόμβους παλαβής και το C^- τους αντίστοιχους κόμβους προορισμού. Επίσης, θεωρούμε ότι για κάθε αίτημα μεταφοράς επιβάτη με κινητικές δυσκολίες με κόμβο παραλαβής $i \in C^+$ αντιστοιχεί ο κόμβος προορισμού $i + n \in C^-$, όπου n είναι ο συνολικός αριθμός των αιτημάτων μεταφοράς των επιβατών με κινητικές δυσκολίες.

Για το ΠΠΛΕΔ-Τ ορίζουμε το σύνολο V των μισθωμένων οχημάτων - ταξί, και για λόγους διατύπωσης του μαθηματικού μοντέλου, θωρούμε ότι υπάρχει τουλάχιστον ένα διαθέσιμο

ταξί για κάθε αίτημα μεταφοράς επιβάτη με κινητικές δυσκολίες. Επιπρόσθετα, νέοι κόμβοι ορίζονται σχετικά με τα σημεία εκκίνησης και τερματισμού των διαδρομών των ταξί $\bar{S} = \{\bar{S}_k | k \in V\}$ και $\bar{E} = \{\bar{E}_k | k \in V\}$, αντίστοιχα. Οπότε, το σύνολο N ορίζεται πλέον από την ένωση των παραπάνω συνόλων κόμβων, δηλαδή $N = B \cup C \cup \bar{S} \cup \bar{E}$.

Το σύνολο $\bar{A}$ περιέχει τριάδες που ορίζονται από τον συνδυασμό των ακμών του A και των δρομολογίων του K, ή των ταξί του V. Οι δυνατές τριάδες του $\bar{A}$ ορίζονται με βάση τις απαιτήσεις του προβλήματος ως εξής:

$\bar{A} = \cup_{k\in K}\{(i,j,k)|i \in (S \cap B_k), j \in (\{i+1\} \cup C^+)\} \cup \cup_{k\in K}\{(i,j,k)|i \in B_k \backslash (E \cup S), j \in (\{i+1\} \cup C)\} \cup \cup_{k\in K}\{(i,j,k)|i \in C^+, j \in (C \cup B_k)\backslash(S \cup E \cup \{i\})\} \cup \cup_{k\in K}\{(i,j,k)|i \in C^-, j \in (C \cup B_k)\backslash(S \cup \{i-n, i\})\} \cup \cup_{k\in V}\{(i,j,k)|i \in \bar{S}_k, j \in (C^+ \cup \bar{E}_k\} \cup \cup_{k\in V}\{(i,j,k)|i \in C^+, j \in C\backslash\{i\}\} \cup \cup_{k\in V}\{(i,j,k)|i \in C^-, j \in (C \cup \bar{E}_k)\backslash\{i-n, i\}\}$

Επιπρόσθετα, ορίζονται τα εξής:

- $[a_i, b_i], i \in N\backslash\{\bar{S} \cup \bar{E}\}$ είναι το χρονικό διάστημα εντός του οποίου οι κόμβοι της ένωσης των συνόλων $C \cup B$ πρέπει να εξυπηρετηθούν σε περίπτωση που εξυπηρετούνται από τα λεωφορεία του συστήματος, δηλαδή από $k \in K$
- $[\hat{a}_i, \hat{b}_i], i \in C$ είναι το χρονικό διάστημα εντός του οποίου κάθε κόμβος πρέπει να εξυπηρετηθεί σε περίπτωση που εξυπηρετείται από ταξί, δηλαδή $k \in V$
- $d_i, i \in N$ είναι ο αριθμός των επιβατών με κινητικές δυσκολίες που αντιστοιχούν σε κάθε κόμβο. Θεωρούμε ότι $d_i = 0$ για $i \in B \cup \bar{S} \cup \bar{E}$. Για τους κόμβους παραλαβής θεωρούμε ότι $d_i = 1, i \in C^+$, και για τους κόμβους προορισμού θεωρούμε ότι $d_i = -1, i \in C^-$
- Q είναι η χωρητικότητα κάθε λεωφορείου για επιβάτες με κινητικές δυσκολίες
- $\bar{Q}$ είναι η χωρητικότητα κάθε ταξί για επιβάτες με κινητικές δυσκολίες
- $T_{ij}, i, j \in N$ είναι ο χρόνος ταξιδίου από τον κόμβο i στον κόμβο j. Θεωρούμε ότι οι χρόνοι ταξιδίου ικανοποιούν την τριγωνική ανισότητα

Οι μεταβλητές απόφασης του ΠΠΛΕΔ-Τ ορίζονται ως εξής:

- $x_{ijk}, (i,j,k) \in \bar{A}$ λαμβάνει την τιμή 1 αν η ακμή από τον κόμβο i προς τον κόμβο j διασχίζεται από λεωφορείο κατά τη διάρκεια του δρομολογίου k, ή από ταξί. Αλλιώς λαμβάνει την τιμή 0
- $s_i, i \in N$ είναι ο χρόνος που εξυπηρετείται ο κόμβος i

- $q_i, i \in N$ είναι ο αριθμός των επιβατών με κινητικές δυσκολίες που επιβαίνουν, είτε στο λεωφορείο κατά το ταξίδι $k \in K$, είτε στο ταξί $k \in V$, αμέσως μετά την εξυπηρέτηση του κόμβου i
- $w_i, i \in N \setminus (\bar{S} \cup \bar{E})$ είναι μια δυαδική μεταβλητή που χρησιμοποιείται για τη γραμμικοποίηση του μαθηματικού μοντέλου.

Η αντικειμενική συνάρτηση του προβλήματος στοχεύει στην ελαχιστοποίηση των συνιστωσών του κόστους μεταφοράς. Τα κόστη αυτά περιλαμβάνουν: (α) τη συνολική απόσταση που καλύπτεται από τα ταξί, και (β) την επιπλέον απόσταση, εκτός της τυπικής διαδρομής, που καλύπτεται από τα λεωφορεία κατά τη διάρκεια των εκτελούμενων δρομολογίων. Επιπλέον, η τελευταία απόσταση πολλαπλασιάζεται με τον συντελεστή a, οποίος χρησιμοποιείται για να εκφράσει το σχετικά υψηλότερο κόστος λειτουργίας των λεωφορείων σε σχέση με αυτό των ταξί, καθώς και την ενδεχόμενη απώλεια εσόδων της υπηρεσίας λεωφορείων λόγω του επιπλέον χρόνου ταξιδίου που οφείλεται στις απαραίτητες παρακάμψεις του δρομολογίου.

Κατά συνέπεια, η αντικειμενική συνάρτηση ορίζεται ως εξής:

$$TC = min \sum_{\{(i,j,k)|k \in V,(i,j,k) \in \bar{A}\}} c_{ij} x_{ijk} + \sum_{\{(i,j,k)|k \in K,(i,j,k) \in \bar{A}\}} a c_{ij} x_{ijk} - a|K|F \tag{24}$$

όπου F είναι το κόστος για την ολοκλήρωση ενός τυπικού δρομολογίου χωρίς παρακάμψεις. Η ελαχιστοποίηση της (24) υπόκειται στους περιορισμούς (25) έως (46), κατωτέρω:

$$\sum_{\{j|(i,j,k) \in \bar{A}\}} x_{ijk} = 1, \quad i \in B_k \setminus E, k \in K \tag{25}$$

$$\sum_{\{i|(i,j,k) \in \bar{A}\}} x_{ijk} = 1, \quad j \in (E \cap B_k), k \in K \tag{26}$$

$$\sum_{\{j|(i,j,k) \in \bar{A}\}} x_{ijk} = 1, \quad i \in \bar{S}, k \in V \tag{27}$$

$$\sum_{\{i|(i,j,k) \in \bar{A}\}} x_{ijk} = 1, \quad j \in \bar{E}, k \in V \tag{28}$$

$$\sum_{\{(j,k)|(i,j,k) \in \bar{A}\}} x_{ijk} = 1, \quad i \in C^+ \tag{29}$$

$$\sum_{\{i|(i,h,k) \in \bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k) \in \bar{A}\}} x_{hjk} = 0, \quad h \in B_k \setminus (S \cup E), k \in K \tag{30}$$

$$\sum_{\{i|(i,h,k)\in\bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k)\in\bar{A}\}} x_{hjk} = 0, \quad h \in C, k \in K \cup V \tag{31}$$

$$\sum_{\{i|(i,h,k)\in\bar{A}\}} x_{ihk} - \sum_{\{j|(j,h+n,k)\in\bar{A}\}} x_{j,h+n,k} = 0, \quad h \in C^+, k \in K \cup V \tag{32}$$

$$s_{i-1} + T_{i-1,i} \le s_i, \quad i \in B_k \backslash S, k \in K \tag{33}$$

$$s_i + T_{i,i+n} \le s_{i+n}, \quad i \in C^+ \tag{34}$$

$$s_i + T_{ij} - M(1 - x_{ijk}) \le s_j, \quad (i,j,k) \in \bar{A} \tag{35}$$

$$s_j \le a_j + (b_j - a_j)w_j + M(1 - x_{ijk}), \quad k \in K, (i,j,k) \in \bar{A} \tag{36}$$

$$s_j \le s_i + t_{ij} + a_j(1 - w_j) + M(1 - x_{ijk}), \quad k \in K, (i,j,k) \in \bar{A} \tag{37}$$

$$\sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}\,\hat{a}_i + \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk}\,a_i \le s_i, \quad i \in C \tag{38}$$

$$s_i \le \sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}\,\hat{b}_i + \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk}\,b_i, \quad i \in C \tag{39}$$

$$a_i \le s_i \le b_i, \quad i \in B \tag{40}$$

$$q_i + d_j - M(1 - x_{ijk}) \le q_j, \quad (i,j,k) \in \bar{A} \tag{41}$$

$$q_i \le Q \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk} + \bar{Q} \sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}, \quad i \in N\backslash\{S \cup E \cup \bar{S} \cup \bar{E}\} \tag{42}$$

$$\sum_{i\in S\cup E\cup\bar{S}\cup\bar{E}} q_i = 0 \tag{43}$$

$$x_{ijk} \in \{0,1\}, \quad (i,j,k) \in \bar{A} \tag{44}$$

$$w_i \in \{0,1\}, \quad i \in N(\bar{S} \cup \bar{E}) \tag{45}$$

$$q_i \in \mathbb{N}^0, \quad i \in N \tag{46}$$

Οι περιορισμοί (25) και (26) ορίζουν ότι οι στάσεις της τυπικής διαδρομής εξυπηρετούνται άπαξ ανά δρομολόγιο. Οι περιορισμοί (27) και (28) διασφαλίζουν ότι όλα τα διαθέσιμα ταξί θα εκκινήσουν από την τοποθεσία αναμονής και θα επιστρέψουν στην ίδια ή άλλη τοποθεσία αναμονής. Ο περιορισμός (29) καθορίζει ότι οι κόμβοι παραλαβής των επιβατών με κινητικές δυσκολίες, και κατά συνέπεια οι αντίστοιχοι κόμβοι τελικού προορισμού, θα εξυπηρετηθούν μια φορά. Μέσω των περιορισμών (30) και (31) διασφαλίζεται η συνέχεια εξυπηρέτησης των κόμβων, εκτός των αφετηριών και των τερματικών στάσεων των λεωφορείων, και των κόμβων εκκίνησης και τερματισμού των ταξί. Ο περιορισμός (32) εξασφαλίζει ότι οι κόμβοι

παραλαβής και προορισμού κάθε επιβάτη με κινητικές δυσκολίες εξυπηρετείται από το ίδιο δρομολόγιο ή από το ίδιο ταξί.

Ο περιορισμός (33) διασφαλίζει ότι οι στάσεις της τυπικής διαδρομής θα εξυπηρετηθούν με την προκαθορισμένη σειρά. Ο περιορισμός (34) ορίζει ότι οι κόμβοι προορισμού των επιβατών με κινητικές δυσκολίες θα εξυπηρετηθούν μετά τους αντίστοιχους κόμβους παραλαβής. Μέσω του περιορισμού (35) ορίζεται η ροή του χρόνου, θεωρώντας ότι $M \gg 1$. Οι περιορισμοί (36) και (37) διασφαλίσουν ότι όταν ένας κόμβος εξυπηρετείται, τότε αυτό συμβαίνει τον νωρίτερο δυνατό χρόνο, όπως και στην περίπτωση του ΠΠΛΕΔ.

Οι περιορισμοί (38) και (39) διασφαλίζουν ότι κάθε κόμβος που αντιστοιχεί σε επιβάτη με κινητικές δυσκολίες, εξυπηρετείται εντός του προκαθορισμένου διαστήματος. Συγκεκριμένα, ο περιορισμός (38) διασφαλίζει ότι η εξυπηρέτηση του κόμβου θα αρχίσει το νωρίτερο τη χρονική στιγμή $\hat{a}_i$ αν εξυπηρετηθεί από ταξί ή τη χρονική στιγμή a_i αν εξυπηρετηθεί από δρομολόγιο λεωφορείου. Αντίστοιχα, ο περιορισμός (39) διασφαλίζει ότι η εξυπηρέτηση του κόμβου θα ολοκληρωθεί προ της χρονικής στιγμής $\hat{b}_i$ αν εξυπηρετηθεί από ταξί ή προ της χρονικής στιγμής b_i αν εξυπηρετηθεί από δρομολόγιο λεωφορείου. Ο περιορισμός (40) ορίζει την επιτρεπόμενη χρονική περίοδο κατά την οποία πρέπει να εξυπηρετηθούν οι στάσεις του τυπικού δρομολογίου.

Ο περιορισμός (41) ορίζει τη μεταβολή του αριθμού των επιβατών με κινητικές δυσκολίες, στο όχημα που εξυπηρετεί των κόμβο i. Ο περιορισμός (42) διασφαλίζει ότι ο αριθμός των επιβατών με κινητικές δυσκολίες που επιβαίνουν στο όχημα, είτε λεωφορείο είτε ταξί, θα είναι μικρότερος ή ίσος με Q ή $\bar{Q}$, ανάλογα με ποιο μέσο εξυπηρετεί τον κόμβο i. Ο περιορισμός (43) ορίζει ότι κάθε όχημα εκκινεί και τερματίζει χωρίς επιβάτες με κινητικές δυσκολίες.

Τέλος, μέσω των περιορισμών (44) και (45) ορίζεται η δυαδική φύση των μεταβλητών απόφασης x_{ijk} και w_iκαι μέσω του περιορισμού (46) διασφαλίζεται ότι οι μεταβλητές απόφασης q_i είναι ακέραιοι αριθμοί μεγαλύτεροι ή ίσοι με το μηδέν.

Προτεινόμενη μέθοδος επίλυσης

Για τη βέλτιστη επίλυση του ΠΠΛΕΔ-Τ υλοποιήθηκε νέο πλαίσιο ΚΑ. Επιπλέον η προτεινόμενη μέθοδος επίλυσης χρησιμοποιήθηκε για να μελετηθεί πειραματικά η απόδοση του συστήματος. Συγκεκριμένα για το ΠΠΛΕΔ-Τ το κυρίως πρόβλημα είναι πρόβλημα δυαδικού προγραμματισμού, στο οποίο οι μεταβλητές απόφασης σχετίζονται με τη χρήση δρομολόγιου λεωφορείου /διαδρομής ταξί ή όχι, λαμβάνοντας υπόψη όλα τα πιθανά

δρομολόγια/διαδρομές και των δυο διαθέσιμων μέσων μεταφοράς. Για τη μαθηματική διατύπωση του προβλήματος θεωρούμε ότι το δ_k περιέχει όλες τις πιθανές εφικτές διαδρομές του δρομολογίου k, ενώ το σύνολο $\delta_{|K|+1}$ περιέχει όλες τις πιθανές διαδρομές των ταξί. Συνεπώς όλες οι πιθανές, εφικτές διαδρομές και των δύο μέσων ανήκουν στο σύνολο $\delta = \bigcup_{k \in K} \delta_k \cup \delta_{|K|+1}$.

Το πρόβλημα αποτίμησης στοχεύει στην εύρεση της καλύτερης τιμής του μειωμένου κόστους (reduced cost) το οποίο υπολογίζεται με βάση τις δυικές τιμές της λύσης του κύριου προβλήματος. Στην περίπτωση του ΠΠΛΕΔ-Τ για την επίλυση του προβλήματος αποτίμησης συνδυάστηκαν οι διαδρομές που παράγονται από δύο διαφορετικούς ΑΑΕ: (α) για την αναζήτηση διαδρομών των λεωφορείων χρησιμοποιήθηκε ο ΑΑΕ που αναπτύχθηκε για το ΠΠΛΕΔ, στον οποίο δεν λαμβάνονται υπόψη οι πληροφορίες σχετικά με τον συνολικό χρόνο μετακίνησης των επιβατών με κινητικές δυσκολίες, και (β) για την αναζήτηση των διαδρομών ταξί χρησιμοποιήθηκε ο ΑΑΕ των Ropke and Cordeau (2009), που σχεδιάστηκε για το πρόβλημα αποτίμησης του Προβλήματος Παραλαβής και Παράδοσης με Χρονικά Παράθυρα Εξυπηρέτησης (Pick-up and Delivery Problem with Time Windows).

Ενδεικτικά αποτελέσματα και συμπεράσματα

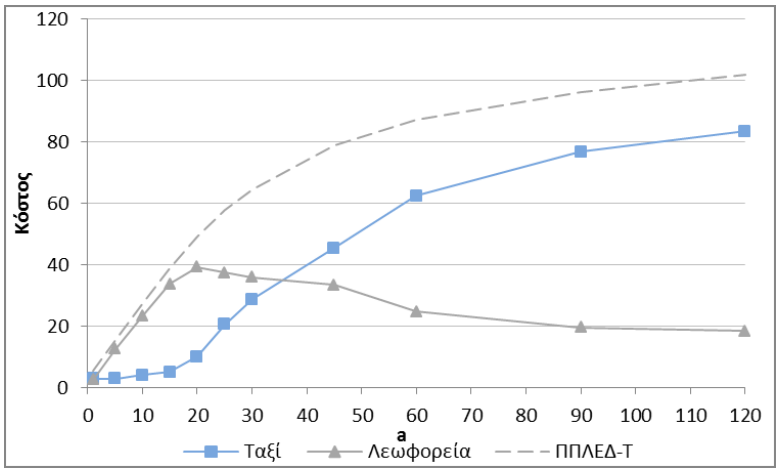
Η μελέτη σχεδιασμού του παραπάνω συστήματος έγινε μέσω εκτενούς πειραματικής διερεύνησης, στην οποία χρησιμοποιήθηκε το μοντέλο και ο αλγόριθμος επίλυσης του ΠΠΛΕΔ-Τ. Η πειραματική διερεύνηση υπέδειξε ότι το δημόσιο σύστημα λεωφορείων μεταφέρει άτομα με κινητικές δυσκολίες και, συνεπώς, συνεισφέρει στο συνολικό κόστος ακόμα και όταν το κόστος λειτουργίας του συστήματος λεωφορείων είναι υψηλό σε σχέση με αυτό των ταξί (υψηλές τιμές του συντελεστή α). Το συμπέρασμα αυτό είναι εμφανές στο Σχήμα 4, στο οποίο απεικονίζεται ο καταμερισμός του κόστους στα δυο μεταφορικά μέσα και το αντίστοιχο συνολικό κόστος του συστήματος για την περίπτωση όπου οι περιορισμοί που διασφαλίζουν την ποιότητα εξυπηρέτησης, πρακτικά, δεν επηρεάζουν το σύστημα.

Επίσης, σχετικά με την ανάπτυξη του συστήματος έχοντας ορίσει την τιμή του συντελεστή a και την τιμή της χωρητικότητας των λεωφορείων για επιβάτες με κινητικές δυσκολίες Q, ο σχεδιαστής θα πρέπει να επικεντρωθεί στις τιμές των εξής παραμέτρων:

(a) d - καθορίζει τον επιτρεπόμενο χρόνο παράκαμψης του κανονικού δρομολογίου λεωφορείου, καθώς και τον μέγιστο χρόνο επιβίβασης των πελατών που θα εξυπηρετηθούν με ταξί

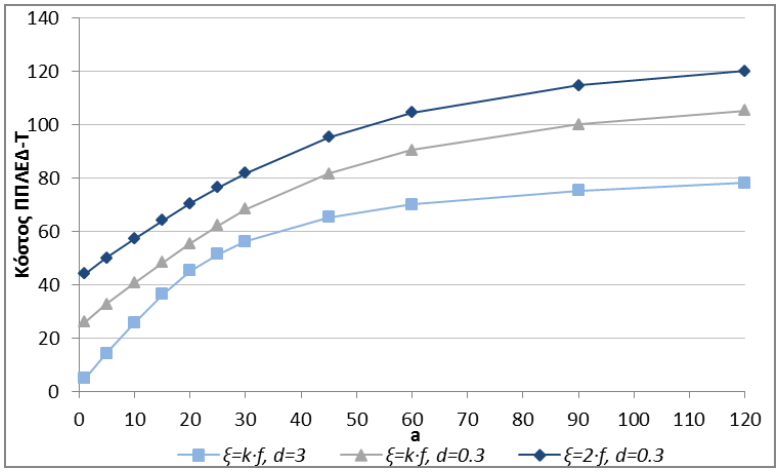
(b) ξ - επιτρεπόμενος χρόνος εξυπηρέτησης των επιβατών με κινητικές δυσκολίες και

(c) τη χωρητικότητα των ταξί.



Σχήμα 4: Κόστος ανά μεταφορικό μέσο για αυξανόμενες τιμές του συντελεστή a και για ξ=υψηλό, d= υψηλό και $\bar{Q} = 1$

Στο Σχήμα 5 εμφανίζονται τρεις συνδυασμοί των παραμέτρων d και ξ για χωρητικότητα ταξί $\bar{Q} = 2$. Είναι εμφανές ότι για χαμηλότερο επίπεδο ποιότητας εξυπηρέτησης, το συνολικό κόστος του μεταφορικού συστήματος είναι χαμηλότερο για όλες τις τιμές του συντελεστή a. Επιπροσθέτως, αυξάνοντας τις τιμές των σχετικών παραμέτρων άνω κάποιων ορίων, με αντίστοιχη μείωση της ποιότητας εξυπηρέτησης, δεν βελτιώνεται σημαντικά το αντίστοιχο κόστος του συστήματος.



Σχήμα 5: Κόστος του μεταφορικού συστήματος για τρεις περιπτώσεις του επιπέδου ποιότητας εξυπηρέτησης, όπου $\bar{Q} = 2$ και για αυξανόμενες τιμές της παραμέτρου a

Η μελέτη σχεδιασμού ανέδειξε, επίσης, τη μέθοδο καθορισμού των τιμών των παραμέτρων του συστήματος ώστε να επιτευχθεί ισορροπία μεταξύ της ποιότητας εξυπηρέτησης και του κόστους του συστήματος.

Προγραμματισμός εκκένωσης περιοχών με τη χρήση λεωφορείων δημόσιων συγκοινωνιών

Αρχικά αντιμετωπίσαμε την περίπτωση εκκένωσης περιοχής λόγω επικείμενης απειλής. Συγκεκριμένα μελετήσαμε το Πρόβλημα Προγραμματισμού Εκκενώσεως Περιοχής με τη Χρήση Λεωφορείων (ΠΠΕΠΧΛ, Bus Evacuation Problem, BEP) που αρχικά προτάθηκε από τον Bish (2011).

Στο ΠΠΕΠΧΛ, στόλος λεωφορείων με ίδια χαρακτηριστικά καλείται να παραλάβει πολίτες από προκαθορισμένες τοποθεσίες και να τους μεταφέρει στα διαθέσιμα καταφύγια. Κάθε καταφύγιο μπορεί να φιλοξενήσει συγκεκριμένο αριθμό πολιτών. Ο αντικειμενικός σκοπός του προβλήματος είναι να ολοκληρωθεί η διαδικασία στον ελάχιστο χρόνο εκκένωσης. Ο τελευταίος ορίζεται ως η περίοδος από τη χρονική στιγμή που το πρώτο λεωφορείο εκκινεί από την αφετηρία του έως τη χρονική στιγμή που ο τελευταίος πολίτης θα μεταφερθεί σε καταφύγιο. Επίσης, ανάμεσα στις πιθανές λύσεις με τον ίδιο χρόνο εκκένωσης προτιμάται η λύση με το ελάχιστο κόστος λειτουργίας.

Αν και το ΠΠΕΠΧΛ αναφέρεται σε περιπτώσεις εκκένωσης εν όψει απειλής, παρόμοιες συνθήκες επικρατούν στη μεταφορά μετά την εκδήλωση της απειλής ή άλλου γεγονότος, όπως η μεταφορά τραυματιών σε νοσοκομειακές μονάδες ως αποτέλεσμα πολεμικής σύγκρουσης. Οι νοσοκομειακές μονάδες ενδέχεται να είναι μόνιμες εγκαταστάσεις μεγάλης δυναμικότητας ή προσωρινές μονάδες μικρής δυναμικότητας. Στα πλαίσια της διδακτορικής έρευνας μελετήσαμε το σχετικό Πρόβλημα Εκκένωσης Τραυματιών (ΠΕΤ).

Διατύπωση μαθηματικού μοντέλου ΠΠΕΠΧΛ

Θεωρείστε ότι $M = \{1, \dots, \mu\}$ είναι το σύνολο των διαθέσιμων καταφυγίων, το καθένα δυναμικότητας $P^m, m \in M$, και ότι $K = \{1, \dots, v\}$ είναι το σύνολο των διαθέσιμων οχημάτων (λεωφορείων), το καθένα χωρητικότητας Q. Όλα τα οχήματα εκκινούν και τερματίζουν τις διαδρομές τους από ένα σταθμό. Για τη διατύπωση του μοντέλου με δύο δείκτες ορίζουμε ένα σταθμό εκκίνησης και ένα σταθμό τερματισμού για κάθε όχημα μέσω των συνόλων S και E, όπου $S = \{s^k | k \in K\}$ και $E = \{e^k | k \in K\}$. Σημειώνεται, ότι αν θεωρήσουμε ότι όλες οι αποστάσεις προς τους σταθμούς του E είναι μηδενικές τότε εξετάζουμε την περίπτωση που μελετήθηκε από τον Bish (2011).

Θεωρήστε ότι C είναι το σύνολο των σημείων αναμονής των πολιτών προς μεταφορά. Τα σημεία αυτά ορίζονται ως σημεία παραλαβής. Επιπλέον, θεωρήστε $d_i, i \in C \cup S$ τον αριθμό

των πολιτών στα σημεία παραλαβής, θεωρώντας ότι $d_i = 0, i \in S$ και $d_i \in \mathbb{N}^+, i \in C$. Σημειώνεται ότι εξετάζουμε την περίπτωση για την οποία σε κάθε σημείο παραλαβής ο αριθμός των πολιτών που περιμένουν είναι μικρότερος ή ίσος της χωρητικότητας του λεωφορείου ($d_i \leq Q, i \in C$). Στην περίπτωση που ο αριθμός των πολιτών σε αναμονή είναι μεγαλύτερος της χωρητικότητας του λεωφορείου, δημιουργούμε πολλαπλά σημεία παραλαβής με βάση τον αριθμό των πολιτών σε αναμονή.

Για τη μαθηματική διατύπωση των πιθανών πολλαπλών επισκέψεων των οχημάτων στα καταφύγια, στο πλαίσιο των δυο δεικτών, ορίζεται ένα νέο σύνολο κόμβων $D = \{D_i^m | i \in C \cup E, m \in M\}$. Το σύνολο D περιέχει πολλαπλούς *ψευδοκόμβους* που ορίζονται για κάθε σημείο παραλαβής ή τερματικό σταθμό και κάθε διαθέσιμο καταφύγιο. Επίσης, το σύνολο των δυνατών ακμών ορίζεται ως εξής:

$$A = \{(i,j) | i \in S, j \in C \cup E\} \cup \{(i,j) | i \in C, j \in (C \backslash \{i\}) \cup (D \backslash \{D_i^m | m \in M\})\} \cup \{(i,j) | i = D_h^t, h \in C \cup E, t \in M, j \in \{h\} \cup \{D_h^m | m \in M \backslash \{t\}\}\}$$

Οπότε ένα καταφύγιο ενδέχεται να δεχτεί πολίτες από τα οχήματα το πολύ $|C| + |E|$ φορές.

Συνεπώς, ορίζεται ο κατευθυνόμενος γράφος $G(V, A)$ με σύνολο κόμβων $V = S \cup E \cup C \cup D$.

Θεωρήστε ότι t_{ij} είναι χρόνος που απαιτείται ώστε ένα όχημα να διασχίσει την ακμή $(i,j) \in A$, με τους χρόνους να ικανοποιούν την τριγωνική ανισότητα. Θεωρήστε, επίσης, τις παρακάτω μεταβλητές απόφασης, όπου:

- $w_i, i \in V$ είναι ο χρόνος άφιξης ενός οχήματος στον κόμβο i. Σημειώνεται ότι στο πλαίσιο της μαθηματικής διατύπωσης δύο δεικτών ένας κόμβος ενδέχεται να εξυπηρετηθεί από οποιοδήποτε όχημα, αλλά μόνο μια φορά από ένα όχημα
- $q_i, i \in V$ είναι ο αριθμός των πολιτών που επιβαίνουν στο όχημα που εξυπηρετεί τον κόμβο i αμέσως πριν την εξυπηρέτηση του
- $d_i^m, i \in C \cup E, m \in M$ είναι ο αριθμός των πολιτών που αποβιβάζονται στο καταφύγιο m όταν ένα όχημα επισκέπτεται τον ψευδοκόμβο D_i^m
- x_{ij} λαμβάνει την τιμή 1 αν η ακμή $(i,j) \in A$ διασχίζεται από κάποιο όχημα, αλλιώς λαμβάνει την τιμή 0
- T_{evac} είναι ο συνολικός χρόνος εκκένωσης.

Η αντικειμενική συνάρτηση του ΠΠΕΠΧΛ ορίζεται ως εξής:

$$TC = \min\{T_{evac} + \frac{1}{L}\sum_{i \in E} w_i\} \quad (47)$$

όπου το δεύτερο μέρος είναι ο συνολικός χρόνος (κόστος) λειτουργίας των οχημάτων και το L διασφαλίζει ότι το πρώτος μέρος της συνάρτησης κυριαρχεί του δεύτερου (υπό λεξικογραφική έννοια). Αυτό επιτυγχάνεται θέτοντας την τιμή του L να υπερβαίνει τον συνολικό χρόνο λειτουργίας: $L > \sum_{(i,j)\in A} t_{ij}$. Οι περιορισμοί του προβλήματος έχουν ως εξής:

$$T_{evac} \geq w_j - \sum_{i \in V|(i,j)\in A} t_{ij} x_{ij},\ j \in E \quad (48)$$

$$\sum_{j \in V|(i,j)\in A} x_{ij} = 1,\ i \in C \cup S \quad (49)$$

$$\sum_{j \in V|(i,j)\in A} x_{ij} \leq 1,\ i \in D \quad (50)$$

$$\sum_{i \in V|(i,j)\in A} x_{ij} = 1,\ j \in E \quad (51)$$

$$\sum_{i \in V|(i,h)\in A} x_{ih} - \sum_{j \in V|(h,j)\in A} x_{hj} = 0,\ h \in V \backslash (S \cup E) \quad (52)$$

$$w_i + t_{ij} - B(1 - x_{ij}) \leq w_j,\ (i,j) \in A \quad (53)$$

$$q_i + d_i - B(1 - x_{ij}) \leq q_j,\ i \in C \cup S, \{j \in V|(i,j) \in A\} \quad (54)$$

$$q_{D_i^m} - d_i^m - B\left(1 - x_{D_i^m j}\right) \leq q_j,\ i \in C \cup E, m \in M, \{j \in V|(D_i^m, j) \in A\} \quad (55)$$

$$\sum_{i \in C \cup E} d_i^m \leq P^m,\ m \in M \quad (56)$$

$$\sum_{m \in M} \sum_{i \in C \cup E} d_i^m = \sum_{i \in C} d_i \quad (57)$$

$$\sum_{i \in S \cup E} q_i + \sum_{i \in S} w_i = 0 \quad (58)$$

$$0 \leq q_j \leq Q \sum_{i \in V|(i,j)\in A} x_{ij},\ j \in V \backslash (S \cup E) \quad (59)$$

$$0 \leq w_j \leq B \sum_{i \in V|(i,j)\in A} x_{ij},\ j \in V \backslash S \quad (60)$$

$$d_i^m \in \mathbb{N}_0,\ i \in C \cup E, m \in M \quad (61)$$

$$x_{ij} \in \{0,1\},\ (i,j) \in A \quad (62)$$

Ο περιορισμός (48) διασφαλίζει ότι ο χρόνος εκκένωσης είναι μεγαλύτερος του χρόνου της τελευταίας επίσκεψης οχήματος σε κάποιο καταφύγιο. Ο περιορισμός (49) διασφαλίζει ότι τα οχήματα θα εκκινήσουν από τους σταθμούς εκκίνησης και θα εξυπηρετήσουν όλα τα σημεία

παραλαβής. Ο περιορισμός (50) ορίζει ότι ο κάθε ψευδοκόμβος εξυπηρετείται μία ή καμία φορά από τα οχήματα. Ο περιορισμός (51) ορίζει ότι όλα τα οχήματα θα επισκεφτούν τον τερματικό σταθμό. Ο περιορισμός (52) διασφαλίζει ότι αν ένα όχημα φθάσει σε ένα κόμβο, εκτός των σταθμών εκκίνησης και τερματισμού, τότε θα αναχωρήσει και από αυτόν.

Ο περιορισμός (53) διατηρεί τη ροή του χρόνου όταν ένα όχημα ταξιδεύει από τον κόμβο i στον κόμβο j, θεωρώντας ότι $B \gg 1$. Ο περιορισμός (54) ορίζει τη μεταβολή του αριθμού επιβατών στο όχημα που επισκέπτεται τον κόμβο i, αμέσως μετά την αναχώρησή του από αυτόν. Ο περιορισμός (55) ορίζει τη μεταβολή του αριθμού επιβατών στο όχημα που επισκέπτεται ένα καταφύγιο, αμέσως μετά την αναχώρηση του από αυτό. Ο περιορισμός (56) διασφαλίζει ότι ο συνολικός αριθμός των πολιτών που θα αποβιβαστούν σε όλους τους ψευδοκόμβους που αντιστοιχούν στο ίδιο πραγματικό καταφύγιο δεν θα ξεπεράσει τη δυναμικότητα του. Σύμφωνα με τον περιορισμό (57) ο αριθμός των πολιτών που θα αποβιβαστούν σε όλα τα καταφύγια θα είναι ίσος με τον αριθμό των πολιτών που περιμένουν στα σημεία παραλαβής. Ο περιορισμός (58) ορίζει ότι όλα τα οχήματα εκκινούν από τους σταθμούς εκκίνησης χωρίς επιβάτες τη χρονική στιγμή 0 και τερματίζουν στους τερματικούς σταθμούς και πάλι χωρίς επιβάτες. Ο περιορισμός (59) διασφαλίζει ότι σε κάθε κόμβο, εκτός των σταθμών εκκίνησης και τερματισμού, ο αριθμός των πολιτών που επιβαίνουν σε κάθε όχημα είναι μεγαλύτερος ή ίσος του μηδενός και μικρότερος ή ίσος της χωρητικότητας του. Επιπλέον, αν ένας κόμβος δεν εξυπηρετείται από κάποιο όχημα τότε ο αριθμός αυτός ισούται με το μηδέν. Παρομοίως, ο περιορισμός (60) διασφαλίζει ότι ο χρόνος άφιξης στους κόμβους είναι μεγαλύτερος ή ίσος με το μηδέν, και συγκεκριμένα ίσος με μηδέν στην περίπτωση που ο κόμβος δεν εξυπηρετείται από κανένα όχημα. Ο περιορισμός (61) διασφαλίζει ότι ο αριθμός των πολιτών που αποβιβάζονται στα καταφύγια είναι μεγαλύτερος ή ίσος με μηδέν. Τέλος, ο περιορισμός (62) ορίζει τη δυαδική φύση των μεταβλητών που σχετίζονται με τις ακμές.

Διατύπωση μαθηματικού μοντέλου ΠΕΤ

Ο αντικειμενικός σκοπός του προβλήματος εκκένωσης τραυματιών είναι η ελαχιστοποίηση του αθροίσματος των χρόνων άφιξης των τραυματιών στις νοσοκομειακές μονάδες (ή ισοδύναμα η ελαχιστοποίηση του μέσου χρόνου άφιξης). Για το πρόβλημα αυτό ορίζονται επιπλέον τα εξής:

- R το σύνολο που περιέχει κάθε τραυματία h που περιμένει να μεταφερθεί

- $\tilde{d}_i^h, i \in V, h \in R$ λαμβάνει την τιμή 1 αν ο τραυματίας h περιμένει στην τοποθεσία i, αλλιώς λαμβάνει την τιμή 0
- $a_h, h \in R$ ο χρόνος άφιξης του τραυματία h σε νοσοκομειακή μονάδα
- $\tilde{d}_i^{mh}$, $i \in C \cup E, m \in M, h \in R$ λαμβάνει την τιμή 1 αν ο τραυματίας h μεταφέρεται στον ψευδοκόμβο της νοσοκομειακής μονάδας m ο οποίος αναφέρεται στον κόμβο i
- $\tilde{q}_i^h$ for $i \in V, h \in R$ λαμβάνει την τιμή 1 αν ο τραυματίας h επιβαίνει στο όχημα ακριβώς πριν αυτό φθάσει στον κόμβο 1, αλλιώς λαμβάνει την τιμή 0.

Η αντικειμενική συνάρτηση του ΠΕΤ ορίζεται ως:

$$TC = \min \sum_{h \in R} a_h \tag{63}$$

και υπόκειται στους περιορισμούς (49) – (53), καθώς και στους ακόλουθους:

$$\tilde{q}_i^h + \tilde{d}_i^h - B(1 - x_{ij}) \leq \tilde{q}_j^h, \ \ i \in C \cup S, \{j \in V | (i,j) \in A\}, h \in R \tag{64}$$

$$\tilde{q}_{D_i^m}^h - \tilde{d}_i^{mh} - B\left(1 - x_{D_i^m j}\right) \leq \tilde{q}_j^h, \ i \in C \cup E, m \in M, \{j \in V | (D_i^m, j) \in A\}, h \in R \tag{65}$$

$$w_{D_i^m} - B(1 - \tilde{d}_i^{mh}) \leq a_h, \ i \in C \cup E, m \in M, \ h \in R \tag{66}$$

$$\sum_{i \in C \cup E} \sum_{m \in M} \tilde{d}_i^{mh} = 1, \ h \in R \tag{67}$$

$$\sum_{h \in R} \sum_{i \in C \cup E} \tilde{d}_i^{mh} \leq P^m, \ m \in M \tag{68}$$

$$\sum_{i \in S} w_i + \sum_{i \in S \cup E} \sum_{h \in R} \tilde{q}_i^h = 0, \tag{69}$$

$$0 \leq \sum_{h \in R} \tilde{q}_i^h \leq Q \sum_{j \in V | (i,j) \in A} x_{ij}, \ i \in V \backslash (S \cup E) \tag{70}$$

$$\tilde{d}_i^{mh}, \tilde{q}_i^h, x_{ij} \in \{0,1\} \tag{71}$$

Ο περιορισμός (64) ορίζει τη μεταβολή του αριθμού επιβατών στο όχημα που επισκέπτεται τον κόμβο i, αμέσως μετά την αναχώρησή του από αυτόν. Ο περιορισμός (65) ορίζει τη μεταβολή του αριθμού επιβατών σε όχημα που επισκέπτεται νοσοκομειακή μονάδα. Ο περιορισμός (66) ορίζει τον χρόνο άφιξης του κάθε τραυματία σε νοσοκομειακή μονάδα. Ο περιορισμός (67) διασφαλίζει ότι όλοι οι τραυματίες θα μεταφερθούν σε νοσοκομειακές μονάδες. Ο περιορισμός (68) διασφαλίζει ότι μια νοσοκομειακή μονάδα δεν θα δεχθεί περισσότερους τραυματίες από τη δυναμικότητα της. Ο περιορισμός (69) διασφαλίζει ότι όλα τα οχήματα εκκινούν χωρίς επιβάτες τη χρονική στιγμή 0. Ο περιορισμός (70) διασφαλίζει

ότι ο αριθμός των τραυματιών που επιβαίνουν σε όχημα δεν υπερβαίνει τη χωρητικότητά του.

Για το ΤΕΠ εξετάζουμε και μια εναλλακτική αντικειμενική συνάρτηση παρόμοια με αυτή του ΠΠΕΠΧΛ, η οποία στοχεύει στην ελαχιστοποίηση του χρόνου εκκένωσης. Επιπλέον ανάμεσα στις λύσεις με τον ίδιο χρόνο εκκένωσης επιλέγεται αυτή με το μικρότερο άθροισμα των χρόνων άφιξης των τραυματιών στις νοσοκομειακές μονάδες.

$$TC = \min T_{evac} + \frac{1}{\tilde{L}} \sum_{h \in R} a_h \tag{72}$$

Σε αυτή την περίπτωση $\tilde{L} = |R| \cdot \sum_{(i,j) \in A} t_{ij}$.

Υβριδική προσέγγιση επίλυσης για το ΠΠΕΠΧΛ

Για την επίλυση του ΠΠΕΠΧΛ αναπτύξαμε ένα πλαίσιο επίλυσης, το οποίο συνδυάζει κατάλληλα γνωστές ευρετικές και ακριβείς μεθόδους για την παραγωγή αποτελεσματικών λύσεων με μικρή απόκλιση από τη βέλτιστη. Το προτεινόμενο πλαίσιο αυτό περιλαμβάνει απλό ευρετικό αλγόριθμο (NN-BEP) για την παραγωγή αρχικής εφικτής λύσης και δύο προηγμένες φάσεις βελτίωσης.

Κατά τη διάρκεια της πρώτης φάσης παράγεται μια λύση αποδεκτής ποιότητας χρησιμοποιώντας την αλληλουχία των εξής βημάτων:

- Μετευρετικό αλγόριθμο Large Neighborhood Search (LNS) μέσω του οποίου εφικτές λύσεις αποσυντίθενται και ανασυντίθενται για συγκεκριμένο αριθμό επαναλήψεων με σκοπό να βελτιωθεί σημαντικά η αρχική λύση. Επίσης, όλες οι μοναδικές διαδρομές οχημάτων που αναγνωρίζονται κατά τις διαδικασίες του αλγορίθμου εισάγονται στο σύνολο εναλλακτικών διαδρομών Φ', το οποίο χρησιμοποιείται στο επόμενο βήμα της πρώτης φάσης για την παραγωγή μια καλής ολοκληρωμένης λύσης
- Πρόγραμμα μεικτού γραμμικού ακέραιου προγραμματισμού (RMP) που αναπτύχθηκε για το ΠΠΕΠΧΛ, το οποίο συνθέτει καλή ολοκληρωμένη λύση συνδυάζοντας τις εναλλακτικές διαδρομές του συνόλου Φ'
- Ευρετικό αλγόριθμο που βασίζεται στο Πρόβλημα Περιοδεύοντος Πωλητή (Travelling Salesman Problem Improvement, TSPI), ο οποίος εφαρμόζεται σε κάθε διαδρομή της λύσης του RMP για την περαιτέρω βελτίωσή της.

Η πρώτη φάση του πλαισίου εφαρμόζεται έως ότου δεν είναι δυνατή η βελτίωση της καλύτερης λύσης. Αυτό επιδιώκεται από τη δεύτερη φάση του πλαισίου χρησιμοποιώντας

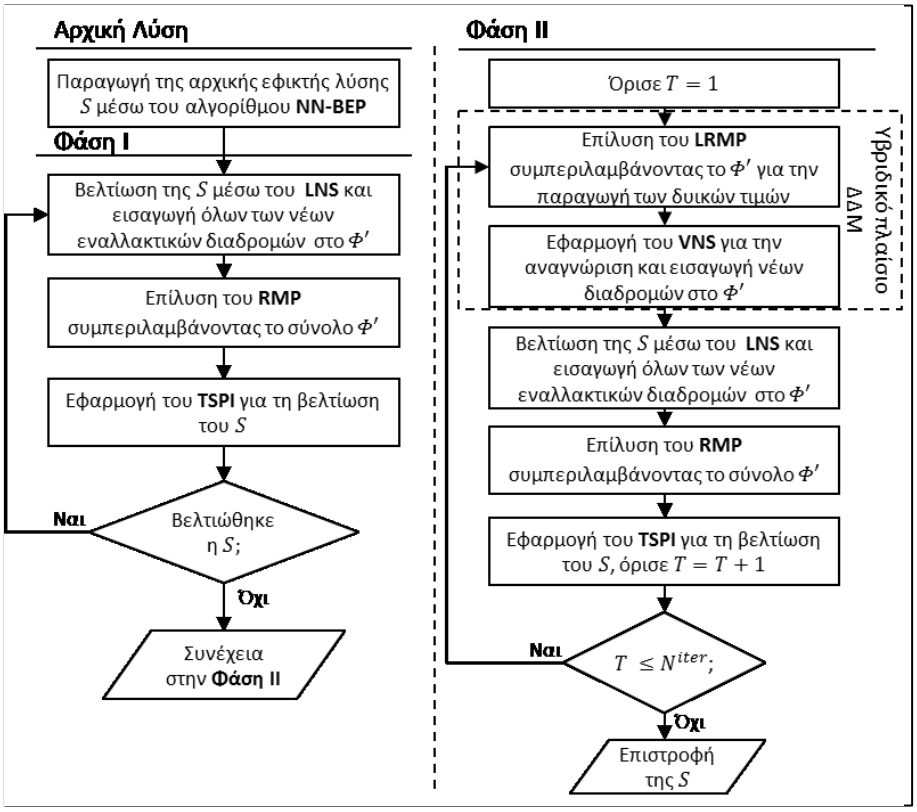
επιπλέον αλγόριθμο, ο οποίος προσομοιάζει τη δομή της Μεθόδου Δυναμικής Δημιουργίας Μεταβλητών (Column Generation). Ο αλγόριθμός αυτός περιλαμβάνει δύο βήματα:

(1) Επίλυση του χαλαρωμένου RMP (LRMP) με τη χρήση του συνόλου Φ' για την παραγωγή των δυικών τιμών της λύσης του προγράμματος

(2) Εφαρμογή του μετευρετικού αλγορίθμου Variable Neighborhood Search (VNS), ο οποίος εκμεταλλεύεται τις δυικές τιμές για να αναγνωρίσει και να προσθέσει στο σύνολο Φ' επιπλέον εναλλακτικές διαδρομές.

- Η αλληλουχία μεθόδων που ακολουθεί η δεύτερη φάση είναι:
  - Επίλυση του χαλαρωμένου RMP
  - Εφαρμογή του VNS
  - Εφαρμογή του LNS
  - Επίλυση του RMP
  - Εφαρμογή του TSPI

Η δεύτερη φάση εφαρμόζεται για συγκεκριμένο αριθμό επαναλήψεων (N^{iter}) κατά τις οποίες το εύρος αναζήτησης επεκτείνεται με βάση την παράμετρο T. Το Σχήμα 6 απεικονίζει το διάγραμμα ροής του πλαισίου επίλυσης που αναπτύχθηκε για το ΠΠΕΠΧΛ.



Σχήμα 6: Διάγραμμα ροής του υβριδικού πλαισίου επίλυσης

Ενδεικτικά αποτελέσματα και συμπεράσματα

Η πειραματική μελέτη σχετικά με τα προβλήματα ΠΠΕΠΧΛ και ΠΕΤ αφορούσε την αξιολόγηση του προτεινόμενου υβριδικού πλαισίου. Αρχικά, επικυρώθηκε η ποιότητα των λύσεων που παράγονται από το υβριδικό πλαίσιο συγκρίνοντάς τες με τις βέλτιστες λύσεις που παράγονται από την επίλυση του μαθηματικού μοντέλου για προβλήματα μικρής κλίμακας. Σύμφωνα με τα αποτελέσματα. το υβριδικό πλαίσιο παράγει λύσεις πλησίον των βέλτιστων σε πολύ ικανοποιητικούς υπολογιστικούς χρόνους.

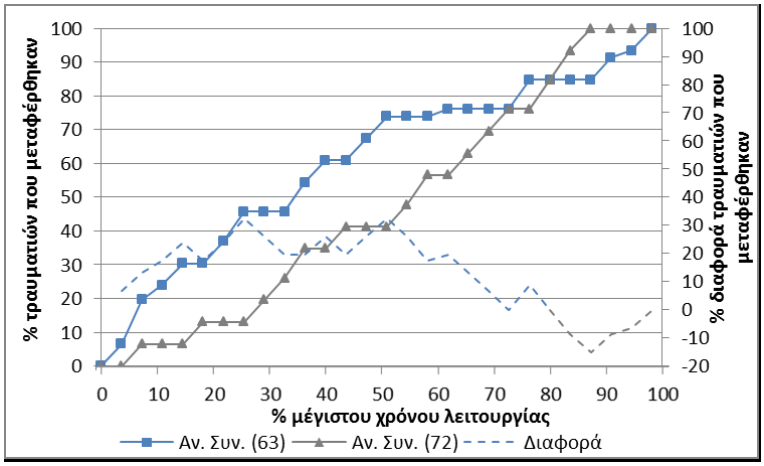
Το Σχήμα 7 απεικονίζει την εξέλιξη της απόκλισης της τρέχουσας λύσης από την καλύτερη λύση. Από το Σχήμα αυτό διαφαίνεται ότι η Φάση ΙΙ συνεισφέρει σημαντικά στη βελτίωση της λύσης. Συγκεκριμένα, η Φάση ΙΙ, κατά την οποία διευρύνεται το φάσμα αναζήτησης γειτονικών λύσεων, συνεισφέρει περίπου 5% στη βελτίωση της τελικής λύσης.

Σχήμα 7: Σύγκλιση στην τελική λύση του προτεινόμενου πλαισίου κατά τη διάρκεια των δύο φάσεων για το πρόβλημα "coord100-5-1"

Επίσης, μελετήθηκε η επίπτωση του επιπέδου ομαδοποίησης των πολιτών (ή τραυματιών) που αναμένουν να μεταφερθούν στα καταφύγια (ή στις νοσοκομειακές μονάδες) στην ποιότητα των λύσεων που παράγονται. Τα αποτελέσματα υποδεικνύουν ότι όταν οι πολίτες ομαδοποιούνται σε μικρότερες ομάδες, ο απαιτούμενος υπολογιστικός χρόνος αυξάνεται σημαντικά. Ωστόσο, η ποιότητα των λύσεων επηρεάζεται σημαντικά μόνο σε περιπτώσεις στις οποίες η χωρητικότητα των διαθέσιμων οχημάτων δεν αξιοποιείται ικανοποιητικά.

Για την περίπτωση του ΠΕΤ εξετάστηκε η επίδραση των δυο διαφορετικών αντικειμενικών συναρτήσεων στην αποτελεσματικότητα των λύσεων. Όπως φαίνεται και στο Σχήμα 8, ακολουθώντας την αντικειμενική συνάρτηση (63) περισσότεροι τραυματίες μεταφέρονται στις νοσοκομειακές μονάδες στα πρώτα στάδια της εκκένωσης, ενώ ακολουθώντας την αντικειμενική συνάρτηση (72) η συνολική διαδικασία ολοκληρώνεται σε μικρότερο χρονικό διάστημα. Αυτή η διαφορά είναι σημαντική και θα πρέπει να λαμβάνεται υπόψη ανάλογα με

τις συνθήκες που επικρατούν. Όταν υπάρχει οριστική και αυστηρή διορία (π.χ. λήξη ανακωχής ή παύση πυρός), τότε η αντικειμενική συνάρτηση (72) είναι η καταλληλότερη. Όταν δεν υφίσταται τέτοια διορία η καταλληλότερη αντικειμενική συνάρτηση είναι εκείνη της εξίσωσης (63), καθότι ελαττώνει τον μέσο χρόνο μεταφοράς των τραυματιών στις νοσοκομειακές μονάδες.



Σχήμα 8: Ποσοστό των τραυματιών που μεταφέρθηκαν στις νοσοκομειακές μονάδες κατά τη διάρκεια της διαδικασίας εκκένωσης

Τέλος, η προτεινόμενη μέθοδος εξετάστηκε σε προβλήματα πραγματικής κλίμακας με 4.000 πολίτες στην περίπτωση του ΠΠΕΠΧΛ και με 100 τραυματίες στην περίπτωση του ΤΕΠ. Τα πειραματικά αποτελέσματα υπέδειξαν ότι η υβριδική μέθοδος παράγει λύσεις σταθερής ποιότητας σε αποδεκτούς υπολογιστικούς χρόνους. Συνεπώς, είναι κατάλληλη για την αντιμετώπιση προβλημάτων εκκενώσεων περιοχών.

Αποφάσεις επιστροφής οχήματος για αναπλήρωση αποθέματος σε περίπτωση δύο σημείων αναπλήρωσης και προκαθορισμένης σειράς επίσκεψης πελατών

Εκμεταλλευόμενοι μερικά από τα σημαντικά χαρακτηριστικά των παραπάνω προβλημάτων, καθώς και των μεθόδων επίλυσης που αναπτύχθηκαν για αυτά, μελετήσαμε ειδικές περιπτώσεις στον τομέα των logistics. Συγκεκριμένα, τα εν λόγω χαρακτηριστικά είναι τα εξής: (α) η προκαθορισμένη και αυστηρή σειρά επίσκεψης πελατών, όπως στην περίπτωση των στάσεων των δρομολογίων των λεωφορείων, και (β) η περιορισμένη δυναμικότητα των σημείων αναπλήρωσης του αποθέματος, όπως και στην περίπτωση των καταφυγίων και των νοσοκομειακών μονάδων.

Οι ειδικές περιπτώσεις του τομέα των logistics που μελετήθηκαν αφορούν τη δρομολόγηση ενός οχήματος και δύο σημείων αναπλήρωσης του αποθέματος του οχήματος. Συγκεκριμένα,

μελετήθηκαν τρεις παραλλαγές ως προς τη χωρητικότητα των σημείων αναπλήρωσης. Στην πρώτη παραλλαγή (UU) και τα δύο σημεία αναπλήρωσης έχουν απεριόριστη χωρητικότητα, στη δεύτερη (UL) το ένα σημείο έχει απεριόριστη χωρητικότητα και το δεύτερο περιορισμένη, ενώ στην τρίτη παραλλαγή (LL) και τα δύο σημεία αναπλήρωσης έχουν περιορισμένη χωρητικότητα.

Σημειώνεται ότι στην παραλλαγή UU, η απόφαση που πρέπει να ληφθεί είναι εάν μετά από την εξυπηρέτηση ενός πελάτη θα πρέπει το όχημα να επιστρέψει ή όχι στο πλησιέστερο σημείο αναπλήρωσης αποθέματος για πλήρη αναπλήρωση. Στις άλλες δύο παραλλαγές UL και LL, εφόσον τουλάχιστον ένα από τα δύο σημεία αναπλήρωσης είναι περιορισμένης χωρητικότητας, οι αποφάσεις που πρέπει να ληφθούν μετά την εξυπηρέτηση ενός πελάτη είναι: (i) επιστροφή ή όχι σε σημείο αναπλήρωσης, (ii) σε ποιο σημείο αναπλήρωσης πρέπει να κατευθυνθεί το όχημα, και (iii) η ποσότητα προϊόντος που θα αναπληρωθεί κατά την επίσκεψη, σε περίπτωση που το επιλεγμένο σημείο αναπλήρωσης έχει περιορισμένη δυναμικότητα.

Διατύπωση μαθηματικού μοντέλου UU

Θεωρήστε ότι:

- Το σύνολο $C = \{1, \dots, n\}$ περιέχει τους κόμβους πελατών του δικτύου
- Ο κόμβος 0 είναι το σημείο εκκίνησης του οχήματος και ο κόμβος $n + 1$ είναι το σημείο τερματισμού του οχήματος
- Οι κόμβοι D^1 και D^2 είναι τα δυο διαθέσιμα σημεία ανεφοδιασμού.

Θεωρήστε τον κατευθυνόμενο γράφο $G(V, A)$, όπου $V = C \cup \{0, D^1, D^2, n + 1\}$ είναι το σύνολο των κόμβων και $A = \{(i, i + 1), (i, D^1), (i, D^2), (D^1, D^2), (D^2, D^1), (D^1, i + 1), (D^2, i + 1): i \in V \setminus \{n + 1\}\}$ είναι το σύνολο των ακμών. Το όχημα πρέπει να εξυπηρετήσει όλους τους πελάτες ακλουθώντας τη σειρά $1, \dots, n$. Ο κάθε πελάτης έχει γνωστή ζήτηση $r_i, \forall i \in C$ για κοινό προϊόν και $c_{ij} \geq 0$ είναι το κόστος (απόσταση) του οχήματος για να διασχίσει την ακμή (i, j). Το αρχικό απόθεμα των δύο σημείων αναπλήρωσης είναι Π^1 και Π^2, αντίστοιχα, ενώ η χωρητικότητα του οχήματος ισούται με Q μονάδες προϊόντος. Η ζήτηση του κόμβου 0 είναι $r_0 = 0$, και το φορτίο του οχήματος είναι γνωστό κατά την εκκίνησή του.

Αντικειμενικός σκοπός του προβλήματος είναι να ικανοποιηθεί η ζήτηση όλων των πελατών ακολουθώντας την προκαθορισμένη σειρά και διασχίζοντας την ελάχιστη απόσταση.

Όπως αναφέρθηκε και προηγουμένως, στην περίπτωση UU η χωρητικότητα των δύο σημείων αναπλήρωσης αποθέματος είναι απεριόριστη ($\Pi^1, \Pi^2 \to \infty$). Συνεπώς η διαδοχική επίσκεψη των δύο σημείων αναπλήρωσης δεν είναι απαραίτητη, διότι το όχημα μπορεί να αναπληρώσει όλο το απόθεμα του με επίσκεψη σε ένα μόνο σημείο αναπλήρωσης. Οπότε, το σύνολο τον απαραίτητων ακμών είναι το A', το οποίο είναι υποσύνολο του συνόλου A και ορίζεται ως εξής:

$$A' = \{(i, i+1), (i, D^1), (i, D^2), (D^1, i+1), (D^2, i+1)\} : i \in V \setminus \{n+1\}, A' \subseteq A$$

Οι μεταβλητές απόφασης του είναι οι εξής:

- $x_{i,i+1}, i = 0, \dots, n$, που σχετίζεται με την ακμή $(i, i+1)$ και αντιστοιχεί σε απευθείας μετάβαση από τον πελάτη i στον πελάτη $i+1$
- $x_i^D, i = 0, \dots, n$, που σχετίζεται με το σύνολο ακμών $\{(i, D^1), (D^1, i+1)\}$ στην περίπτωση που $c_{iD^1} + c_{D^1 i+1} \leq c_{iD^2} + c_{D^2 i+1}$. Αλλιώς η μεταβλητή σχετίζεται με το σύνολο ακμών $\{(i, D^2), (D^2, i+1)\}$.

Κάθε μία από τις παραπάνω μεταβλητές απόφασης λαμβάνει την τιμή 1 αν η αντίστοιχη ακμή διατρέχεται από το όχημα, αλλιώς λαμβάνει την τιμή 0. Το κόστος που σχετίζεται με τη μεταβλητή $x_{i,i+1}$ είναι $c_{i,i+1}$, ενώ το κόστος που σχετίζεται με τη μεταβλητή x_i^D είναι:

$$\tilde{c}_i^D = \begin{cases} c_{iD^1} + c_{D^1 i+1}, & c_{iD^1} + c_{D^1 i+1} \leq c_{iD^2} + c_{D^2 i+1} \\ c_{iD^2} + c_{D^2 i+1}, & \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

Οι μεταβλητές απόφασης L_i ορίζουν την ποσότητα προϊόντος που φορτώνεται στο όχημα κατά τη διάρκεια της επίσκεψης του σε σημείο αναπλήρωσης αποθέματος (αποθήκη). Το φορτίο του οχήματος αμέσως μετά την επίσκεψη στον κόμβο i ορίζεται ως q_i. Η μαθηματική διατύπωση του προβλήματος για την περίπτωση UU είναι η εξής:

$$minTC = \sum_{i=0}^{n-1} \left(c_{i,i+1} x_{i,i+1} + \tilde{c}_i^D x_i^D \right) + c_{n,n+1} \tag{73}$$

Δεδομένου ότι:

$$x_{i,i+1} + x_i^D = 1, \ i = 0, \dots, n \tag{74}$$

$$q_i = q_{i-1} - r_i + L_{i-1}, \ i = 1, \dots, n \tag{75}$$

$$q_i + r_i \leq Q, \ i = 0, \dots, n \tag{76}$$

$$L_0 \geq Q - M \times (1 - x_0^D) \tag{77}$$

$$L_i \geq Q - q_{i-1} - M \times (1 - x_i^D),\ \ i = 1, \dots, n \tag{78}$$

$$L_i \leq M \times x_i^D,\ \ i = 0, \dots, n \tag{79}$$

$$q_i, L_i \in \mathbb{N}_0,\ \ i = 0, \dots, n \tag{80}$$

$$x_{i,i+1}, x_i^D \in \{0,1\},\ \ i = 0, \dots, n \tag{81}$$

Ο περιορισμός (74) διασφαλίζει ότι το όχημα μετά την αναχώρηση του από πελάτη θα κατευθυνθεί είτε στον επόμενο πελάτη της προκαθορισμένης σειράς, είτε σε σημείο αναπλήρωσης αποθέματος. Ο περιορισμός (75) ορίζει τη μεταβολή του φορτίου του οχήματος κατά την επίσκεψή του στον πελάτη i. Ο περιορισμός (76) διασφαλίζει ότι η χωρητικότητα του οχήματος δεν θα παραβιαστεί σε κανένα σημείο του δρομολογίου. Οι περιορισμοί (75) με (78) διασφαλίζουν ότι κάθε φορά που το όχημα επισκέπτεται ένα από τα δύο σημεία αναπλήρωσης αποθέματος, θα πληρωθεί στο μέγιστο της χωρητικότητας του. Ο μεγάλος αριθμός $M \gg 1$ χρησιμοποιείται για λόγους γραμμικοποίησης. Ο περιορισμός (79) επίσης χρησιμοποιείται για λόγους γραμμικοποίησης, ώστε να αποφευχθούν οι πολλαπλασιασμοί μεταξύ των μεταβλητών που αντιστοιχούν στις ακμές και των μεταβλητών που αντιστοιχούν στην ποσότητα του προϊόντος που φορτώνεται στο όχημα. Ο περιορισμός (80) ορίζει ότι οι μεταβλητές που σχετίζονται με την ποσότητα του προϊόντος που φορτώνεται στο όχημα είναι ακέραιοι αριθμοί μεγαλύτεροι ή ίσοι με το μηδέν. Τέλος, ο περιορισμός (81) ορίζει τη δυαδική φύση των μεταβλητών που αντιστοιχούν στις ακμές.

Τα μαθηματικά μοντέλα των προβλημάτων UL και LL έχουν σημαντικές ομοιότητες με το παραπάνω μοντέλο, αλλά διαφέρουν ως προς το σύνολο των ακμών του γράφου και ως προς ορισμένους περιορισμούς χωρητικότητας.

Προτεινόμενες μέθοδοι επίλυσης

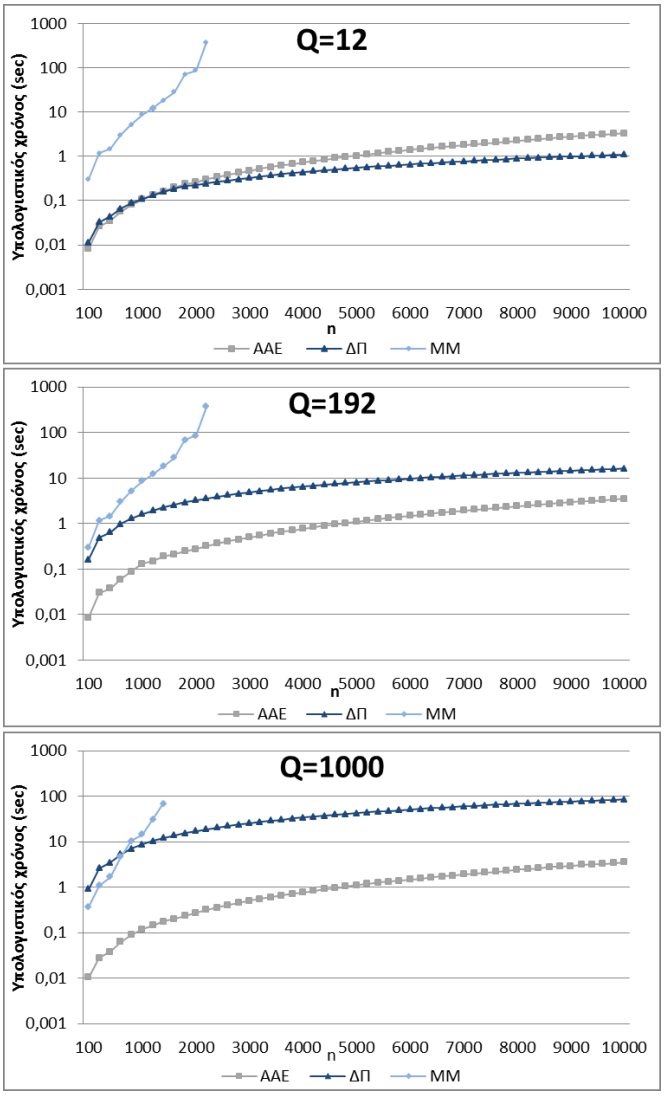
Για την επίλυση των προβλημάτων UU, UL και LL αναπτύχθηκαν εναλλακτικές μέθοδοι που περιλαμβάνουν τον καθιερωμένο Δυναμικό Προγραμματισμό (ΔΠ) για την πιο απλή παραλλαγή (UU), προσεγγίσεις Αλγορίθμου Ανάθεσης Ετικετών (ΑΑΕ, Labeling Algorithms) και νέο ευρετικό Αλγόριθμο Τμηματοποίησης του Δρομολογίου (ΑΤΔ, Partitioning Heuristic).

Οι προσεγγίσεις ΑΑΕ περιλαμβάνουν εξειδικευμένους κανόνες κυριαρχίας σε σχέση με το φορτίο του οχήματος και με το απόθεμα των σημείων αναπλήρωσης. Συγκεκριμένα, για τις παραλλαγές UU και UL οι κανόνες κυριαρχίας εγγυώνται τον εντοπισμό τις βέλτιστης λύσης, ενώ δεν εγγυώνται τον εντοπισμό της βέλτιστης λύσης για την παραλλαγή LL.

Ο νέος ευρετικός αλγόριθμος ΑΤΔ αποσυνθέτει το πρόβλημα ως εξής: (α) Αρχικά η ακολουθία των πελατών χωρίζεται σε N διαδοχικές ακολουθίες (σκέλη). (β) Για κάθε σκέλος, εφαρμόζεται ο ΑΑΕ για την παραγωγή όλων των εφικτών δρομολογίων. (γ) Ο καλύτερος συνδυασμός για τη δημιουργία ενός ολοκληρωμένου δρομολόγιου επιλέγεται από την επίλυση ειδικού προβλήματος δυαδικού προγραμματισμού.

Ενδεικτικά αποτελέσματα και συμπεράσματα

Η απόδοση όλων των αλγορίθμων αξιολογήθηκε χρησιμοποιώντας μεγάλο αριθμό προβλημάτων, συγκρίνοντας τις παραγόμενες λύσεις των προτεινόμενων μεθόδων με τις βέλτιστες λύσεις που παράγονται επιλύοντας τα αντίστοιχα μαθηματικά μοντέλα (ΜΜ).

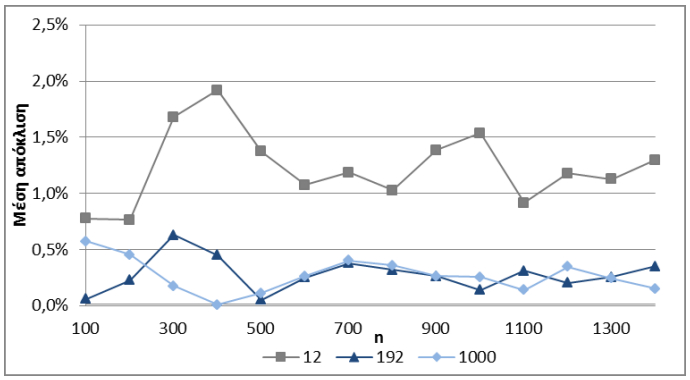


Σχήμα 9: Πρόβλημα UU – Υπολογιστικός χρόνος ως προς τον αριθμό των πελατών *n* για τρεις τιμές της χωρητικότητας του οχήματος *Q*

Οι ΑΑΕ παράγουν βέλτιστες λύσεις για τα προβλήματα UU και UL, ενώ παράγουν αποδοτικές λύσεις για το πρόβλημα LL. Όσον αφορά τους απαιτούμενους υπολογιστικούς

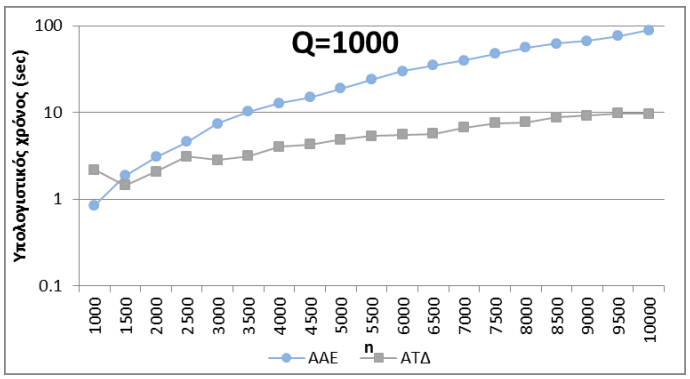
χρόνους, οι ΑΑΕ απαιτούν σημαντικά μικρότερο υπολογιστικό χρόνο σε σχέση με τον ΔΠ. Ενδεικτικά, σύμφωνα με το Σχήμα 9 για το πρόβλημα UU, η απόδοση του ΔΠ εξαρτάται από τη χωρητικότητα του οχήματος, ενώ ο αντίστοιχος ΑΑΕ (UU) δεν εξαρτάται από την παράμετρο αυτή.

Για το πρόβλημα UL τροποποιήσαμε τον αντίστοιχο ΑΑΕ, ο οποίος παράγει αποδοτικές (όχι βέλτιστες) λύσεις σε χρόνο μικρότερο έως και τρεις τάξεις μεγέθους από τον ΑΑΕ που παράγει βέλτιστες λύσεις, καθώς και από το αντίστοιχο ΜΜ. Στο σχήμα 10 απεικονίζεται η μέση απόκλιση των λύσεων του τροποποιημένου ΑΑΕ ως προς τη βέλτιστη λύση σε σχέση με τον αριθμό των πελατών n. Όπως φαίνεται ο τροποποιημένος ΑΑΕ παράγει ποιοτικές λύσεις.



Σχήμα 10: Πρόβλημα UL - Μέση απόκλιση των λύσεων του τροποποιημένου ΑΑΕ από τη βέλτιστη λύση σε σχέση με τον αριθμό των πελατών n για τρεις τιμές της χωρητικότητας του οχήματος

Για το πρόβλημα LL, ο ευρετικός ΑΑΕ και ο ΑΤΔ απαιτούν σημαντικά λιγότερο χρόνο από ότι η επίλυση του ΜΜ (1 με 3 τάξεις μεγέθους λιγότερο), ενώ συγκλίνουν σε αποδοτικές λύσεις (απόκλιση κάτω του 1% από τη βέλτιστη λύση). Τέλος, σύμφωνα με το Σχήμα 11, ο ΑΤΔ παρουσιάζει την καλύτερη απόδοση για προβλήματα LL μεγάλης κλίμακας, συγκλίνοντας σε ποιοτικές λύσεις εντός υπολογιστικού χρόνου έως και μίας τάξης μεγέθους μικρότερο σε σχέση με τον ΑΑΕ.



Σχήμα 11: Πρόβλημα LL – Υπολογιστικός χρόνος ως προς τον αριθμό των πελατών n για τιμή χωρητικότητας οχήματος $Q = 100$

Abstract

In this dissertation we investigate paratransit services, which support mobility for people with a disability, under normal and emergency conditions. The demand for such services is expected to increase significantly, as life expectancy and the proportion of senior citizens in the general population are growing globally.

For paratransit services under normal conditions, we introduce and study Scheduled Paratransit Transport Systems (SPTS), in which a public transport bus is allowed to diverge from its nominal path to pick-up paratransit passengers and drop them off at their destination. To study the design aspects of the SPTS system we defined and modeled a new routing problem, the Demand Responsive Bus Routing Problem (DRBRP). The latter has been solved by an exact Branch and Price framework, which includes a labeling algorithm with appropriate dominance rules in order to reach the optimal solution within reasonable computational times. Using this solution framework, we have studied SPTS with respect to its significant design parameters, and proposed guidelines for the efficient development of such systems in various transport environments. Subsequently, we investigate SPTS-T, an interesting enhancement of SPTS, which aims to satisfy the entire paratransit demand by combining SPTS with accessible taxi (or private hire) services. The problem objective is to minimize the cost of the combined transportation system, while serving all paratransit requests and offering sufficient level of service. We have modeled SPTS-T by combining attributes of the DRBRP and Dial-a-Ride Problem (DARP), and solved it by an exact Branch and Price framework. We have performed an extensive experimental study that has resulted in specific guidelines to system designers.

In the area of paratransit services under emergency conditions we investigated the problem of evacuating the residents of a region in case of an emergency - the bus evacuation problem. We have developed a hybrid algorithm (HCG) that incorporates the column generation framework, with the well-known meta-heuristic methods LNS and VNS. We also have studied an interesting variation of this problem, which deals with the evacuation of casualties from several triage points. This case may be encountered in a battlefield or after disasters (natural catastrophes, terrorist attacks, etc.), in which a significant number of casualties should be transported to available medical facilities. Extensive experimental results have indicated that for both cases the proposed framework provides near optimal solutions for

small scale instances, and efficient solutions in reasonable computational times for problems of practical scale.

Finally, we study the implications of the above research in logistics cases, in which conditions similar to those of the aforementioned transportation problems hold, including routes with strict sequences, or limitations for the capacities of collection/gathering centers. Specifically, we have examined three interesting cases of the single vehicle routing problem with a predefined client sequence and two load replenishment warehouses. For each case we have developed appropriate solution algorithms to address their increasing complexity. The efficiency of these algorithms has been studied by solving a wide range of problem instances, and by comparing the results with those obtained from a state-of-the-art MILP solver.

Contents

Chapter 1. Introduction....1

Chapter 2. Background and dissertation objectives....6

2.1 Dissertation motivation and research focus areas....7

2.2 Flexible demand responsive transportation systems....13

2.3 Evacuation planning for transit depended populations....14

2.4 Routing problems related to paratransit....16

2.5 Branch and price method....24

2.6 Research contribution of the dissertation....28

Chapter 3. Scheduled paratransit transport systems (SPTS)....31

3.1 Description of SPTS....32

3.2 Mathematical formulation of DRBRP....34

3.3 A Branch and price approach for the DRBRP problem....39

3.4 Tests and results....45

Chapter 4. Scheduled paratransit transport enhanced by accessible taxis (SPTS-T)....55

4.1 Mathematical formulation of DRBRP-T....58

4.2 A branch and price approach for the DRBRP-T problem....62

4.3 Tests and results....65

Chapter 5. Solving the bus evacuation problem and its extensions....76

5.1 Mathematical formulations for BEP and CEP....78

5.2 A hybrid solution framework for BEP....86

5.3 Tests and results....97

Chapter 6. Application of paratransit transport aspects to logistics....109

6.1 Description of the problem variants under consideration....110

6.2 Modeling the three problem cases....112

6.3 Solution methods....117

6.4 Experimental analysis 125

Chapter 7. Conclusions and recommendations for future research 131

7.1 Conclusions 131

7.2 Recommendations for future research 134

References 137

Appendices 157

Appendix A: Generation of test data for scheduled paratransit transport systems 157

Appendix B: Tuning the probabilistic return to shelters (PRS) heuristic 162

List of Figures

Figure 2.1. Estimated prevalence of moderate and severe disability by region and age (WHO, 2011) 8

Figure 3.1: Why paratransit ride time is important: 1^+ request pick-up, 1^- request drop-off 33

Figure 3.2: Illustration of the graph of a simple example 35

Figure 3.3: Service performance of SPTS for three different cases of d and assuming no other constraint 48

Figure 3.4: Service performance of SPTS for three different cases of Q, assuming that $d = 1$ and no other constraint 49

Figure 3.5: Service performance of SPTS, for three different cases of time window width, assuming that $d = 1$ and $Q = 2$ 49

Figure 3.6: Service performance of SPTS as *d* increases, for three different cases of Q, assuming that $n/k = 2$ and $\xi = \infty$ 50

Figure 3.7: SPTS performance as *d* increases, for three different cases of ξ, assuming that $n/k = 2$ and $Q = 2$ 50

Figure 3.8: SPTS performance as *d* increases, for combinations of ξ and Q, assuming that $n/k = 2$ 50

Figure 3.9: Service level for $n = 30$ and $Q = 1$ for the three different environments 53

Figure 3.10: Service level for $n = 30$ and for different cases of the rural environment 53

Figure 3.11: Ratio of served requests provided by SPTS for different parameters combinations assuming that $n/k = 2$ 54

Figure 4.1a: Modal split of paratransit clients for increasing values of α and for ξ=high, d=high and $Q = 1$ 70

Figure 4.1b: Cost per mode for increasing values of a and for ξ=high, d=high and $Q = 1$ 70

Figure 4.2a: Bus modal split for the three cases studied, $Q = 1$, and for increasing values of a 70

Figure 4.2b: Cost of the SPTS-T for and the three cases studied, $Q = 1$, and for increasing values of a 70

Figure 4.3a: Bus modal split for the three cases studied, Q=2 and for increasing values of a 71

Figure 4.3b: Cost of SPTS-T for the three cases studied, $Q = 2$, and for increasing values of a 71

Figure 4.4a: Bus modal split for different cases of ξ and Q for increasing values of d and $a = 10$ and $n/(k \cdot Q) = 2.1$ 72

Figure 4.4b: Bus modal split for different cases of ξ and Q for increasing values of d and $a = 10$ and $n/(k \cdot Q) = 0.75$ 72

Figure 4.5a: Cost of SPTS-T for different cases of ξ and Q for increasing values of d, $a = 10$ and $n/(k \cdot Q) = 2.1$ 73

Figure 4.5b: Cost of SPTS-T for different cases of ξ and Q for increasing values of d, $a = 10$ and $n/(k \cdot Q) = 0.75$ 73

Figure 5.1: Illustration of the problem nodes in a simple example 80

Figure 5.2: Illustration of the graph of the simple example of Fig. 5.1 81

Figure 5.3: Outline of the hybrid framework 87

Figure 5.4: Effect of U on computational time: Ratio $\tau(U)$ of time required when $U = 1$ with time required when $U \in \{2,4,6,8,10\}$ 101

Figure 5.5: HCG convergence during the two framework phases for problem “coord100-10-2” 103

Figure 5.6: HCG convergence during the two framework phases for problem “coord100-5-1” 104

Figure 5.7: Percentage of casualties transported to the medical facilities during the transport operation 108

Figure 6.1: The network of the problem 113

Figure 6.2: Problem UU - Computational time as a function of the number of clients n for three values of the capacity Q 126

Figure 6.3: Problem UL - Computational time as a function of the number of clients n for three values of the capacity Q 127

Figure 6.4: Average deviation of *la(UL-non-opt)* solutions from the optimal w.r.t. n for $Q = \{12,192,1000\}$ 128

Figure 6.5: Problem LL - Computational time as a function of n for three values of the capacity Q 129

Figure 6.6: Problem LL - Average deviation of the solution from the optimal as a function of n for three values of the capacity Q 129

Figure 6.7: Problem LL - Computational time as a function of the number of clients n for three values of the capacity Q 130

Figure A.1: Area for the locating bus stop $i + 1$ w.r.t. to bus stop i 159

Figure A.2: Creation of a pick-up and a drop-off location in urban environment 159

Figure A.3: Creation of a pick-up and a drop-off location for a suburban environment 160

Figure A.4: Creation of a pick-up and a drop-off location in rural environment 161

Figure B.1: Return probability P_Q as a function of evacuees on board for different values of the temperature T 163

Figure B.2: Return probability P_D as a function of Δt over the average travelling time for different values of temperature T 163

Figure B.3: Convergence of the solution with respect to temperature change 164

Figure B.4: Average deviation from the best solution for the two probability functions 164

Figure B.5: Impact of the frequency of feasibility check to the efficiency and performance of the total framework 165

List of Tables

Table 2.1: Relationship of routing problems with the research areas of the dissertasion 16

Table 3.1: How DRBRP treats service levels 34

Table 3.2: Comparison of BSS and SS performance 47

Table 3.3: Parameters settings per mobility environment 51

Table 3.4: Key SPTS characteristics per application environment 52

Table 4.1: Time window settings per transportation mode and per trip type for outbound (inbound) trips 57

Table 4.2: Quality of service per client type of SPTS-T (for the time window widths, please refer to Table 4.1) 57

Table 4.3: Parameters settings used for problem generation (urban environment) 67

Table 4.4: Problem cases with progressively increasing service levels 69

Table 5.1: Inputs and outputs of the constituent elements of the hybrid framework 97

Table 5.2: Comparison of HCG performance with MILP for BEP 99

Table 5.3: Effect of U on the quality of the solution; i.e. the deviation $\delta(U)$ for $U \in \{2,4,6,8,10\}$ from the solution with $U = 1$ 100

Table 5.4: Theoretical empty capacity of the vehicles if the demand d is broken into groups of U 102

Table 5.5: Parameters of the example problems and performance of HCG for the pre-disaster evacuation case 103

Table 5.6: Comparison of HCG and MILP performance for CEP 105

Table 5.7: Comparison of HCG and MILP performance for CEP-MO 105

Table 5.8: Parameters of casualties evacuation instances 106

Table 5.9: HCG performance on casualty evacuation problems 107

Table B.1: Parameters of the test problems 162

List of Abbreviations

Abbreviation	Description
AARP	American Association of Retired Persons
AGV	Automated Guided Vehicle
BB	Branch and Bound
BEP	Bus Evacuation Problem
BP	Branch and Price
BSS	Bus Stop Structure
CEP	Casualty Evacuation Problem
CEP-MO	Multi Objective – Casualty Evacuation Problem
CG	Column Generation
CVRP	Capacitated Vehicle Routing Problem
DARP	Dial-a-Ride Problem
DAS	Demand Adaptive System
DHS	Department of Homeland Security
DRBRP	Demand Responsive Bus Routing Problem
DRBRP-T	Demand Responsive Bus Routing Problem combined with the Dial-A-Ride problem
DRT	Demand Responsive Transit
ESPPRC	Elementary Shortest Path Problem with Resource Constraints
ESPPTWCPD	Elementary Shortest Path Problem with Time Windows, Capacity, and Pickup and Delivery
FTA	Federal Transit Administration
FV	First Version
GAO	Government Accountability Office (US)
HCG	Hybrid Column Generation
IBEP	Integrated Bus Evacuation Problem
LA/la	Labeling algorithm
LB	Lower Bound
LL	Limited – Limited

Abbreviation	Description
LMP	Linear Master Problem
LNS	Large Neighborhood Search
LRP	Location Routing Problem
m-	Multiple bus routes
MAST	Mobility Allowance Shuttle Transit
MILP	Mixed Integer Linear Programming
NN-BEP	Nearest Neighbor like heuristic (for BEP)
OP	Orienteering Problem
OP-PDP	Orienteering Problem with Pick-up and Deliveries
PA/pa	Partition Algorithm
PDP	Pickup and Delivery Problem
PDPTW	Pickup and Delivery Problem with Time Windows
PRS	Probabilistic Return to Shelters heuristic
RBEP	Robust Bus Evacuation Problem
RLMP	Restricted Linear Master Problem
RMP	Restricted Master Problem
SDRBRP	Elementary Shortest Path Problem with Time Windows, Capacity, Pick-up and Delivery as well as Sequence Restrictions
SP	Sub-Problem
SPTS	Scheduled Paratransit Transport Systems
SPTS-T	Scheduled paratransit transport enhanced by accessible taxis
SS	This latter algorithm (hereafter called
SV	Second Version
TOP	Team Orienteering Problem
TRB	Transportation Research Board
TSP	Traveling Salesman Problem
TSPI	Travelling Salesman Problem Improvement heuristic

Abbreviation	Description
UB	Upper Bound
UBEP	Un-capacitated Bus Evacuation Problem
UL	Unlimited – Limited
US	United States
UU	Unlimited – Unlimited
VNS	Variable Neighborhood Search
VRPSF	Vehicle Routing Problem with Satellite Facilities
VRPTW	Vehicle Routing Problem with Time Windows
WHO	World Health Organization

Chapter 1. Introduction

In 2013 the share of people over 60 in the global population reached 11.7% (United Nations, 2013), while in Europe similar statistics indicate that the proportion of senior citizens is even higher reaching 20% (World Health Organization, 2014). By 2050, the global proportion of senior citizens is estimated to grow significantly reaching 21.1% (United Nations, 2013). However, according to the same reference, as the world population ages, the relative occurrence of some form of disability also increases. According to the World Health Organization (2011) and Avendano *et al.* (2009) in the US, and for the year 2004, mobility limitations were reported in 59% of men and women in the age group of 50 to 74 years; the corresponding percentage in Europe is 43%. Thus, in addition to the requirement of special mobility services for people with disabilities, the demand for such services is expected to increase dramatically as life expectancy increases in the near future. This becomes a

significant social priority, since mobility is directly related to quality of life, affecting almost all the activities of daily living.

The need for special transportation becomes more acute during emergency situations, such as natural or manmade disasters demanding emergency evacuation. Hence, it is important to include special initiatives for the transit-depended population into emergency preparedness plans. This need was tragically displayed in New Orleans during Hurricane Katrina in 2005. New Orleans was estimated to be the residence of 200,000 transit-depended people (Wolshon, 2002). Approximately 971 deaths were attributed to Hurricane Katrina, with 47% of the victims in the age group of 75 years old and older. Moreover, less than 10% of the victims were younger than 45 years old, while the mean age of the victims was 69 years (Brunkard *et al.*, 2008).

Significant research has focused on improving the transportation conditions of transit-depended populations based on advances in various fields, including Operational Research, Information Technology, Statistics, Meteorology, etc. Following the results of research initiatives, government and the private sector have developed and introduced novel individualized transportation services for these populations. According to the National Transit Database (2013), "(the number of) *Demand response* (transport) *agencies* (was) *increased by 32.2% over the ten year period (2003 to 2012), reflecting the need to continue providing special transit service for elderly individuals and individuals with disabilities*". Demand response transport agencies include transit systems that are scheduled in response to calls from passengers, and typically include minibuses or taxis.

This dissertation focuses on the design of such systems. We first consider the case of providing individualized services using existing modes of public transport. We also explore how services of such demand responsive systems may be enhanced by using chartered special vehicles (taxis or private hire vehicles). In both cases, the goal is to provide high quality door-to-door transportation for clients in need of special transport, while mitigating the related costs.

We also consider the case of transportation under emergency situations, which call for population evacuation. Such evacuations are commonly pre-planned and mostly car-based. This practice, however, does not address the needs of transit-dependent individuals (e.g. the elderly, people with medical or mobility issues etc.). Furthermore, car-based evacuations may also result in high traffic congestion levels due to low vehicle occupancy. To deal with such

issues, we investigate bus-based evacuation and the related transportation problems that consider significant constraints raised by the capacity of shelters. We extend this work to an interesting transportation problem that involves the transportation of casualties after an emergency or terrorist incident, or from a battlefield, to temporary or permanent medical facilities. This latter problem shares similar characteristics and restrictions with bus-based evacuation, but it differs in terms of objectives and the scale of demand.

Finally, we investigate applications of the above four paratransit problems to logistics, and specifically to the case of a single distribution vehicle that uses two warehouses (or depots) for load replenishment. Three cases with respect to the inventory level of the two warehouses have been investigated.

In addressing the above problems, this dissertation has made the following research contributions: (a) Definition of novel transportation problems, for which appropriate mathematical formulations have been introduced. (b) Development of exact Branch and Price and hybrid meta-heuristic methods to solve these problems in an efficient manner, taking into full consideration their characteristics. (c) Provision of design guidelines for the development of special paratransit transportation systems. These guidelines have been based on the outcomes of extensive experimental investigations that studied the tuning of key system parameters in various application environments.

The structure of the dissertation is summarized as follows:

Chapter 2 first discusses the motivation and the main focus areas of the dissertation, and reviews important research in these areas. Subsequently, the literature review focuses on the fundamentals of related routing problems, as well as on the core operations research methods leveraged to develop the solution methods proposed by the dissertation. Finally, Chapter 2 identifies the research gaps and outlines the contributions of this work.

Chapter 3 focuses on ways to provide individualized services to people with mobility challenges using existing modes of public transport. This chapter studies the design of an interesting system, in which buses operating in a public transport route may diverge from their nominal path to pick-up passengers with limited mobility and drop them off at their destination. The design problem has been modeled by a mixed integer-linear program, and an exact Branch and Price approach has been developed to solve it to optimality. The proposed approach includes a labeling algorithm in which we introduced appropriate dominance rules, which do not compromise optimality. The efficiency of our approach has been compared with

that of related algorithms from the recent literature. Furthermore, the proposed approach has been used to study key aspects of the system design problem, such as the effect of various constraints on the service level, and the tuning of the system's parameters to address different transport environments.

Chapter 4 studies an enhancement of the above system, which aims to satisfy the entire door-to-door paratransit demand. To do so, the system of Chapter 3 is supplemented by accessible taxi services. The scope of the enhanced system is to minimize overall transport cost under service quality imperatives. The proposed system and its aspects have been studied by modeling the related transportation problem as a mixed integer-linear program, which is solved to optimality by a new exact Branch and Price approach. An extensive experimental study has been performed to provide design guidelines for achieving an appropriate balance between the costs of the public bus and taxi system, and the various service quality levels.

Chapter 5 addresses the problem of transporting paratransit passengers from their origins to multiple gathering centers. Two interesting and related cases of this transportation problem have been considered: (a) the case of evacuation in anticipation of major natural threats, and (b) the case of transporting casualties after an emergency or terrorist incident, or from a battlefield. Both problem cases share common characteristics and restrictions, i.e. limited available vehicles and limited capacity of the collection/gathering centers, but differ in terms of objectives and the scale of demand. To address the evacuation problem and its extensions, two-index, mixed integer, linear formulations have been proposed, and a Hybrid Column Generation solution an framework has been developed. Extensive experimental results indicate that for both cases the proposed framework provides near optimal solutions for small scale instances, and efficient solutions in reasonable computational times for problems of practical scale.

Chapter 6 examines interesting applications in the area of logistics that share common features with the above problems; specifically, (a) predefined and strict route sequences similar to those of public bus services, and (b) limitations in the capacity of warehouses, or inventory replenishment facilities, similar to the limitations faced by the shelters during evacuations. Specifically, we investigate three cases of the single vehicle routing problem with a predefined client sequence and two load replenishment warehouses. Given the location and demand of the clients, the goal is to identify the minimal cost route, which includes optimal load replenishment visits to the warehouses in order to fully satisfy the client demand. The cases studied vary with respect to inventory availability at each warehouse and

are of increasing complexity. Solution algorithms have been developed to address this complexity, ranging from a standard dynamic programming algorithm for the simplest case, to labeling algorithms and a new partitioning heuristic. The efficiency of these algorithms has been studied by solving a wide range of problem instances, and by comparing the results with those obtained from a state-of-the-art MILP solver.

Finally, Chapter 7 presents the conclusions of this dissertation along with directions for further research.

Chapter 2. Background and dissertation objectives

In this dissertation we investigate the following: (a) Design of flexible demand responsive services using existing modes of public transport to provide quality door-to-door transport for paratransit clients. (b) Planning bus-based evacuation for transit-dependent people in anticipation of major natural threats. In this area, we also investigate the similar problem of transporting casualties after an emergency or terrorist incident, or from a battlefield to temporary or permanent medical facilities.

Prior to focusing on these research problems, Chapter 2 reviews the background and previous research in this area, and sets the research goals. Specifically, Section 2.1 discusses the characteristics and the necessity for paratransit transport systems and the potential improvements in the quality of life of the targeted groups. The following two Sections (2.2 and 2.3) present the literature on the focus areas of the dissertation. Subsequently, Section 2.4

presents the respective literature for vehicle routing problems that are closely related to the systems under consideration. Firstly fundamental routing problems are reviewed, and, subsequently, variations of these problems related to the characteristics of paratransit transport are discussed. Section 2.5 presents a targeted discussion on the core operations research method that forms the foundation for the solution approaches applied in the current dissertation. Finally, Section 2.6 highlights research gaps, and the contributions of this dissertation.

2.1 Dissertation motivation and research focus areas

In this Section we discuss the motivation for the research conducted in this dissertation. We first consider the general context, that is, the targeted population groups and the importance of the ability to move. Subsequently, we discuss major issues concerning people in need of special transportation conditions under both normal and emergency conditions.

2.1.1 Mobility, Disabilities and ageing population

Mobility is a form of capital, the availability of and access to which depends on age, gender, class, and ethnicity (Tesfahuney, 1998; Stjernborg *et al*., 2014). It enhances social networks and social interaction, creates a sense of control and independence, denotes status and role, and helps people interact with nature and exercise cognitive skills (Musselwhite and Haddad, 2010). Thus, it is a key element in the quality of life for all (Spinney *et al.*, 2009).

Mobility is more challenging for people with a disability, defined as a physical or mental condition that restricts the basic physical, spiritual and mental functions (Todorov and Kirchner, 2000, Mollaoğlu *et al*., 2010). Figure 2.1 captures results of the World Health Survey conducted by World Health Organization (2011) across 59 countries related to rates of disability in the general population. It is clear that disability increases significantly with age, averaging 18.2% among all age groups.

Especially for senior citizens, researchers have concluded that: (a) mobility, (b) the ability to leave home, and (c) the ability to maintain social relations are key to their wellbeing (Irwin, 1969; Farquhar, 1995; Marottoli *et al*., 1997; Metz, 2000; Fonda *et al*., 2001; Metz, 2003; Gabriel and Bowling, 2004; Hardy *et al*., 2005; Mollaoğlu *et al*., 2010; Schwanen and Ziegler, 2011; Ziegler and Schwanen, 2011; Ravulaparthy *et al*., 2013; Hjorthol, 2013; Stjernborg *et al*., 2014). Kim and Ulfarsson (2004) state that for the older population the need

to travel of may be even greater than that of the younger population, since senior citizens must consume more time for non-home activities and are in need for more social services and health care. Moreover, according to Musselwhite and Haddad (2010), the older population travels to fulfill the desire for independence, control, maintaining status, inclusion and travel for its own sake.

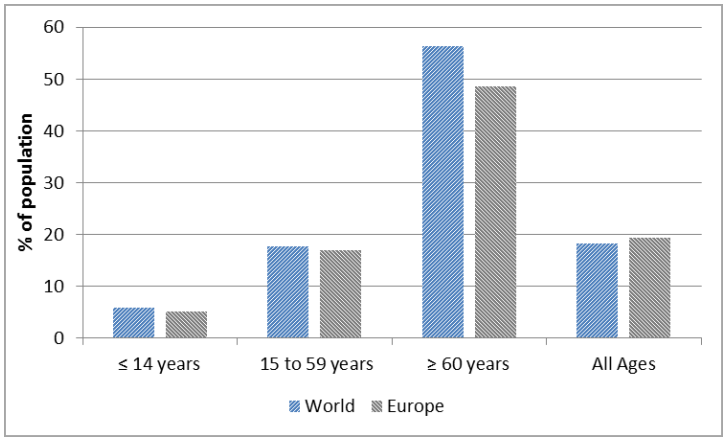


Figure 2.1. Estimated prevalence of moderate and severe disability by region and age (WHO, 2011)

The absence of mobility in older age promotes feelings of social deprivation and exclusion (Church *et al*., 2000; Schonfelder and Axhausen, 2003; Urry, 2007; Ravulaparthy *et al*., 2013; Stjernborg *et al*., 2014), but also contributes to depression (Fonda *et al*., 2001) and isolation (Johnson, 2002). In the study of Mollaoğlu *et al*. (2010) a significant correlation between mobility and life satisfaction has been identified. Furthermore, the older population living in rural areas faces greater difficulties than those living in cities, since the elderly in these areas are more isolated and usually live at greater distances from medical and other services (Glasgow and Blakely, 2000; Schwarzlose *et al*., 2014).

Ageing has been also linked to a decrease in driving, an important skill that promotes mobility (Mercado and Páez, 2009; Newbold *et al*., 2005; Raitanen *et al*., 2003). Marottoli *et al*. (1997) identified that the decision to stop driving is a strong predictor for symptoms of depression in older age. Although, the need for self-dependency and the desire to travel motivates the elderly to continue driving (Burns, 1999; Schwarzlose *et al*., 2014), statistics indicate that older drivers are more likely to be involved in accidents (Schwanen and Páez, 2010); in those accidents the elderly are more likely to be at fault (Dellinger *et al*., 2002; Schwarzlose *et al*., 2014).

Apart from the driving skills and difficulties, Kim (2011) indicated that a significant proportion of the older population in the US limits their travel due to the deficiency of the transportation system. It is noted that the ability to use the transportation system without

assistance has been defined as one of the seven most important areas among the Instrumental Activities of Daily Living of older people's welfare and wellbeing (Fillenbaum, 1985; Hjorthol, 2013).

Thus, it is vital to provide transportation alternatives for the vulnerable population of people with disabilities, including the elderly. More so since, as already mentioned, the global share of people older than 60 is expected to rise significantly, and especially in developed countries as a result of increased longevity and declining fertility (Alsnih and Hensher, 2003; Spinney *et al*., 2009; Ravulaparthy *et al*., 2013; World Health Organization, 2011; United Nations, 2013).

2.1.2 Paratransit services under normal conditions

Novel transportation systems are necessary to allow people with disability, including the elderly, to maintain quality of life and inclusion in society (Su and Bell, 2009; Li *et al*., 2012). In addition, society benefits by transportation systems that will allow older people to live independently in their own homes for longer, since the cost of long-term care is reduced. Su and Bell (2009) state that: "*in many cases it is more economic and effective to provide special transport services than to improve the accessibility of traditional public transport to older people*". The latter is in line with the study of Stjernborg *et al*. (2014) that indicates the need of society to invest in special transport services to keep the most vulnerable group of people mobile when they are not able to use other modes of transport.

Offering special transport services is not a new approach. For instance, flexible route bus transport in terms of the "telebus" received positive feedback by people using a wheel chair in Berlin (Pajonk, 1982; Broome *et al*., 2012). Alsnih and Hensher (2003) pointed out that the market for flexible public transport services for the elderly is large enough to become a core business component for mainstream public bus operators.

Evidence suggests that older travellers will use public transport where it is accessible and is provided at a reasonable service level (Currie and Delbosc, 2010). Giuliano *et al*. (2003) point out that the development of paratransit options may be an effective strategy to address the mobility of older people. Su and Bell (2009) mention that modal choice is limited for people with mobility impairments, and state that this choice could be more flexible by improving traditional public transport services and by introducing innovative special services.

These types of services are the focus of the first part of the current dissertation. Specifically, we focus on flexible public transport services to address the mobility needs of the elderly, as well as those of people with a disability. Of interest are ways to provide individualized bus services to these populations using existing modes of public transport. Such services promote equality between client groups, while, at the same time, offer high quality transport to people with a disability, and limit public transport expenditures.

The characteristics of such flexible transportation services have been studied, discussed and proposed in the literature over the recent years. Specifically, such transport systems should ensure:

Proximity/accessibility: Public transportation becomes impractical if an individual needs transportation to a location that is at a distance from a fixed route (Hjorthol, 2013; Schwarzlose *et al*., 2014). Kim and Ulfarsson (2004) found that older people are more likely to use public transport if they live within five blocks of bus stops. Proximity has been found to be a strong indicator of public transport use in several other studies (Su *et al*., 2008; Hess, 2009; Currie and Delbosc, 2010).

Flexibility: Rosenbloom (2004) observed that responsive transit options might effectively substitute for car trips for the elderly and the disabled. Broome *et al*. (2012) evaluated the efficiency of providing flexible bus transport services to the elderly. They concluded that demand may be doubled, while overall satisfaction will increase significantly. Moreover, Rosenbloom (2001) suggests that specialized transport services could offer a good alternative to traditional public transport services.

Cost vs. time: Su *et al*. (2008) studied the effects of travel time and travel cost on the modal choice of the elderly during shopping trips. The study concluded that for the elderly travel cost is more significant than travel time. This is in line with the findings of Su and Bell (2009), who suggested that the elderly have more free time to use public transport. According to Su (2007) this has important implications for transport policy, which should provide specialized transport services considering the travel cost rather than travel time.

Cost of transportation services: Although efforts have been made to improve rural public transportation, rising costs and limited funding continue to hinder the growth of related programs (Schwarzlose *et al*., 2014). Thus, specialized transport services should not further burden the budget of public transportation. According to Nguyen-Hoang and Yeung (2010)

the benefits of a flexible demand-responsive form of public transportation exceed the related costs.

2.1.3 Paratransit services under emergency conditions

As stated in Murray-Tuite and Wolshon (2013), evacuations of populations in anticipation of natural threats are frequent. According to the study of the Sandia National Laboratory for the US Nuclear Regulatory Commission (Sandia National Laboratories, 2004), in the US an evacuation of 1,000 or more persons is required every two to three weeks.

A significant number of reports and research papers have recognized and documented the risks that populations without access to a car are facing during an evacuation (Fischetti, 2001). For example, the evacuation of New Orleans, Louisiana in 2007 is considered as one of the most successful cases of evacuation by car in U.S. history with more than 1.2 million people evacuating the region in a 48hr period (Renne *et al*., 2009b). Unfortunately, it was also one of the most unsuccessful evacuations for those who were not able to evacuate by their own means (Renne *et al*., 2009b). As a result, 71% of those who lost their lives were over the age of 60, and 47% over the age of 75 (AARP, 2006). In particular, the 2006 Nationwide Plan Review of the U.S. Department of Homeland Security (DHS) concluded that very few US states and large urban areas have prepared sufficient plans for evacuating populations in need of assisted transportation, and highlighted that most evacuation plans consider personal vehicles, without considering the role of public transportation systems (DHS, 2006; GAO, 2007).

The understanding of the aforementioned facts promoted regulatory changes in emergency response research and practice (Wolshon *et al*., 2005a, 2005b; DHS, 2006; Litman, 2006; Hess and Gotham, 2007; Renne *et al*., 2008, 2009a, 2009b; An *et al*., 2013). In particular, significant changes in evacuation planning concern those residents who do not have access to a car or hesitate to use it during an evacuation. This population group is referred as the "carless" and mainly includes those who have mobility or health problems (Hess and Gotham, 2007; Renne *et al*., 2008; TRB, 2008; An *et al*., 2013).

Evacuating the carless population in a large scale urban area is a significant logistical challenge (Hess *et al*., 2013). The focus groups of Renne *et al*. (2009b) identified the variety of needs and the diversity of the carless population, which includes seniors, disabled adults, the homeless, the socially isolated, those that do not speak the local language (English in this case), those in postoperative recovery, the illiterate, and the tourists. According to TRB

(2008), these groups are spatially dispersed and difficult to reach efficiently in an emergency evacuation. However, Hess and Gotham (2007) suggested that the carless tend to be neighbors to people that are also carless. Despite their location, these populations rely on government assistance to safely evacuate (Hess *et al*., 2013), highlighting the critical role of public transportation in evacuations (Hess and Gotham, 2007).

According to TRB (2008) this role has only recently been recognized. This is also highlighted by Hess *et al*. (2013), who suggest that alternatives to private car evacuations are feasible if adequately planned and coordinated, since over 90% of carless households in the 100 largest metropolitan areas in the US have access to public transit service.

Apart from the need to evacuate the carless, there are certain risks on relying solely on private car. For instance, as stated in Gerber *et al*. (2010) when the private car is the primary mode used during an evacuation, then traffic congestion is likely to limit evacuation capability.

Wolshon *et al. (*2005a) indicates that evacuation strategies, such as the coordination of traffic control systems on arterial routes and the use of public transit and other vehicles, should be explored. Several other studies also note the potentially significant role of transit in emergency management planning (FTA, 2007; Schwartz and Litman 2008; Renne *et al*., 2008; Clarke and Habib, 2010).

Specifically, Higgins *et al*. (1999) identified the following key activities that may be performed by transit agencies during emergencies:

- Transportation: Evacuate people and particularly carless populations, transport emergency personnel (workers and volunteers) to/from emergency locations, transport passengers and supplies during the recovery from the disaster
- Sheltering: The air-conditioned/heated vehicles (buses) may be utilized as respite facilities for emergency personnel or as shelter for evacuees
- Communications and monitoring: The radio-equipped vehicles may assist to communication and to monitoring roads and weather conditions.

Similarly, the National Consortium on the Coordination of Human Services Transportation (2006) summarized in a strategy paper measures to serve populations with special needs in an emergency evacuation: identification and location of these populations, advance planning, communication, transportation assistance.

In the shame fashion, TRB (2008) stated that transit agencies can be important contributors in all four activities of emergency response, i.e. mitigation, preparedness, response, and recovery. TRB (2008) also identified the need for research in evacuation modeling and specifically with respect to the use of buses during evacuations.

Bus-based evacuation planning is the focus of this dissertation regarding paratransit transport under emergency situations. Our work focuses on ways to efficiently provide transportation and shelter for transit depended populations, such as populations:

- that are in institutions including hospitals, nursing homes, and prisons (Hess *et al*., 2013)
- that need medical attention but not hospitalization (TRB, 2008); transport to shelters with appropriate medical resources should be provided to them
- without access to private cars that need to be transported to shelters planned for the general population.

2.2 Flexible demand responsive transportation systems

Our work in this area is motivated from the need to serve the first and last leg of any public transport trip (i.e. the so-called *last mile*). This includes the legs between the origin (e.g. home) and the departure bus stop, as well as the arrival bus stop and the destination (e.g. place of work, hospital). Often such distances are considerable, especially in suburban or rural areas. In order to eliminate these trip legs, some municipalities have introduced semi-flexible bus routes (Flipper Project, 2011). The latter serve a reference (nominal) route that includes fixed regular bus stops; however, when there are requests from clients with limited mobility, one or more of the vehicles serving the route may diverge from this nominal path, pick-up the clients from their origin and eventually drop them to a regular bus stop or to a special destination. Obviously such diversions are possible within a certain distance from the nominal path.

The concept of Demand Responsive Transit (DRT) services has attracted attention over the last decade. Mageean and Nelson (2003) evaluated the application of DRT using telematics in five cities across Europe; their results indicated that DRT services may provide efficient public transportation, as well as an excellent level of service. Similar studies have been carried out by Brake *et al*. (2007) and Nelson *et al.* (2010). They concluded that the application of DRT services involve changes in business models, citizen behavior, and require technology support. Recently, Nourbakhsh and Ouyang (2012) proposed a structured

flexible route transit system, which eliminates the walking time to and from bus stops. This system allows buses to pick-up or drop-off passengers within predefined service areas, called bus tubes. The authors' quantitative analysis indicated that system operational costs are lower than those of an equivalent system that combines fixed-transit routes and taxi services, the latter to serve the first and last trip legs.

The Demand Adaptive System (DAS) of Daganzo (1984), Crainic *et al.* (2000, 2005), and Malucelli *et al.* (2001) is a special case of DRT and is related to the problems being addressed in this dissertation. In their case, optional bus stops are defined and are activated in order to provide improved service. Another case closely related to this dissertation is the Mobility Allowance Shuttle Transit (MAST) services of Quadrifoglio *et al.* (2006, 2007, 2008) and Quadrifoglio and Dessouky (2008). In MAST, a vehicle moves repeatedly from an origin point to a destination point and serves additional intermediate bus stops (checkpoints), while it is allowed to deviate within a certain range of the main route to serve additional requests. A latest departure time is defined for each checkpoint and cannot be violated. Chandra and Quadrifoglio (2013) also discuss the importance of first and last mile connectivity and propose an analytic queuing model for estimating a near optimal terminal-to-terminal cycle length for a demand responsive feeder service to maximize quality of passenger service.

2.3 Evacuation planning for transit depended populations

Over the last decade, significant research has also been dedicated to evacuation planning; interested readers may refer to Hamacher and Tjandra (2001), Altay and Green (2006), Bretschneider (2013), and Murray-Tuite and Wolshon (2013) for research advances in this area. Related aspects involve demand estimation, departure scheduling, route choice, contra-flow operations and relief operations.

However, studies on planning transit-based evacuation for the carless population are limited (Zhang and Chang, 2014). This activity commonly involves a fleet of buses, including public transport buses, school buses etc., and a number of available shelter locations. Using public transport in emergency evacuation operations has been considered in Sayyady and Eksioglu (2010). Their work aims to maximize the number of evacuees, assuming that not all evacuees may be transported, and vehicles are not allowed to perform multiple routes. A similar case has been presented by Abdelgawad *et al*. (2010) for no-notice evacuation in congested urban

areas. The authors applied heuristic rules in a constraint programming approach based on the ILOG - dispatcher. Kaufmann (2014) addressed the special case of the uncapacitated bus evacuation problem (UBEP), which involves the scheduling of a fleet of buses to evacuate people from a fixed set of gathering points to a fixed set of shelters. In their study the capacity of the shelters is assumed to be unlimited.

In Zheng's (2014) emergency evacuation strategy, the buses serve a set of pickup requests and delivery points using a certain routing strategy that aims to minimize the exposed casualty time. The delivery nodes of this case are of limited capacity and include both train stations and shelters. Recently, Zhang and Chang (2014) developed a model for an interesting evacuation planning case in response of no-noticed threads considering capacitated shelters. Their model consists of two phases: The first phase selects pick-up locations from candidate ones based on the spatial distribution of the evacuees. The second phase allocates available buses to transport the evacuees from the selected pick-up locations to the available shelter locations. Zhang and Chang (2014) also considered a fixed route policy, in which the buses start from the bus depot to visit at most two pick-up locations of evacuees and drop them off at the safety area. He *et al*. (2009) developed an optimization model to generate evacuation plans for transit-dependent residents considering uncertain demand. Their model was solved by a genetic algorithm in a random search manner.

Song *et al*. (2009) explored the Location Routing Problem (LRP) with uncertain demand, in order to provide a decision tool for evacuation planning. In this case, transit buses are used and the objective is to minimize total travelling time, while ensuring a maximum evacuation time is not exceeded. The case studied by Song *et al*. (2009) considers uncapacitated shelters, and that each vehicle is dedicated to one shelter.

Bish (2011) presented, discussed and studied the Bus Evacuation Problem (BEP). In BEP, a number of homogenous buses, which are located initially at one or more origins (depots), must pick-up known numbers of evacuees from known pick-up locations and transport them to available shelters. Each of the latter may accept up to a certain number of evacuees. The objective is to complete the evacuation in the minimum possible time span, which begins when the first bus departs from its depot and ends when the last evacuee arrives to a shelter. In this case the routes of the buses are flexible, i.e. they may visit multiple pick-up locations, as well as any available shelter. The authors provided two heuristics to address this problem: The first produces a feasible solution for the problem and, subsequently, improves it by route swapping and route reassignment. The second heuristic decomposes the solutions of the first

one in route segments, which are then reassigned to complete routes through a mathematical formulation that is solved within a certain time limit. The authors note that further research in improving the solution methods is warranted.

A simplified version of BEP has been discussed by Goerigk *et al*. (2013), who investigated approaches to identify lower and upper bounds within a Branch and Bound framework. In this simplified version, the fixed routing policy was also applied considering that the number of evacuees waiting at the pick-up locations are integer multiples of bus loads. Recently, Goerigk *et al*. (2014) introduced the Integrated Bus Evacuation Problem (IBEP) that extends the latter simplified model by determining both the pick-up points and the shelter points for evacuating a region with the help of buses. To address this problem, they developed a branch-and-price strategy and compared its efficiency using a commercial IP solver. Goerigk and Grün (2014) addressed the robust bus evacuation problem (RBEP), in which the exact numbers of evacuees are not known in advance. The decision is whether to dispatch a bus right away or to wait until more information becomes known.

2.4 Routing problems related to paratransit

Prior to discussing our contributions to paratransit transport under normal and emergency conditions, we present the fundamentals of routing problems, with particular emphasis on those problems closely related to the research in this dissertation. Table 2.1 relates the problems reviewed below to our research in demand responsive systems and bus-based evacuation planning.

Table 2.1: Relationship of routing problems with the research areas of the dissertasion

	Research Focus Area	
Routing problem	**Demand Responsive Systems**	**Bus-Based Evacuation**
Traveling salesman	-	Used as sub-problem in solution approach
Capacitated vehicle routing	Similarity in some model constraints	Similarity in some model constraints
Team orienteering	Some similarity in model objectives	-
Pickup and delivery	Parallels in methods used in solution approach	-
Dial-a-ride	Similarity in some model constraints, parallels in methods used in solution approach	Parallels in solution approach

Routing problem	Research Focus Area	
	Demand Responsive Systems	Bus-Based Evacuation
Location-routing	-	Similarities in model constraints and entities

2.4.1 Traveling salesman problem

The Traveling Salesman Problem (TSP; Dantzig *et al.,* 1954) captures the basic aspects of a routing problem, i.e. the generation of a route that serves certain objectives. Even if the TSP is probably the simplest routing problem, it is still an NP-complete problem (Papadimitriou, 1977). Specifically, in TSP the salesman starts from an origin city, visits exactly once all cities of a certain set, and returns to the origin city again. The objective is to identify the route with the shortest distance.

A way to formally describe the problem is the formulation of Miller *et al.* (1960). Consider a graph $G(V, A)$, where $V = \{0,1,2, \dots, v\}$ is the set of nodes with 0 been the selected origin city and v be a duplicate of 0 used as the destination node. Let A be the set of arcs that connect node 0 with all nodes except node v, and all nodes $V\backslash\{0, v\}$ between them and with node v. Let x_{ij} receive the value 1 if arc $(i, j) \in A$ participates in the solution, otherwise let it receive the value 0. Let $n_i, i \in V\backslash\{0, v\}$ be the position of node i in the path of the optimal solution, considering that $n_0 = 0$ and $n_v = v$. Also, let d_{ij} be the distance to travel from node i to node j. Then, the problem is to identify the arcs that minimize the total distance:

$$min \sum_{(i,j)\in A} d_{ij} x_{ij} \tag{2.1}$$

subject to:

$$\sum_{i|(i,j)\in A} x_{ij} = 1, \ j \in V\backslash\{0\} \tag{2.2}$$

$$\sum_{j|(i,j)\in A} x_{ij} = 1, \ i \in V\backslash\{v\} \tag{2.3}$$

$$n_i + 1 - M(1 - x_{ij}) \leq n_j, \ (i,j) \in A \tag{2.4}$$

$$x_{ij} \in \{0,1\}, \ (i,j) \in A \tag{2.5}$$

$$n_i \in \mathbb{N}_0, \ i \in V \tag{2.6}$$

Constraint (2.2) ensures that the vehicle (salesman) will arrive once to all nodes, excluding the starting one. Similarly, Constraint (2.3) ensures that the vehicle will also depart once from all nodes except the last one. Constraint (2.4) ensures that if the vehicle travels from node i to

node j, then the position of j in the sequence will be at least greater by 1 than the position of node i. The latter ensures that no subtours will be present at the final solution. Furthermore, in many variations of TSP, n_i may be replaced by other variables to account for the consumption of resources that increase along with the progress of the route, e.g. travel time.

TSP has been widely studied, and a vast number of methods have been developed to provide efficient solutions. These include: heuristic methods, such as the Nearest Neighbor algorithm (Hurkens and Woeginger, 2004), Clarke and Wright Savings (Clarke and Wright, 1964), Christofides (1979); metaheuristic methods, such as Genetic Algorithms (Holland, 1975), Simulated Annealing (Kirkpatrick *et al*., 1983), Tabu Search (Glover, 1989, 1990, 1996); and exact methods, such as Branch and Cut (Fischetti *et al.,* 2003; Applegate *et al.*, 2007).

2.4.2 Capacitated vehicle routing problem

The Capacitated Vehicle Routing Problem (CVRP) is a generalization of the TSP, although it uses different entity descriptions. It was firstly introduced by Dantzig and Ramser (1959), who proposed a mathematical programming formulation and an algorithmic approach to solve a real-life problem for the delivery of gasoline to service stations. The definition of CVRP and its variants, as well as an extensive analysis of solution methods, have been presented by Toth and Vigo (2002). CVRP is considered as the basis for a large class of routing problems.

The model presented below is a two-index vehicle flow formulation, as presented in Toth and Vigo (2002), assuming a complete graph $G(V, A)$. Let x_{ij} take the value 1 if arc $(i, j) \in A$ participates to the optimal solution, or take the value 0 otherwise. Also, let c_{ij} be the cost to traverse arc (i, j).

$$min \sum_{(i,j)\in A} c_{ij} x_{ij} \tag{2.7}$$

subject to

$$\sum_{i\in V} x_{ij} = 1, \ j \in V\backslash\{0\} \tag{2.8}$$

$$\sum_{j\in V} x_{ij} = 1, \ i \in V\backslash\{0\} \tag{2.9}$$

$$\sum_{i\in V} x_{i0} = K \tag{2.10}$$

$$\sum_{j \in V} x_{0j} = K \tag{2.11}$$

$$\sum_{i \notin S} \sum_{j \in S} x_{ij} \geq r(S), \;\; S \subseteq V \backslash \{0\}, S \neq \emptyset \tag{2.12}$$

$$x_{ij} \in \{0,1\}, \;\; i,j \in V \tag{2.13}$$

Constraints (2.8) and (2.9) ensure that exactly one arc arrives to and departs from each client node, respectively. Constraints (2.10) and (2.11) ensure that the number of arcs arrive and depart the depot node is equal to the number of available vehicles. Constraints (2.12) are the so-called capacity cut constraints and impose connectivity of the solution. In these constraints, a node set S is crossed by a number of arcs greater than or equal to $r(S)$, that is the minimum number of vehicles required to serve the nodes of set S.

CVRP is an NP-hard problem since it is a generalization of the TSP (Garey and Johnson, 1979). The latter becomes more obvious if one considers that a VRP problem with only one vehicle and capacity sufficient to satisfy all clients is actually a TSP.

There is extensive literature regarding methods for solving the VRP. Interested readers may refer to the work of Christofides *et al. (*1981), Desrochers *et al*. (1992), Laporte (1992), Laporte and Osman (1995), Toth and Vigo (2002), and Laporte (2009).

VRP is the seed for a vast number of problem variations that concern the literature of Operations Research. These variations correspond to different features with respect to:

- The clients, i.e.: their location, the nature of their demand, time windows for their service etc.
- The vehicles, i.e.: their capacity, the respective costs, the option to perform multiple trips etc.
- The objective function, i.e.: minimize the number of vehicles, minimize penalties related to the quality of service, maximize profit, maximize quality of service etc.

Well known variations of the VRP are the following:

- VRP with Time Windows (Cordeau *et al.,* 2002)
- VRP with Pickup and Delivery (Toth and Vigo, 2002; Daganzo and Hall, 1993)
- Distance Constrained VRP (Toth and Vigo, 2002)
- Multi-Depot VRP (Bianco *et al*., 1994)
- Heterogeneous Capacitated VRP (Taillard, 1999)
- VRP with Backhauls (Toth and Vigo, 2002)

- Periodic VRP (Christofides and Beasley, 1984)
- Open VRP with time windows (Repoussis *et al.*, 2007).

2.4.3 Team orienteering problem

In the orienteering routing problems, a vehicle starts from an origin and attempts to visit as many nodes as possible, striving to maximize the score collected from these visits, while covering the related distance or time within a predefined limit. Orienteering routing problems could be considered as a combination of the Traveling Salesman Problem (TSP) and the Knapsack Problem (Vansteenwegen *et al*., 2011). Early applications of the Orienteering Problem (OP) were presented by Tsiligirides (1984), in which a travelling salesman within limited time serves those clients who maximize total sales. Golden *et al*. (1987) addressed a related inventory routing problem regarding delivery of heating fuel on a daily basis in order to maximize profit. They demonstrated that OP is also an NP-hard problem.

The Team Orienteering Problem (TOP) is an extension of the OP, in which a team of players attempt to maximize the total points collected by the team within the time or distance limit. TOP was first defined by Butt and Cavalier (1994) by the term Multiple Tour Maximum Collection Problem. The term Team Orienteering Problem proposed by Chao *et al*. (1996b) is the prevalent one.

In order to describe the TOP formally, consider a directed graph $G = (N, A)$ where $N = \{0,1,2, \dots, n+1\}$ is the node set. Additionally, let n be the number of clients, and 0 and $n+1$ be the origin and destination depot(s), respectively. Assume arc set A that contains all arcs (i, j) connecting the nodes of set N. Furthermore, let $p_i, i \in N$ be the profit collected after serving node i, t_{ij} be the time required to traverse arc (i, j) assuming that the triangular inequality is satisfied, $[a_i, b_i]$ be the time interval within which each node must be served and $d_i, i \in N$ is the demand of each node. Finally, let V be the number of available vehicles, each of them of capacity Q units.

Then, the TOP with capacity and time window constraints consists of finding the set of vehicles routes M for which the total profit collected is maximized. The latter should be achieved considering that each vehicle route must start from node 0 and end up to $n+1$, and that each node i is served at most once, with service starting within time interval $[a_i, b_i]$. Furthermore, the load on the vehicle should not exceed Q units at any time during its route, and the number of vehicles to be used is $|M| \leq V$ with V a predefined number.

To formulate TOP with capacity and time window constraints as a (route based) binary problem, let Ω be the set of all routes that meet the above considerations. In this fashion, let $\hat{p}_r = \sum_{i \in V} a_{ir} p_i$ be the profit collected from route r, in which a_{ir} indicates the number of times node i is served in route r. Let the binary decision variable x_r be equal to 1 if route r participates in the solution, otherwise be equal to 0. The objective function aims to maximize the total profit collected:

$$TP = max \sum_{r \in \Omega} \hat{p}_r \, x_r \tag{2.14}$$

Subject to:

$$\sum_{r \in \Omega} a_{ir} x_r \leq 1, \;\; i \in N \backslash \{0, n+1\} \tag{2.15}$$

$$\sum_{r \in \Omega} x_r \leq V \tag{2.16}$$

$$x_r \in \{0,1\}, \;\; r \in \Omega \tag{2.17}$$

Inequality (2.15) limits each node to be served by up to one route, and Inequality (2.16) ensures that the number of vehicles do no exceed the predefined number.

Interested readers may refer to Chao *et al*. (1996a), Tang and Miller-Hooks (2005) for heuristic solution methods, and to Boussier *et al*. (2007) and Archetti *et al*. (2009) for exact methods. Vansteenwegen *et al*. (2011) also provide an interesting review concerning orienteering problems.

2.4.4 Pickup and delivery problems

In pickup and delivery problems (PDPs), goods or passengers are transported between an origin and a destination. The latter are involved in two types of special constraints: (a) pairing constraints that refer to the relationship between the origin and the destination locations, i.e. both locations should be served by the same route, and (b) precedence constraints to ensure that an origin location should be visited prior to the corresponding destination location. Different settings of these latter constraints generate a wide range of interesting practical problems, which present challenges regarding both the related formulation and appropriate solution methods.

Berbeglia *et al*. (2007, 2010), introduced a classification scheme for PDPs based on the number of origins and destinations involved. According to this scheme, PDPs may be classified into three (3) different categories:

- Many-to-Many (M-M) problems, in which any vertex may serve as a source or as a destination
- One-to-Many-to-One (1-M-1) problems, in which goods/ passengers initially available at a depot are to be transported to multiple sites, while other goods/ passengers available at these sites need to be transported back to the depot
- One-to-One (1-1) problems, in which each item/ passenger is associated with a certain origin and a certain destination.

Variations of the PDP involve the following problems:

- Swapping Problem (Anily and Hassin, 1992)
- Pickup and Delivery Traveling Salesman Problem (Hernandez-Perez and Salazar-Gonzalez, 2007)
- Multi-Vehicle PDP with Single Demands and Backhauls (Deif and Bodin, 1984)
- Multi-Vehicle PDP with Single Demands and Mixed Solutions (Golden *et al*., 1988)
- Single Vehicle PDP with Single Demands and Mixed Solutions (Anily and Mosheiov, 1994)
- PDP with Combined Demand (Gribkovskaia *et al*., 2007)
- Multi-Vehicle Routing Problem with Pick-ups and Deliveries (Dumas *et al*., 1991)
- Vehicle Routing Problem with Pick-ups, Deliveries and Transshipments (Mitrovic-Minic and Laporte, 2006)
- Dial a Ride Problem (Cordeau and Laporte, 2003, 2007).

Several surveys of the PDPTW literature have been presented in the last decades (see Savelsbergh and Sol, 1995; Mitrovic-Minic, 1998; Desaulniers *et al*., 2002; Ropke *et al*., 2007; Parragh *et al*., 2008a).

2.4.5 Dial-a-ride problems

The Dial-a-Ride Problem (DARP) is a variant of the PDP that concerns the transportation of passengers between an origin and a destination. In a typical DARP, a fleet of identical vehicles initially located at the same depot serve the client needs. The objective of the problem is to designate a set of vehicle routes that minimize the total operation cost, but also satisfy a set of constrains that safeguard the quality of service provided to the passengers. Many cases of DARP concern door-to-door transportation services for elderly or disabled people (Madsen *et al*., 1995; Toth and Vigo, 1997; Borndörfer *et al*., 1999; Diana and Dessouky, 2004; Rekiek *et al*., 2006; Melachrinoudis *et al*., 2007).

The most interesting aspect of DARP is the quality of service and has attracted the interest of the literature in the field of Operation Research for more than four decades, as stated in Paquette *et al.* (2009). The different quality attributes used in the literature involve (Paquette *et al.*, 2009):

- Difference between actual and desired delivery time
- Time windows
- The waiting time during the ride, i.e. the time on board during which the vehicle does not move
- Waiting time before pickup
- The total waiting and ride time
- The maximum ride time
- The average ride time
- The ratio of actual ride time vs. direct ride time
- The excess of ride time over direct time
- The total ride time
- The total cycle time (between the call and the delivery time)
- The maximum number of stops while a user is on board.

Interested readers may refer to Cordeau (2006) for exact formulations, and to Cordeau and Laporte (2003, 2007), and to Parragh *et al*. (2008b) for extensive surveys and reviews.

2.4.6 Location-routing problems

Another class of problems related to this dissertation includes the Location-Routing Problems (LRP), in which the appropriate number and location of distribution centers are determined simultaneously while optimizing the routing costs to serve a set of clients. The most common problem of this class is the Capacitated LRP (CLRP), which consists of: (a) determining which depots must be established (opened), (b) assigning the clients to open depots, and (c) develop vehicle routes for each depot and its clients. In most cases the latter are constructed by minimizing alternative cost functionals, such as the cost to establish a depot, the fixed costs of the vehicles, the cost of the routes etc., and are subject to constraints regarding the clients demand and the capacities of the vehicles and the depots, i.e.:

- the total demand of clients assigned to one depot must not exceed its capacity
- each route begins and ends at the same depot
- each vehicle performs at most one trip

- each client is served by one single vehicle (no split delivery)
- the total demand of clients visited by one vehicle fits the vehicle capacity.

An extensive review of LRP is provided in Nagy and Salhi (2007), who classified the literature on location-routing problems by distinguishing methodologies for deterministic variants, stochastic or dynamic problems, and versions with more complex networks (several echelons, inter-depot routes, etc.). Most recent literature on LRPs is presented by Lopes *et al*. (2013) and by Prodhon and Prins (2014).

2.5 Branch and price method

In this Section we review Branch and Price (BP), a significant solution method that have inspired some of the approaches developed to address the problems under study. Branch and Price is regarded as one of the most promising exact methods for addressing vehicle routing problems. It has, therefore, attracted considerable attention by the relevant community.

In a Branch and Price framework, a problem is solved using the Column Generation (CG) method by relaxing the integrality constraints of the original problem, and solving the resulting large linear program within a Branch and Bound (BB) tree (Desaulniers *et al*., 2005). Desrosiers *et al*. (1984) and Agarwal *et al*. (1989) addressed the VRPTW without capacity constraints, and the CVRP, respectively using BP. Following the work of Desrochers *et al*. (1992), a large number of references studied VRPTW using CG (Feillet *et al*., 2004; Chabrier, 2006). A related survey of this work is provided in Kallehauge *et al*. (2005).

In the present Section we first focus on the main component of BP, i.e. the CG method. Subsequently, we present the resource constrained shortest path problems that are commonly used as sub-problems within the CG, and we review the role of the BB in the overall method. Finally, we review references that incorporate BP with heuristic approaches to deal with well-known routing problems.

2.5.1 Column generation

Column generation is a decomposition technique for solving linear programing problems. This technique is used to solve large scale problems by using only a subset of the variables (or columns) of the model under consideration. The advantages of decomposing linear programming problems were initially pointed out by Ford and Fulkerson (1958), and the

decomposition scheme of Dantzig and Wolfe (1960) generalized the column generation method. Interested readers may refer to Lübbecke and Desrosiers (2005) for a formal historical review of the column generation method, and Desaulniers *et al*. (2005) for a vast number of column generation applications.

In Column Generation a linear programming problem is decomposed into two problems, the master problem and the pricing problem, which are solved iteratively. The master problem is a reduced problem in terms of the variable set, and the pricing problem seeks new variables to enhance this reduced variable set. To identify those new variables that improve the objective function, the pricing problem is guided by the dual variables of the master problem.

In order to better describe the column generation technique, consider the way the simplex method solves the following linear problem comprising (2.18) to (2.20). For this problem, let Φ be the set of all variables. In routing problems each variable could represent a feasible route, and, thus in this case Φ is the set of all feasible routes. In this fashion, assume that $p_\varphi, \varphi \in \Phi$ is a profit associated with route φ, and let $a_{l\varphi}$ receive the value 1 if route $\varphi \in \Phi$ visits location $l \in L$, where L is the set of all locations.

$$\max \sum_{\varphi \in \Phi} p_\delta \, x_\delta \tag{2.18}$$

subject to:

$$\sum_{\varphi \in \Phi} a_{l\varphi} \, x_\varphi \leq 1, \;\; l \in L \tag{2.19}$$

$$0 \leq x_\varphi \leq 1, \; \varphi \in \Phi \tag{2.20}$$

During the iterations of the simplex algorithm a feasible solution is maintained, which is called basic. In each iteration, a new variable (column) with negative reduced cost enters the basis aiming to improve the objective function. The reduced cost c_i for the i^{th} variable or the i^{th} column may be calculated as:

$$c_i = \sum_{l \in L} a_{li} \, \pi_l - p_i$$

where π_l are the dual variables of the current solution with respect to constraints (2.19). A common approach is to select the new entering column to be the one with the minimum reduced cost, i.e. $\arg\min_i \{c_i\}$. The simplex algorithm reaches the optimal solution when no column with negative reduced cost exists.

Column generation initiates its search by solving the above linear program considering only $\bar{\Phi}$, a subset of Φ for which the program has a feasible solution. The dual variables π_l of the latter solution may be utilized to seek new variables with negative reduced cost to be added in $\bar{\Phi}$. This may be achieved as in the simplex algorithm, i.e. by calculating all reduced costs for the columns $\Phi \backslash \bar{\Phi}$. Then, following the approach of simplex, the column with the most negative reduced cost may be added to $\bar{\Phi}$, and the linear program is solved again. As in the Simplex algorithm, if there is no such column, then the solution is optimal. Note that the latter solution is also optimal for problem (2.18) to (2.20) that considers the full set of variables Φ.

For the above approach all $\Phi \backslash \bar{\Phi}$ must be known, which is impractical for large scale problems. However, one could identify the column with the most negative reduced cost by solving a different problem, the pricing problem or the sub-problem. The pricing problem for the (2.18) to (2.19) problem is of the form:

$$\min\{c_i x_i | i \in \Phi\}$$

In routing problems the pricing problem is commonly modeled as a resource constrained shortest path problem, and is solved by labeling algorithms which are discussed below.

More information about column generation can be found in Wolsey (1998), Sigurd *et al.* (2004), Vanderbeck (2000), Lübbecke and Desrosiers (2005), Ralphs and Galati (2005).

2.5.2 Resource constrained shortest path problem

Consider a graph $G(V, A)$ where $V = \{0,1, \dots, n\}$ is the set of nodes and A is the set of arcs $(i, j), i, j \in V$. Also, consider a set of resources R, a cost value c_{ij} related to each arc (i, j), and let the consumption of resource $r \in R$ to be equal to l_{ij}^r for each arc. Then the resource consumption level of $r \in R$ for a path arriving at $j \in V$ from $i \in V$ may be designated as $T_j^r = f_{ij}^r(T_i^r, l_{ij}^r)$, where $f_{ij}^r(T_i^r, l_{ij}^r)$ is the related resource consumption function. Furthermore, constraints for each resource $r \in R$ may be imposed by a resource window $[a_i^r, b_i^r]$ for every node $i \in V$. Applying appropriately the resource consumption function and the resource windows, one may formulate a number of commonly applied constraints, such as: each node may be visited at most once (i.e. elementary paths), vehicle capacity, time of arrivals at each node, precedence relations etc.

The objective of the resource constrained shortest path problem is to identify the path that starts from node 0 and reaches node n with the minimum accumulated cost while satisfying

all resource constraints. Typically such problems are used as pricing problems in column generation schemes for routing problems; interested readers may refer to Feillet *et al*. (2004) for the VRPTW, to Boussier *et al*. (2007) for the TOP, and Sigurd *et al*. (2004) and Ropke and Cordeau (2009) for the PDPTW.

Resource constrained shortest path problems are addressed through dynamic programming approaches, called labeling algorithms. These algorithms are based on the Dijkstra (1959) and Bellman-Ford procedures. In label setting shortest path algorithms, a label $L = (i, (T_i^r, r \in R), p)$ comprises: a reference node i, the cumulated consumption T_i^r of the resources until that node, and a pointer p to its parent label used to construct the path from node 0 to node i. The algorithm initiates from a label corresponding to node 0 and gradually creates new labels for each possible partial path until the best feasible path(s) is obtained. These labels may be discarded during the process, if the resource constraints are violated or if they are dominated by other labels. The latter is identified by dominance rules that are applied by comparing two labels at a time.

Labeling algorithms have been used for routing problems in the context of column generation applied on both non-elementary shortest path problems (see Larsen, 2001) and elementary shortest path problems (Feillet *et al*., 2004; Boland *et al*., 2006; Righini and Salani, 2008; Chabrier, 2006). Interested readers may refer to Pugliese and Guerriero (2013) for a survey of recent advances in exact solution approaches for resource constrained shortest path problems.

2.5.3 Branch and bound

Branch and Bound (BB) is a technique that provides integer solutions dividing the feasible solution space into subspaces by adding new constraints (bounds). Each resulting problem (branch) is treated as a new problem from which new branches may be created. This procedure is repeated until the optimal integer solution is obtained. Interested readers may refer to Lawler and Wood (1966) and Lee and Mitchell (2001).

BB guarantees optimality if all columns are known, which is not the case when applying CG. In this latter case, optimality may be achieved by applying CG to each branch of BB, allowing in practice the unknown columns to be created. The latter procedure that combines CG with BB is the Branch and Price framework (Soumis, 1997; Barnhart *et al*., 1998; Danna and Le Pape, 2005).

2.5.4 Hybrid branch and price approaches

In addition to exact BP approaches, many researchers have proposed mixed ones in order to speed up the solution process (Blum *et al*., 2008; Raidl *et al*., 2010). Danna and Le Pape (2005) proposed a mixed scheme, which uses BP, an MIP heuristic solver and a metaheuristic, while Jepsen *et al*. (2008) used subset row inequalities (the Chvatal rank-1 cuts) in order to obtain better lower bounds through a branch-and-cut-and-price algorithm.

Prescott-Gagnon *et al*. (2009) introduced a branch-and-price based large neighborhood search algorithm for the VRPTW. For the same problem, Muter *et al*. (2010) implemented a hybrid tabu search heuristic, in which the tabu search produces feasible routes, and the search is guided by the current best lower and upper bounds.

More recently, Parragh and Schmid (2013) proposed a hybrid column generation with large neighborhood search algorithm to deal with the DARP. Their framework integrates the variable neighborhood search with column generation rules and combines it with a large neighborhood search.

2.6 Research contribution of the dissertation

Our contribution to the study of paratransit services under normal and emergency conditions involves insights in the design of novel transportation systems and the analysis of their special characteristics, as well as the introduction, formulation and solution of new routing problems.

In the area of paratransit services under normal conditions initially we propose and investigate a semi-flexible public transport system for door-to-door paratransit services. The scope of this system is to serve a portion of the paratransit demand, under quality of service restrictions, utilizing the scheduled public bus services. To address this scope we:

- Introduce a new routing problem, the Demand Responsive Bus Routing Problem (DRBRP), for which we propose a mathematical model that captures all aspects of the system under consideration
- Propose an exact Branch and Price (BP) framework that solves to optimality problems of practical size. In addition for the BP framework, we modify the labeling algorithm developed for the Elementary Shortest Path Problem to meet the special constraints of

DRBRP. The efficiency of the proposed exact method is investigated through extensive testing

- Study the design elements of the semi-flexible public transport system in three representative environments: urban, suburban, and rural. The design study focuses on the tuning of the system parameters to suit the characteristics of each environment.

We also contemplate an interesting new enhancement of the above system, in order to satisfy the total door-to-door paratransit demand. To do so, we supplement the service by accessible taxi (or private hire) services. For this study we:

- Introduce a generalization of the DRBRP that combines the DARP and the DRBRP. A new mathematical model is presented to formally describe this problem
- Implement an enhanced Branch and Price framework that provides exact solutions for the respective problem. This BP framework combines two types of sub-problems, one related to the bus service, and another related to the taxi mode
- Study the integrated system under different cost and quality of service limitations through extensive experimental testing to validate its applicability and to provide guidelines to system designers.

In the area of paratransit services under emergency conditions, we address the problem of planning bus-based evacuation in case of advance notice for an upcoming threat, as well as an interesting extension of this problem. Although considerable progress has recently been made in studying transit based evacuations for the general population, key issues remain to be investigated with respect to the populations in need of special transportation conditions. In this area, the dissertation contributions include the following:

- A new two-index formulation that describes the bus evacuation problem effectively, reducing the number of constraints and variables
- A hybrid column generation framework to address the problem efficiently. The method's efficiency is validated by comparing its solutions with the optimal solutions of a MILP solver. Furthermore, we examine ways to deal with the problem's high complexity by appropriately discretizing the demand at the pick-up locations, and assess the effects on the overall efficiency of the solutions
- Extensive testing of the proposed method to instances of practical scale to verify the consistency of the solutions and the adequacy of the required computational times.

We also study an interesting variation of the above problem that deals with the evacuation of casualties after an event. For this new problem the related contributions include:

- Modifying the two-index formulation for the bus evacuation problem and its objectives to describe the casualty evacuation problem
- Tuning accordingly the proposed hybrid framework to meet special characteristics of the casualties transportation problem
- Validating the efficiency of the algorithm by comparing its solutions to the optimal solutions of a MILP solver and by testing it to a wide range of instances of practical scale.

We have leveraged the learnings of the above problems to address interesting cases in logistics, which share some unique characteristics with the above paratransit problems, i.e. the existence of routes with strict sequences, and capacity limitations for the collection/gathering centers. Related contributions include:

- Modeling and analyzing three cases of the single vehicle routing problem with a predefined client sequence and two replenishment warehouse facilities
- Developing algorithms ranging from standard dynamic programming to labeling algorithms and a new partitioning heuristic
- Evaluating the performance of these algorithms using a large number of randomly generated problem instances and assessing the efficiency of the solutions with a commercial MILP solver.

Chapter 3. Scheduled paratransit transport systems (SPTS)

As discussed in Chapter 2, a significant trend in paratransit services is to leverage public transport to provide high quality, non-discriminative services to paratransit clients. This is precisely the focus of this Chapter. Specifically, we address the case of the Scheduled Paratransit Transport System (SPTS), which aims to serve paratransit requests by a scheduled, yet flexible, public bus service.

For this system, we propose a new routing problem, the Demand Responsive Bus Routing Problem (DRBRP), for which we develop a mathematical model to capture all its aspects, and an exact Branch and Price (BP) framework to solve to optimality problems of practical size. For the sub-problem of the BP framework, the labeling algorithm developed to address the related Elementary Shortest Path Problem is modified to meet the special constraints of DRBRP. Moreover, the design of SPTS is studied in three representative environments:

urban, suburban, and rural. The design study focuses on tuning the system parameters to suit the characteristics of each environment. Finally, the efficiency of the proposed exact method is investigated through extensive testing.

The structure of the rest of this Chapter is as follows: Section 3.1 describes the service delivered by SPTS and its unique aspects with respect to existing literature. Section 3.2 presents a mixed integer linear programming model for the DRBRP. Section 3.3 describes the proposed solution approach, while Section 3.4 presents the experimental study, and the resulting design guidelines.

3.1 Description of SPTS

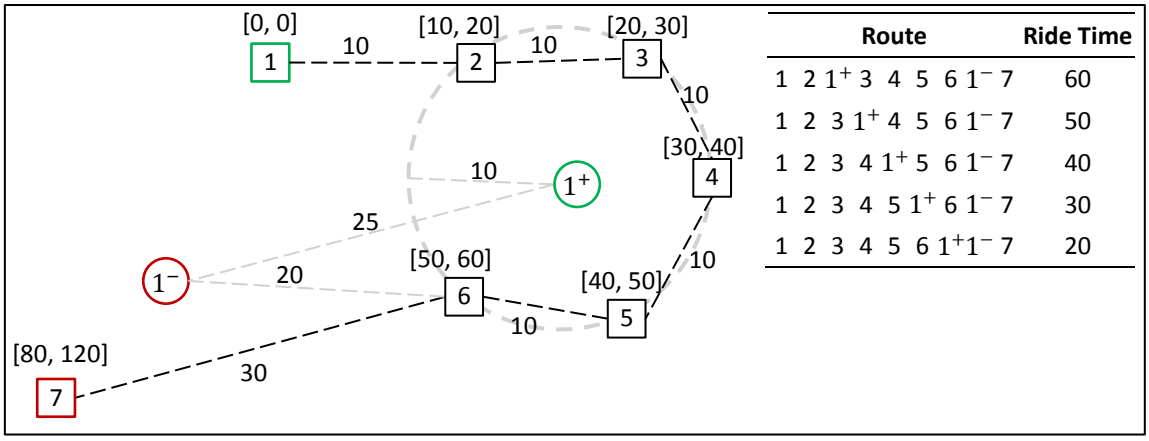
In SPTS, similarly to other dial-a-ride systems (see Paquette *et al.,* 2009), a paratransit request is booked prior to the start of the service period. Typically reservations may be placed for multiple trips to be performed within a week, or for a trip with a departure time within the next few hours. Each paratransit request specifies the desired pick-up and delivery points and the latest time of arrival at the destination. These paratransit requests will be served by allowing buses of a scheduled bus service to deviate from their nominal routes.

The SPTS route includes several predefined regular bus stops, which in every bus trip are visited in a prescribed sequence. The route is served by multiple bus trips (offset in time), and the earliest and latest bus departure times from each regular bus stop are defined per bus trip. For each trip we consider that the earliest departure time from a regular bus stop is equal to the minimum travel time of the bus between the start of the route (first regular bus stop) and the bus stop under consideration. The latest departure time in effect limits the deviation of the bus trip from the nominal route and maintains an adequate service level for the passengers waiting at the regular bus stops.

Each bus in the system has a certain capacity for paratransit clients. The paratransit requests are flexible, i.e. they may or may not be served. Each request is considered to be known before the start of the trip and comprises a pick-up location, a drop-off location, and the latest time that the client may reach its destination (drop-off location). This latest drop-off time is defined by the client. The pick-up time is considered in our case indirectly in the objective function of the problem (see below).

The problem consists of maximizing the total number of paratransit (flexible) requests served, and selecting among the solutions with equal number of served requests the one with

the minimum ride time for the paratransit clients. The latter is also an important issue, since the ride time affects the quality of transportation for the mobility impaired passengers, and affects the pick-up and drop-off times for such requests. The example of Figure 3.1 illustrates the necessity to consider the ride time of the paratransit requests. In this example the request for pick-up (1) may be served between bus stops 2, 3, 4, 5, 6 with almost the same deviation from the nominal route. However, the corresponding client will be offered the best service if s/he would be picked up after bus stop 6. This may be achieved by adding to the objective function a penalty term reflecting the total ride time of paratransit clients along the entire trip, as mentioned above.



Route	Ride Time
1 2 1+ 3 4 5 6 1− 7	60
1 2 3 1+ 4 5 6 1− 7	50
1 2 3 4 1+ 5 6 1− 7	40
1 2 3 4 5 1+ 6 1− 7	30
1 2 3 4 5 6 1+ 1− 7	20

Figure 3.1: Why paratransit ride time is important: 1^+ request pick-up, 1^- request drop-off

Overall, an important issue in designing SPTS is to balance the service level provided to the regular route clients and the service level provided to the paratransit clients. For the service level we consider the following aspects: (a) waiting time, and (b) travel or ride time. Importantly, only for the paratransit clients we consider the number of requests served by the system.

The proposed system deals with the service level of the regular route passengers by setting appropriate time limits. Specifically both aspects (a) and (b) are addressed by setting a time window for each bus stop. Thus, the maximum waiting time of a regular route passenger at an origin bus stop is equal to the time window corresponding to this bus stop. Furthermore, the maximum ride time is always less than the difference of the closing of the destination time window minus the opening of the origin time window.

For the service level of the paratransit clients and for aspect (a), as waiting time we consider the time that a client may arrive at his/her destination earlier to the desired time. The latter is addressed by setting a limit for this earliest time. Thus, the maximum waiting time of a

paratransit client at the destination is equal to the difference of the latest and the earliest times that the client may be transported to the destination. The ride time – aspect (b) – and the number of paratransit requests are dealt in the problem objective as described above.

Another issue that arises from the aforementioned considerations is that during a trip the bus may wait before visiting a paratransit request's location so that the ride time of this request is reduced. This is undesirable since it increases the ride time of all passengers aboard and also increases the waiting time at the bus stops of the regular route. Therefore, we do not allow a bus to wait at any location unless is necessary for the opening of the time window of a paratransit client.

Table 3.1 summarizes how the aforementioned service level aspects of SPTS are incorporated into the DRBRP model described below.

Table 3.1: How DRBRP treats service levels

Service level	Regular route passengers	Paratransit
Waiting time	Constraint	Constraint
Ride time	Constraint	Objective function
Available seats	Constraint	Objective function

3.2 Mathematical formulation of DRBRP

The DRBRP has been formulated as a linear mixed integer problem. We consider the following notation:

- $G(N, A)$ is a directed graph where N is the set of all nodes related to the problem, and A is the set of arcs that connects the nodes
- Let m be the number of the regular bus stops, and v be the total number of bus trips for a certain period, where $K = \{1, \dots, v\}$ is the set of bus trips
- A node is defined per predefined regular bus stop and per bus trip. Thus, let $B = \{1, \dots, m \cdot v\}, B \subset N$ be the set of nodes that model the regular bus stops of the network; the sequence of serving these regular bus stops must be preserved
- Let $B_k = \{1 + m \cdot (k-1), \dots, m + m \cdot (k-1) | \, k \in K\}$ be the set of regular bus stop nodes served during bus trip k
- Let n be the number of paratransit requests. In terms of nodes a paratransit request is characterized by its pick-up node and its delivery node

- Let $S = \{1 + m \cdot (k-1) | \, k \in K\}$ be the set of nodes that correspond to the first regular bus stop of all bus trips
- Let $E = \{m + m \cdot (k-1) | \, k \in K\}$ be the set of nodes that correspond to the last nodes of all bus trips
- Let $C \subset N$ be the set of nodes that correspond to the paratransit requests. In particular $C = C^+ \cup C^-$ where $C^+ = \{m \cdot v + 1, \dots, m \cdot v + n\}$ contains the pick-up nodes of the paratransit requests while $C^- = \{m \cdot v + n + 1, \dots, m \cdot v + 2n\}$ contains the delivery nodes of the paratransit requests. For a paratransit request with pick-up node $i \in C^+$ the respective delivery node is $i + n \in C^-$
- Following the above definitions, the set of nodes N is defined as $N = \{1,2, \dots, m \cdot v + 2n\} = B \cup C$
- Let $\bar{N} = \bigcup_{k \in K} \{(i,k) | i \in C \cup B_k\}$ be a set of pairs comprising a node i and a trip k. The pairs in $\bar{N}$ are necessary in order to identify which trip k can serve which nodes. In case a node i can be served by multiple trips, then multiple pairs are defined for this node, accordingly
- Let $\bar{A}$ be a set of triplets (i,j,k) that comprise an arc (i,j) and a trip k. Specifically, $\bar{A} = \bigcup_{k \in K} \{(i,j,k) | i \in (S \cap B_k), j \in \{i+1\} \cup C^+\} \cup \bigcup_{k \in K} \{(i,j,k) | i \in B_k \backslash (E \cup S), j \in \{i+1\} \cup C\} \cup \bigcup_{k \in K} \{(i,j,k) | i \in C^+, j \in (C \cup B_k) \backslash (S \cup E \cup \{i\})\} \cup \bigcup_{k \in K} \{(i,j,k) | i \in C^-, j \in (C \cup B_k) \backslash (S \cup \{i-n, i\})\}$. Again in case the arc (i,j) can be traversed by multiple trips, then multiple triplets are defined for this pair. Triplets (i,j,k) in $\bar{A}$ are necessary in order to safeguard that a feasible route consists of arcs belonging to the same trip.

Figure 3.2 illustrates the various sets for a case in which $m = 3$, $n = 1$ and $v = 2$. In this case $B = \{1,2,3,4,5,6\}$, $C = \{7,8\}$, $C^+ = \{7\}$, $C^- = \{8\}$, $B_1 = \{1,2,3\}$, $B_2 = \{4,5,6\}$, $S = \{1,4\}$ and $E = \{3,6\}$. Note that nodes (1,4), (2,5) and (3,6) represent the same physical bus stops respectively and that the two arrows from node 7 to node 8 are the arcs (7,8,1) and (7,8,2).

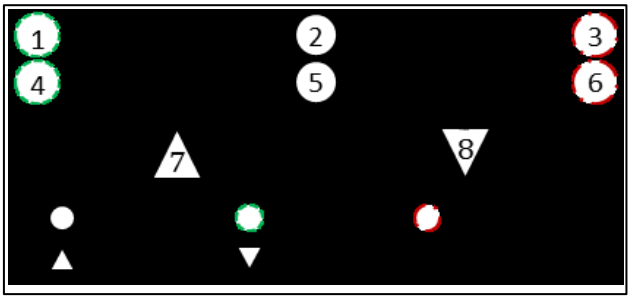


Figure 3.2: Illustration of the graph of a simple example

Continuing the notation, consider that:

- $[a_i, b_i], i \in N$ is the time interval within which each node must be served. For simplicity and without loss of generality we consider that the width $b_i - a_i$ of each time window is equal for every $i \in B$; $a_i, i \in S$ should be defined according to the designated headway between bus trips
- $t_{ij}, i, j \in N$ is the time required to travel directly from node i to node j. We consider that the travel times satisfy the triangle inequality
- $d_i, i \in N$ is the demand of paratransit passengers of each node. If it is a regular bus stop then $d_i = 0, i \in B$. If it is a pick-up node $d_i = 1, i \in C^+$. If it is a delivery node $d_i = -1, i \in C^-$
- Q is the maximum number of paratransit passengers that any vehicle $k \in K$ may carry. Note that only the capacity of the bus in paratransit passengers is of concern, since the regular bus services are assumed to be designed in a way which balances demand and capacity
- $x_{ijk}, (i, j, k) \in \bar{A}$ is assigned the value 1 if the arc from node i to node j is traversed during bus trip k, otherwise 0 is assigned to the variable
- $s_{ik}, (i, k) \in \bar{N}$ is the time that each node i is served during bus trip k
- $q_{ik}, (i, k) \in \bar{N}$ is the number of paratransit passengers on board during trip $k \in K$ immediately after node i is served
- $RT_{ik}, i \in C^+, k \in K$ is the total ride time of paratransit request $i \in C^+$ in bus trip k
- $w_i, i \in N$ is a binary variable that is used for linearization purposes.

The single objective of the problem includes two terms, each addressing a significant service level issue of SPTS, as discussed in Section 4.1: The first term is the number of paratransit requests served, while the second term is the total ride time of these requests. We consider the first term of the objective function as more important than the second one. Accordingly we treat the two terms in a lexicographic order that first maximizes the number of requests served; among the solutions with the same number of requests, it selects the one that minimizes the total ride time of the selected requests. Thus, the objective function is defined as follows:

$$TP = max \sum_{k \in K} \sum_{i \in C^+} \sum_{\{j | (i,j,k) \in \bar{A}\}} x_{ijk} - P \sum_{k \in K} \sum_{i \in C^+} RT_{ik} \tag{3.1}$$

- The value of the penalty constant is given by $P = 1/(n \cdot T_{max})$, where $T_{max} = b_m - a_1 = b_{2\cdot m} - a_{m+1} = \cdots = b_{v\cdot m} - a_{(v-1)\cdot m+1}$ is the maximum ride time of a bus trip. This value guarantees that the last term in Eq. (3.1) is less than 1 and, thus, the maximum possible number of requests will be served irrespective of the savings in total ride time. Furthermore, RT_{ik} for unserved requests is zero as defined below.

Subject to:

$$\sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk} = 1, \quad i \in B_k \backslash E, k \in K \tag{3.2}$$

$$\sum_{\{i|(i,j,k)\in\bar{A}\}} x_{ijk} = 1, \quad j \in (E \cap B_k), k \in K \tag{3.3}$$

$$\sum_{\{(j,k)|(i,j,k)\in\bar{A}\}} x_{ijk} \leq 1, \quad i \in C^+ \tag{3.4}$$

$$\sum_{\{i|(i,h,k)\in\bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k)\in\bar{A}\}} x_{hjk} = 0, \quad h \in C \cup (B_k \backslash S \cup E), k \in K \tag{3.5}$$

$$\sum_{\{i|(i,h,k)\in\bar{A}\}} x_{ihk} - \sum_{\{j|(j,h+n,k)\in\bar{A}\}} x_{j,h+n,k} = 0, \quad h \in C^+, k \in K \tag{3.6}$$

$$s_{i-1,k} + t_{i-1,i} \leq s_{ik}, \quad i \in B_k \backslash S, k \in K \tag{3.7}$$

$$s_{ik} + t_{i,i+n} \sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk} \leq s_{i+n,k}, \quad i \in C^+, k \in K \tag{3.8}$$

$$s_{i+n,k} - s_{ik} = RT_{ik}, \quad i \in C^+, k \in K \tag{3.9}$$

$$s_{ik} + t_{ij} - M(1 - x_{ijk}) \leq s_{jk}, \quad (i,j,k) \in \bar{A} \tag{3.10}$$

$$s_{jk} \leq a_j + (b_j - a_j)w_j + M(1 - x_{ijk}), \quad (i,j,k) \in \bar{A} \tag{3.11}$$

$$s_{jk} \leq s_{ik} + t_{ij} + a_j(1 - w_j) + M(1 - x_{ijk}), \quad (i,j,k) \in \bar{A} \tag{3.12}$$

$$a_i \sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk} \leq s_{ik} \leq b_i \sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk}, \quad i \in C, k \in K \tag{3.13}$$

$$a_i \leq s_{ik} \leq b_i, \quad i \in B_k, k \in K \tag{3.14}$$

$$q_{ik} + d_j - M(1 - x_{ijk}) \leq q_{jk}, \quad (i,j,k) \in \bar{A} \tag{3.15}$$

$$0 \leq q_{ik} \leq Q \sum_{\{j|(i,j,k)\in\bar{A}\}} x_{ijk}, \quad (i,k) \in \bar{N} \tag{3.16}$$

$$\sum_{i\in S\cup E} \sum_{\{k|(i,k)\in\bar{N}\}} q_{ik} = 0 \tag{3.17}$$

$$x_{ijk} \in \{0,1\}, \quad (i,j,k) \in \bar{A} \tag{3.18}$$

$$w_i \in \{0,1\}, \quad i \in N \tag{3.19}$$

$$q_{ik} \in \mathbb{N}^0, \quad (i,k) \in \bar{N} \tag{3.20}$$

$$s_{ik} \geq 0, \quad (i,k) \in \bar{N} \tag{3.21}$$

$$RT_{ik} \geq 0, \quad i \in C^+, k \in K \tag{3.22}$$

Constraints (3.2) and (3.3) ensure that all regular bus stops will be served once per bus trip. Constraint (3.4) ensures that the paratransit pick-up nodes (and thus the respective paratransit requests) may be served once. Equation (3.5) ensures continuity for all nodes except those corresponding to the starting and ending regular bus stops of each bus trip. Equation (3.6) ensures that if a paratransit request is served, then the corresponding pick-up and delivery nodes will be served during the same trip. Constraint (3.7) ensures that regular bus stops will be served with the predefined sequence. Constraint (3.8) ensures that the paratransit delivery nodes will be served after their related pick-up nodes. Constraint (3.9) defines the ride time of paratransit request $i \in C^+$. Note that in case request i is not served $RT_{ik} = 0$ due to Constraint (3.13), and, thus no penalty is added to the objective function. Constraint (3.10) defines the time flow, assuming that $M \gg 1$.

Constraints (3.11) and (3.12) ensure that if a node is served, then the service will be performed at the earliest possible time. Specifically, Constraints (3.10) to (3.14) linearize the following expression, which models the policy for the waiting times mentioned in Section 3.1. The service time of a node is the maximum of the opening of the time window and the minimum time to arrive at that node.

$$x_{ijk} = 1 \implies s_{jk} = \max\{a_j, s_{ik} + t_{ij}\}, (i,j,k) \in \bar{A} \tag{3.23}$$

Note that in case $w_j = 0$ then $s_{jk} = a_j$, otherwise if $w_j = 1$, then $s_{jk} = s_{ik} + t_{ij}$. Constraints (3.13) and (3.14) ensure that if a node is served, this will occur in the predefined time interval $[a_i, b_i]$. Constraint (15) defines the change in the paratransit load of the bus. Constraint (3.16) ensures that the paratransit bus load is greater than zero and less than or equal to Q. Constraint (3.17) designates that every trip starts and ends with an empty bus. Finally, Constraints (3.18) and (3.19) define the binary nature of decision variables x_{ijk} and w_i, respectively, Constraint (3.20) ensures that q_{ik} is a positive integer or zero, and Constraints (3.21) and (3.22) ensure that variables s_{ik} and RT_{ik} will receive values greater or equal than zero.

Concerning the complexity of the above model, consider that Golden *et al.* (1987) have shown that the Orienteering Problem (OP) belongs to the NP-hard class of problems by reducing OP to the Traveling salesman problem. Laporte and Marcelo (1990) also proved that OP belongs to the class of NP-hard problems by reducing it to the Hamiltonian circuit problem. By applying the same reduction of Golden *et al.* (1987), one may deduce that an orienting problem with pick-ups and deliveries is in the same class of complexity as the Pick-up and Delivery Problem (PDP). Specifically, setting the time limit of an Orienteering Problem with Pick-up and Deliveries (OP-PDP) to the exact time needed to serve all clients with the minimum travelling time, the solution of the PDP and the OP-PDP will be the same, and thus, the two problems are of the same complexity. Hence, since the PDP is an NP-hard problem, then the OP-PDP belongs to the same problem class in terms of complexity. DRBRP may be also reduced to an OP-PDP by considering a problem case in which only two regular bus stops (e.g. the first, the last) represent the origin and destination depot. Thus, the DRBRP also belongs to the same complexity class as OP-PDP and is NP-hard.

3.3 A Branch and price approach for the DRBRP problem

In a Branch and Price framework, a problem is solved using the column generation method by relaxing the integrality constraints of the original problem, and solving the resulting large linear program within a branch and bound tree (Desaulniers *et al*., 2005). Column generation comprises two phases that are applied in an iterative manner. The first phase solves the Restricted Master Problem (RMP), which contains a subset of all feasible variables. The second phase solves one or more Sub-Problems (SP), generating additional variables to be added to the RMP. The procedure ends when no more "profitable" variables can be generated by the sub-problem(s). In routing problems, the RMP is usually formulated as a set partitioning problem, and the sub-problem is formulated as a resource constraint shortest path problem. We apply the same principles here.

3.3.1 The set partitioning problem for the DRBRP

In order to formulate the problem presented in Section 3.2 as a set partitioning problem, let δ_k be a set that contains all feasible routes of a single bus trip k, and let $\delta = \delta_1 \cup \delta_2 \cup \ldots \cup \delta_v$ be the set of all possible feasible routes of all bus trips. Each route in δ_k must satisfy constraints (3.2)-(3.22) by setting $K = \{k\}$. For each route $j \in \delta$ define the set A_j of the paratransit requests served by route j. a_{ij} is a binary coefficient that takes the value 1 if

$i \in A_j$. Let rt_{ij} be the ride time of request i served by route j, which is similar but not identical to RT_{ik} of the formulation of Section 3.2; the latter refers to bus trip k and not to route j. Then the profit p_j of route j can be calculated as follows:

$$p_j = \sum_{i \in A_j} a_{ij} - P \sum_{i \in A_j} rt_{ij} \tag{3.24}$$

where P is the penalty of Eq. (3.1). Let y_j be the binary variable which is assigned the value 1 if route j is used, and 0 if not. Then the problem may be formulated by the objective function:

$$max \sum_{j \in \delta} p_j y_j \tag{3.25}$$

Subject to:

$$\sum_{j \in \delta} a_{ij} y_j \leq 1, \quad i \in C^+ \tag{3.26}$$

$$\sum_{j \in \delta_k} y_j = 1, \quad k \in K \tag{3.27}$$

$$y_j \in \{0,1\}, \quad j \in \delta \tag{3.28}$$

Constraint (3.26) ensures that all paratransit requests (pick-up nodes) may be served at most once. Constraint (3.27) ensures that one route must be used per bus trip; thus all regular bus stops will be served once during each bus trip.

3.3.2 The sub-problem for the DRBRP

In the case under consideration one may define $|K|$ sub-problems (SP), one for each bus trip $k \in K$. Each SP will produce additional routes to be added to $\delta'_k \subseteq \delta_k$. The objective function of the SP is to identify the route with the maximum reduced cost of the problem described in the previous section. However, it is more efficient to identify more than one routes with positive reduced cost; these are designated as the routes j satisfying the following condition:

$$p_j - \sum_{i \in A_j} a_{ij} \pi_i - \bar{\pi}_k > 0 \tag{3.29}$$

where π_i are the dual prices related to constraint set (3.26) and $\bar{\pi}_k$ are the dual prices related to constraints (3.27). At the same time, these routes should satisfy constraints (3.2) - (3.22) by setting $K = \{k\}$.

Below we simplify the notation for the SP, which from now on will refer to one bus trip $k \in K$, as follows: Let $\widetilde{N} = B_k \cup C^+ \cup C^-$ be the set of all nodes. Its subset B_k contains the nodes that correspond to the regular bus stops of bus trip k, while $C^+$and C^- are also sub-sets of $\widetilde{N}$ and correspond to the pick-up and drop-off locations of the paratransit requests, respectively.

The SP may be considered as an Elementary Shortest Path Problem with Time Windows, Capacity, Pick-up and Delivery as well as Sequence Restrictions. Hereafter we use the term SDRBRP to represent the sub-problem of DRBRP. To solve such resource constraint routing problems, labeling algorithms are commonly applied; see Feillet *et al.* (2004), Irnich and Desauliners (2005). In fact, this is the critical issue in Column Generation techniques, since the performance of the labeling algorithm affects directly the efficiency of solving the relaxed problem, and thus, the size of problems that may be solved to optimality. In the following subsection, we extend the labeling algorithm presented in Ropke and Cordeau (2009).

3.3.3 Labeling algorithm for the SDRBRP

The proposed approach to solve the SDRBRP adapts the labeling algorithm of Ropke and Cordeau (2009) for solving the Pick-up and Delivery problems with Time Windows. We have modified their algorithm to meet the special characteristics of the SDRBRP and we present a new structure to solve it. Ropke and Cordeau had, in turn, modified the original labeling algorithm proposed by Feillet *et al.* (2004) for the Elementary Shortest Path Problem with Time Windows. The proposed algorithm for the SDRBRP starts from the first regular bus stop and progressively extends all feasible paths until they reach the last regular bus stop. Dominance rules are applied during the latter procedure and paths of inferior quality are discarded without extending them.

Let L be a label that contains information about the consumption of resources along the related path. $L = [ln, pbs, t, rt, h, p, V, O]$, where:

- ln: is the last node of the path
- pbs: is the last regular bus stop visited by the path
- t: is the arrival time at node ln
- rt: is the total ride time of all paratransit requests belonging to the path up to this point
- h: is the number of paratransit passengers on the bus
- p: is the current profit corresponding to the path up to this point

- $V \subseteq \{1, \dots, n\}$: is the set of paratransit requests for which the passenger has been picked up (but possibly not dropped off)
- $O \subseteq \{1, \dots, n\}$: is the set of those paratransit requests for which the passenger has been picked up but not dropped off. Requests O are said to be *open* or *active*

Considering the extension of a label $l = [ln, pbs, t, rt, h, p, V, O]$ to node j one of the following conditions must hold:

$$m < j \leq n \wedge V(l) \cap \{j\} = \emptyset \tag{3.30}$$

$$n < j \leq 2n \wedge j \in O(l) \tag{3.31}$$

$$j < m \wedge pbs(l) = j - 1 \tag{3.32}$$

$$j = m \wedge O(l) = \emptyset \wedge pbs(l) = j - 1 \tag{3.33}$$

where m is the number of regular bus stops and n is the number of paratransit requests. Conditions (3.30) to (3.32) are similar to those used in the Elementary Shortest Path Problem with Time Windows, Capacity, and Pickup and Delivery (ESPPTWCPD) (Ropke and Cordeau, 2009). Condition (3.30) ensures that in case j is a pick-up node it will be added to the path only if the node has not been added previously. In case j is a delivery node, it will be added to the path only if its corresponding pick-up node has been visited previously (3.31). Condition (3.32) defines that a regular bus stop is added to the path only if it is the next one according to the predefined sequence; it also ensures that a regular bus stop is added to the path only if it has not been served previously. The final destination node, which is the last regular bus stop of the sequence, will be added to the path only if no open request exists and all other regular bus stops are served, as defined by Condition (3.33).

If the aforementioned conditions hold, the information of label l_{new} of the path extended to node j is computed as follows:

$$ln(l_{new}) = j \tag{3.34}$$

$$pbs(l_{new}) = \begin{cases} j, & j \in B_k \\ pbs(l), & j \in C^+ \cup C^- \end{cases} \tag{3.35}$$

$$t(l_{new}) = \max(t(l) + t_{n(l),j}, a_j) \tag{3.36}$$

$$rt(l_{new}) = rt(l) + |O(l)| \cdot t_{n(l),j} \tag{3.37}$$

$$h(l_{new}) = h(l) + d_j \tag{3.38}$$

$$p(l_{new}) = \begin{cases} p(l), & j \in B_k \\ p(l) + \bar{p}_{j,} & j \in C^+ \\ p(l) - P \cdot rt_{j,} & j \in C^- \end{cases} \tag{3.39}$$

$$V(l_{new}) = \begin{cases} V(l) \cup \{j\}, & j \in C^+ \\ V(l), & j \in C^- \cup B_k \end{cases} \tag{3.40}$$

$$O(l_{new}) = \begin{cases} O(l), & j \in B_k \\ O(l) \cup \{j\}, & j \in C^+ \\ O(l) \backslash \{j - n\}, & j \in C^- \end{cases} \tag{3.41}$$

Label l_{new} will be discarded if $h(l_{new}) > Q$ or $t(l_{new}) > b_j$, which are the constraints regarding the capacity of the vehicle and the latest time of arrival.

In order to discard dominated labels, dominance rules are established between pairs of labels ending to the same node. These rules ensure that all extensions of the dominated labels do not lead to a better solution than the one of the dominant labels. A label l_1 dominates a label l_2 if all criteria defined by Eqs. (3.42) to (3.48) are satisfied:

$$ln(l_1) = ln(l_2) \tag{3.42}$$

$$pbs(l_1) = pbs(l_2) \tag{3.43}$$

$$t(l_1) \leq t(l_2) \tag{3.44}$$

$$p(l_1) \geq p(l_2) \tag{3.45}$$

$$V(l_1) \subseteq V(l_2) \tag{3.46}$$

$$O(l_1) \subseteq O(l_2) \tag{3.47}$$

$$rt(l_1) \leq rt(l_2) \tag{3.48}$$

Criteria (3.42), (3.44), (3.46) and (3.47) have been used in case of the ESPPTWCPD (Ropke and Cordeau, 2009). We have introduced criteria (3.43), (3.45), and (3.48) to address SDRBRP. To show that the dominance rules guarantee optimality, one must show that: (a) the possible extensions of the dominant label form a superset of all possible extensions of the dominated one, (b) if the dominant and dominated labels extend to the same node, then the dominant one would lead to lower or equal resource consumption.

Statement (a) above is guaranteed if the same amounts of resources are available for use for both labels when these are extended. Consider that both labels concern the same node i (Eq. 3.42). Then if both labels l_1 and l_2 are extended to node j, according to Eq. (3.44) if l_2 is feasible, l_1 will also be feasible, given that the time to traverse arc (i, j) is the same for both

cases. The same consideration can be made for cumulative profit and the total ride time, i.e. resources p and rt, respectively.

These latter considerations also safeguard that statement (b) is valid. It is noted that the same nodes are available when extending either of the two paths, since $V(l_1)$ is a subset of $V(l_2)$ and due to Conditions (3.30) to (3.32), which ensure that all nodes will be visited once. Similarly, due to Condition (3.31), drop-off nodes are available for extension only if the respective pick-up node belongs to the path. In SDRBRP, regular bus stops must be served following a certain sequence, which is expressed by Condition (3.33). l_1 will be able to be extended to the same regular bus stop as l_2, only if Eq. (3.43) is satisfied.

Algorithm 3.1 presents the pseudo-code of the proposed labeling algorithm for the SDRBRP; we call this algorithm "Bus Stop Structure" (BSS). BSS capitalizes on the fact that: (a) a path is feasible only if the bus stop sequence is preserved, and (b) paths can be dominated only by paths that have reached the same regular bus stop. BSS solves SDRBRP by extending all possible paths until the next bus stop and only then discards labels due to dominance criteria (3.44) to (3.48), since criteria (3.42) and (3.43) are already satisfied. Furthermore, if an extension to the next regular bus stop of the sequence is not feasible, then the label is also discarded and no new label is generated.

Algorithm 3.1: Labeling Algorithm – Bus Stop Structure

```
1.  Initialize lists of labels L, L_new, L_final, L_nb
2.  Add first label to L_nb
3.  While L_nb ≠ ∅
4.    L = L_nb, L_nb = ∅
5.    While L ≠ ∅
6.      l = remove(L)
7.      Extend l to next regular bus stop
8.      If extension is feasible
9.        If final bus stop
10.         Add new label to L_final
11.       else
12.         Add new label to L_nb for the next regular bus stop
13.         Add one new label for every feasible pick-up node to L_new
14.         Add one new label for every delivery node to L_new
15.     L = L_new, L_new = ∅
16.   Remove dominated labels form L_nb
17. Return path corresponding to the best label in L_final
```

3.3.4 Branch and bound

Given that column generation does not guarantee the integer solution of the problem, in our approach we implement a typical branch and bound scheme. That is, if the solution provided

by the column generation algorithm contains a variable with a non-integer value, then two new branches are created, each corresponding to a problem that uses a different integer value for this variable. Column generation is applied to solve the problems created in each of these branches. In our case we branch on flow variables based on the value of the variable x_{ijk}. These new branches will include constraints $x_{ijk} = 1$ and $x_{ijk} = 0$ respectively. The variable x_{ijk} that is closest to 0.5 is selected for branching. Interested readers may refer to Desaulniers *et al. (*2005) for alternative branching strategies.

3.4 Tests and results

In this Section we investigate the following:

i. The performance of the proposed Branch and Price approach using as a benchmark a related generic algorithm from the literature
ii. The effects of key SPTS design parameters on system performance
iii. The performance of the SPTS system under conditions that characterize three typical mobility environments.

For the performance studies (i) and (ii) above we consider that the design of the SPTS system is conducted in two stages. The first stage concerns the regular transport service and defines the number m of regular bus stops, the average distance ρ between sequential stops and the headway f of the service (or the number of trips k within a certain period). The values of these parameters depend directly on the mobility environment, i.e. the characteristics of the route and the demand for the regular service. Standard methods are used to define them.

The resulting values are inputs to the second stage, in which the SPTS service is designed in order to satisfy the paratransit demand (number of requests n within a certain time period). The parameters to be defined in this second stage are the focus of our study and include the number of reserved seats Q for paratransit clients per bus, the maximum allowable time deviation d of a bus trip from the nominal schedule, and the width ξ of the drop-off time windows for the paratransit requests. Following this design framework, in study (ii) we consider the effects of Q, d, ξ on system performance given certain values for the parameters of the regular service m, ρ, k, f. Having drawn conclusions on these effects, we study the SPTS design in the typical mobility environments mentioned above.

To provide test instances for all three studies, we developed a random test data generator. The basic input parameters of the generator are those discussed above, i.e. m, f, ρ, u (the nominal vehicle speed), n ξ, Q, and d the coefficient of the allowed delay of each bus; i.e. a bus may serve a bus stop at most $(d \cdot \rho)/u$ after the designated earliest time of arrival to it. Thus, d defines the latest arrival time at each regular bus stop. Note that this imposes a limit to the total trip delay accrued up to each bus stop (in the formulation of Section 3.2 the allowed delay is addressed by setting appropriately the time windows of the regular bus stops). In addition, note that this delay is cumulative with the progression of the bus trip; thus if a bus would deviate from its route at the start of its trip, it would result in less allowed time for deviation for the latter part of the trip.

Another significant input of the generator is the allowable region in which the pick-up and drop-off locations of the paratransit requests may be generated along the service route. The generator considers two cases of allowable regions: ellipses, and circles. For the former case an ellipse is defined considering two sequential regular bus stops as foci and ω as the length of its major axis. For the latter case, a circle is defined considering the location of a regular bus stop as the center and $\bar{\omega}$ as its radius. The generator produces randomly: (a) the locations for the bus stops of the regular route, and (b) the locations of the pick-up and drop-off points for the paratransit requests. To do so, it assumes appropriate probability distributions with respect to parameters ρ and $\omega/\bar{\omega}$ respectively. The exact generation of the test instances is further described in Appendix A

3.4.1 Computational study I: Algorithm efficiency

This first study compares the performance of BSS, the algorithm of Section 3.3, with a suitably adapted version of the generic labeling algorithm proposed in the literature to solve ESPPTWCPD (Ropke and Cordeau, 2009). This latter algorithm (hereafter called SS) essentially shares the same dominance rules with BSS. The difference between the two algorithms is due to the algorithmic steps they follow. Specifically, based on the characteristics of SDRBRP, the BSS algorithm does not need to validate criteria (3.42) and (3.43), since these are satisfied by construction.

The experiments investigate the effects of the following key parameters on computational time and on the maximum size of the problem each algorithm is able of solve:

- $d \in \{0.2, 0.5, 0.8\}$: the maximum allowable delay coefficient

- $m \in \{10, 20\}$: the number of regular bus stops
- $n \in \{10, 15, 20\}$: the number of paratransit requests
- $k \in \{10, 20\}$: the number of bus trips.

The pick-up and drop-off locations have been distributed along the route uniformly within a corridor created by overlapping ellipses. The number of experiments generated is $3 \times 2 \times 3 \times 2 = 36$ and each randomly generated experiment was solved five times (36 experiments $\times$ 5 repetitions/experiment $=$ 180 instances in total) using both methods.

For each algorithm, Table 3.2 presents the average computational time over the five repetitions per experiment, as well as the number of times an experiment did not reach a solution within a run time limit of 2 hours on a PC equipped with an Intel Xeon E5335 CPU, and 8 GB of RAM.

Table 3.2: Comparison of BSS and SS performance

			$d = 0.2$					$d = 0.5$					$d = 0.8$				
			CPU[1]			Failures[2]		CPU[1]			Failures[2]		CPU[1]			Failures[2]	
k	m	n	BSS	SS	%Sav.[3]	BSS	SS	BSS	SS	%Sav.[3]	BSS	SS	BSS	SS	%Sav.[3]	BSS	SS
10	10	10	2.3	2.9	23.4%	-	-	33.3	87.1	61.7%	-	-	111.3	316.1	64.8%	-	-
		15	21.2	33.9	37.5%	-	-	334.6	667.4	49.9%	-	-	1827.2	2132.8	14.3%	-	1
		20	263.7	394.7	33.2%	-	-	1071.0	1388.3	22.9%	2	3	-	-	-	5	5
	20	10	14.0	20.1	30.6%	-	-	43.2	131.7	67.2%	-	-	111.7	418.1	73.3%	-	-
		15	63.2	204.7	69.1%	-	-	501.5	1151.1	56.4%	-	1	1281.3	4821.6	73.4%	-	2
		20	145.7	198.7	26.7%	-	-	1296.9	1085.8	-19.4%	2	3	-	-	-	5	5
20	10	10	2.7	3.5	22.0%	-	-	53.5	116.8	54.2%	-	-	232.9	550.8	57.7%	-	-
		15	69.0	98.6	30.0%	-	-	974.6	2451.8	60.3%	-	-	920.6	1241.1	25.8%	2	3
		20	305.3	648.5	52.9%	-	-	650.7	1478.5	56.0%	1	4	-	-	-	5	5
	20	10	23.9	27.1	11.8%	-	-	107.3	312.1	65.6%	-	-	422.9	2106.2	79.9%	-	-
		15	124.9	232.8	46.3%	-	-	1448.7	4809.1	69.9%	-	2	-	-	-	3	5
		20	611.4	1002.5	39.0%	-	-	-	-	-	4	5	-	-	-	5	5
Average			**137.3**	**239.0**	**42.6%**	**-**	**-**	**592.3**	**1243.6**	**52.4%**	**-**	**-**	**701.1**	**1655.3**	**57.6%**	**-**	**-**
Sum			**-**	**-**	**-**	**0**	**0**	**-**	**-**	**-**	**9**	**18**	**-**	**-**	**-**	**25**	**31**

[1] Average computational time in seconds for problem instances that both algorithms solved, [2] Number of cases not solved, [3] Percent of savings of BSS CPU over SS CPU

In almost all cases BSS obtained the optimal solution in less computational time than SS; on the average BSS required about half the computational time of SS. Furthermore, BSS

managed to solve more problem instances, an indication that BSS can deal with problems of larger scale.

3.4.2 Computational study II: Effects of constraints and parameters

To investigate the effect of significant SPTS constraints and parameters on the efficiency of the problem's solutions, we focus on the SPTS-specific parameters d, Q, ξ. For this study, $m = 10$, $k = 5$, $\rho = 500\text{m}$, $u = 25\text{km/h}$ and $\omega = 500\text{m}$.

Consider first the case in which the capacity of the vehicle is not constrained, and there are no limits regarding the latest drop-off times of the paratransit requests. For this case, Figure 3.3 presents the portion of requests served (as % of total), and the service level for $d = \{0.1, 0.3, 1\}$, and n increasing from 1 to 8 (the ratio n/k of requests per trip increases from 0.2 to 1.6). The service level value is provided as the percentage of the maximum possible value of the objective function; that is, the value obtained by serving all paratransit requests with the minimum ride time per request, i.e. travelling form a pick-up location directly to the respective drop-off location. As expected when d increases the service level also increases. For $d = 1$ all paratransit requests are served.

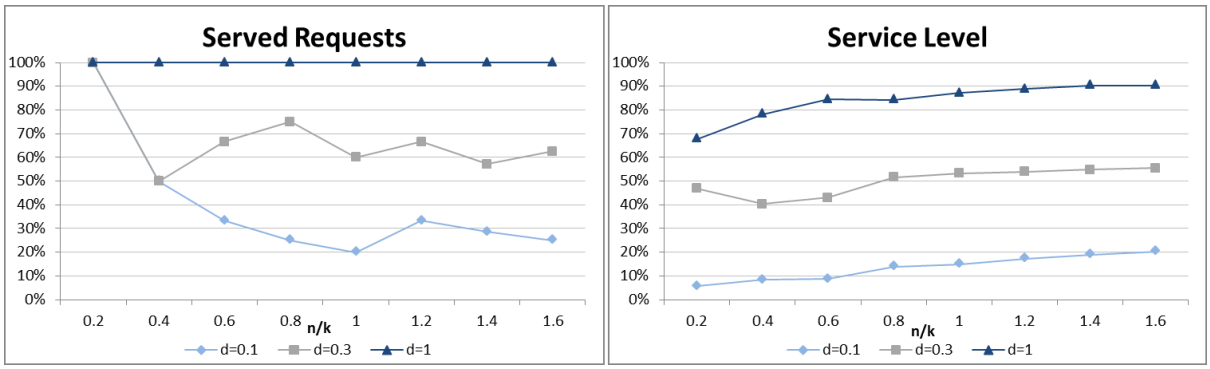


Figure 3.3: Service performance of SPTS for three different cases of d and assuming no other constraint

Consider now the constraint for the vehicle paratransit capacity Q. Figure 3.4 presents the proportion of requests served and the service level for $d = 1$. In this case also the x axis is the number of paratransit requests per trip (n/k). According to Figure 3.4, low values of Q decrease the service level provided by the system, especially with increasing n/k. However, even with only two reserved seats ($Q = 2$). Note that the patterns observed in this Figure are due to the specific restrictions imposed on the system (e.g. not possible to serve all requests).

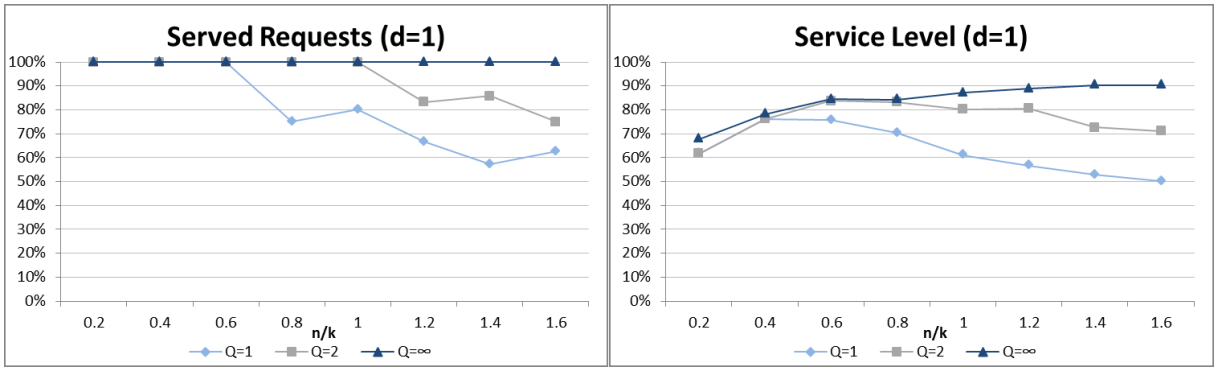


Figure 3.4: Service performance of SPTS for three different cases of Q, assuming that $d = 1$ and no other constraint

For the effect of the time window constraint, consider Figure 3.5, for which $Q = 2$ and $d = 1$. This Figure illustrates the efficiency of the system for the following three cases of time window width (where f is the headway between sequential bus trips):

- **Tight**: $\xi = b_j - a_j = 1 \cdot f$

- **Loose**: $\xi = b_j - a_j = 3 \cdot f$

- ∞: No time windows constraints applied

In this case, the effect of time windows does not appear to be significant, since constraints with respect to d and Q have already affected the performance of the system.

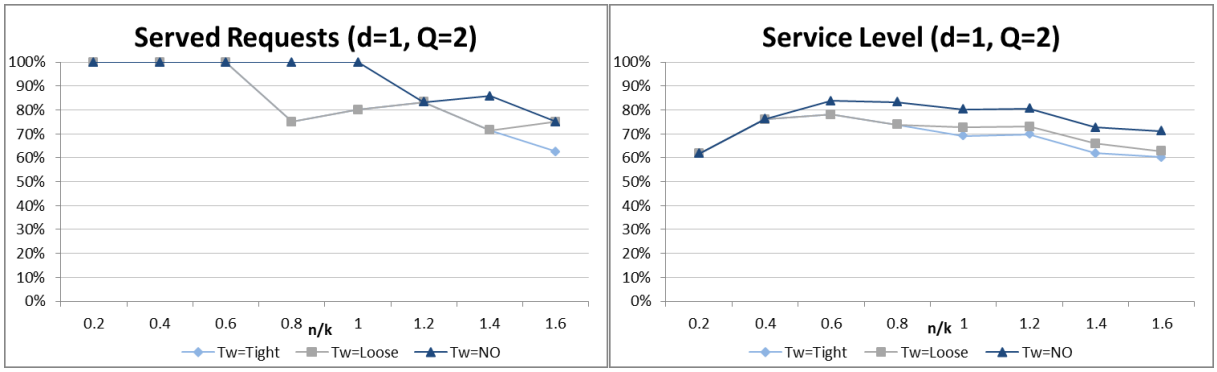


Figure 3.5: Service performance of SPTS, for three different cases of time window width, assuming that $d = 1$ and $Q = 2$

Figure 3.6 presents the interaction between d and Q when no time windows are applied for the paratransit requests, and for $n/k = 2$. It seems that for lower values of d the capacity of the vehicle is not a significant factor; that is d imposes a dominant constraint on system performance, and there is not strong interaction between d and Q. As d increases (up to a point) this interaction becomes stronger, and Q affects substantially SPTS performance. Note that for $Q = 2$ and $Q = 3$ the performance is similar, as expected, under the ratio $n/k = 2$. Thus, if the designer selects to use low d values, in order to limit deviations from the nominal route, increasing Q may not improve system performance.

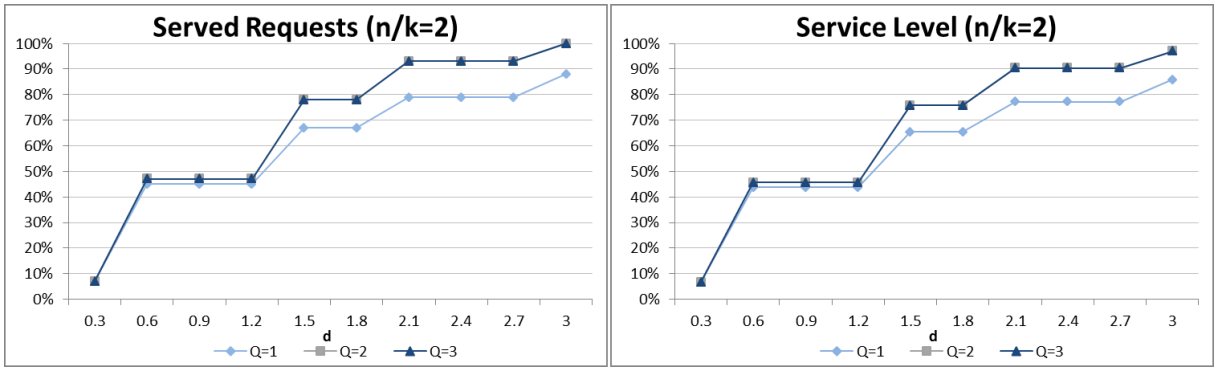


Figure 3.6: Service performance of SPTS as *d* increases, for three different cases of Q, assuming that $n/k = 2$ and $\xi = \infty$

Figure 3.7 presents the interaction between d and ξ assuming that $Q = 2$ and $n/k = 2$. In this case, the interaction is strong for the first two levels of ξ. As a result, performance improves significantly as ξ increases. Thus, even for low values of d, the system designer may improve performance by increasing ξ.

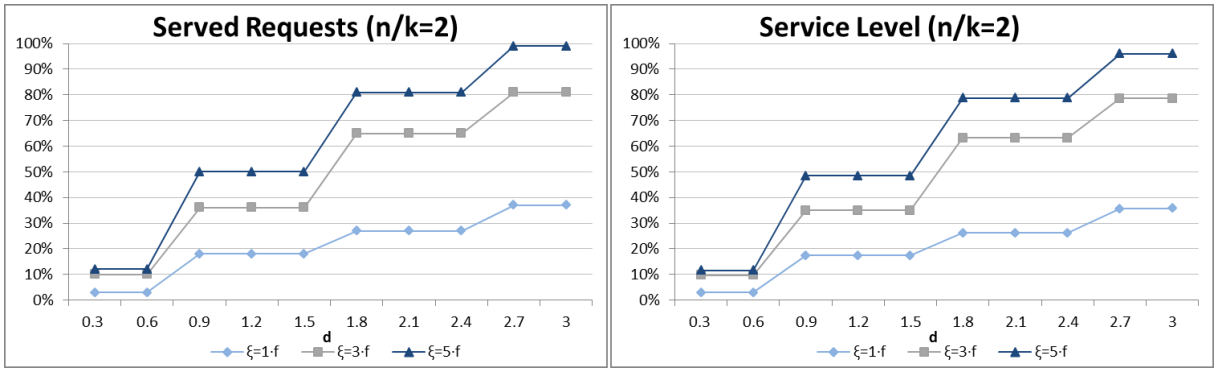


Figure 3.7: SPTS performance as *d* increases, for three different cases of ξ, assuming that $n/k = 2$ and $Q = 2$

Figure 3.8 presents the interactions between Q and ξ. It shows the performance of the system under four combinations of Q and ξ for $n/k = 2$. For low values of d the interaction is not strong; it increases, however, as d increases. As before, in the range of higher d values, the system performance is more sensitive to ξ than Q.

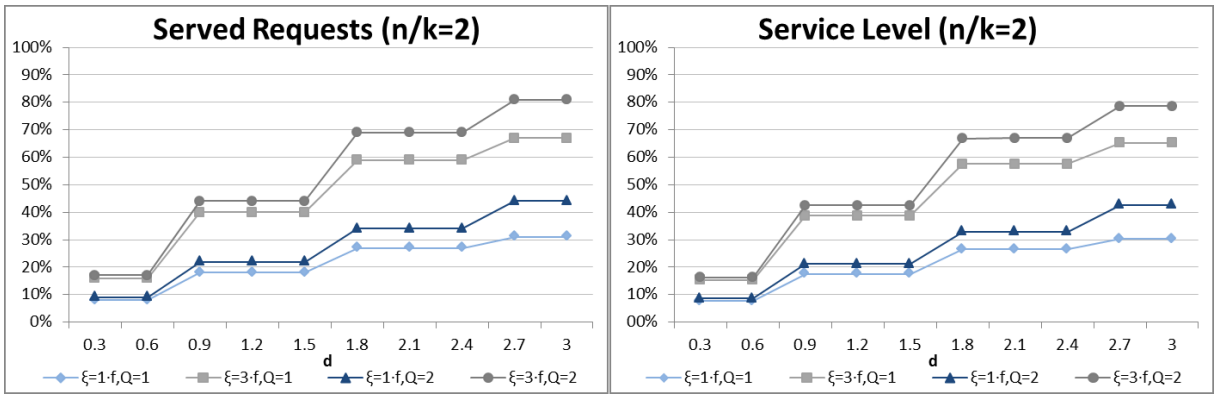


Figure 3.8: SPTS performance as *d* increases, for combinations of ξ and Q, assuming that $n/k = 2$

Overall:

- The service level is very sensitive to the allowable delay d
- Narrow time windows limit the service level significantly
- For lower values of d the capacity Q of the vehicle is not a significant factor. As d increases, up to a point, the interaction between d, Q becomes stronger, and Q now affects system performance
- The interaction between d, ξ is strong. Thus, even for low values of d, the system designer may improve performance by increasing ξ
- For low values of d the interaction between d, Q is not strong; it increases, however, as d increases.

3.4.3 Computational study III: SPTS in various mobility environments

In this Section we investigate the effects of the mobility environment (as characterized by parameters m, k, ρ, u and f) on system performance. Specifically, we study three typical mobility environments: urban, suburban, rural. In all cases we consider a planning horizon of eight hours.

Table 3.3: Parameters settings per mobility environment

Parameters		Environment		
Description	**Symbol**	**Urban**	**Sub-urban**	**Rural**
Average distance between sequential regular bus stops	ρ (m)	500	2000	10000
Headway between bus trips	f (mins)	10	30	60
Total number of bus trips in the planning horizon period	k	48	16	8
Average vehicle speed	u (km/h)	25	30	50
Number of regular bus stops	m	20	20	5
Time window width of the paratransit requests	ξ (mins)	60	60	60
Distribution of pick-up locations of paratransit requests w.r.t. the regular bus stops		Uniform in ellipse, Uniform along route	Uniform in ellipse, Higher at the first and third quarters of the route	Uniform in circle, Uniform along route

Parameters		Environment		
Description	**Symbol**	**Urban**	**Sub-urban**	**Rural**
Distribution of drop-off locations of paratransit requests w.r.t. the regular bus stops	-	Same as above	Uniform in ellipse, Higher at the second and fourth quarter of the route	Same as above
Ellipse major axis	ω (m)	500	1200	-
Circle radius	$\bar{\omega}$ (m)	-	-	4000

In an urban environment the distance between bus stops is low, the service is frequent (i.e. f <15 min), the average speed is low, and the number of regular bus stops is high (m>20). These characteristics change significantly in a rural environment. In this case the service frequency is low (i.e. $f > 30$ min), the average speed is higher and the number of the regular bus stops is low ($m < 10$). The parameter values of the suburban environment fall within the range spanned by the previous two environments. The key characteristics of the three environments are given in Table 3.3 and discussed further in Appendix A.

Table 3.4 uses the parameters of Table 3.3 to determine the values of the SPTS design parameters when operating in the above mobility environments. Figure 3.9 presents the performance of the system for $n = 30$ and $Q = 1$ for the three cases under consideration.

In the urban environment case, adequate service level is achieved for values of $d > 0.6$. This indicates that SPTS is applicable in urban cases, in which limited delay is expected, provided there is adequate capacity to serve the paratransit requests (in this example $Q/(n/k) = 1.6$), and wide time windows for such requests (in this example $\xi/f = 6$).

Table 3.4: Key SPTS characteristics per application environment

Key SPTS Characteristics		Environment		
Description	**Characteristic**	**Urban**	**Sub-urban**	**Rural**
Requests per trip	n/k	0.63	1.88	3.75
Available seats per request per trip	$Q/(n/k)$	1.60	0.53	0.27
Possible trips within the time window of the request	ξ/f	6	2	1

For the suburban environment case, the system displays good service levels for even lower values of d (i.e. $d > 0.3$). This seems to be due to the value of the ratio ω/ρ, which is lower

than that of the urban case. There is no significant improvement to the performance for values of $d > 0.4$. This may be due to the fact that, even if there is enough time to serve the requests, the latter compete for the same bus seat (in this case the performance limiting factor is $Q/(n/k) = 0.53$).

In the rural environment case, it seems that the system provides insufficient service level for the range of d examined. This again is attributed to the fact that the ratio of reserved seats per paratransit request per trip $Q/(n/k) \cong 0.27$ is low, limiting the performance. Another contributing factor may be the relatively tight time windows for such request with respect to the headway between bus trips $\xi/f = 1$.

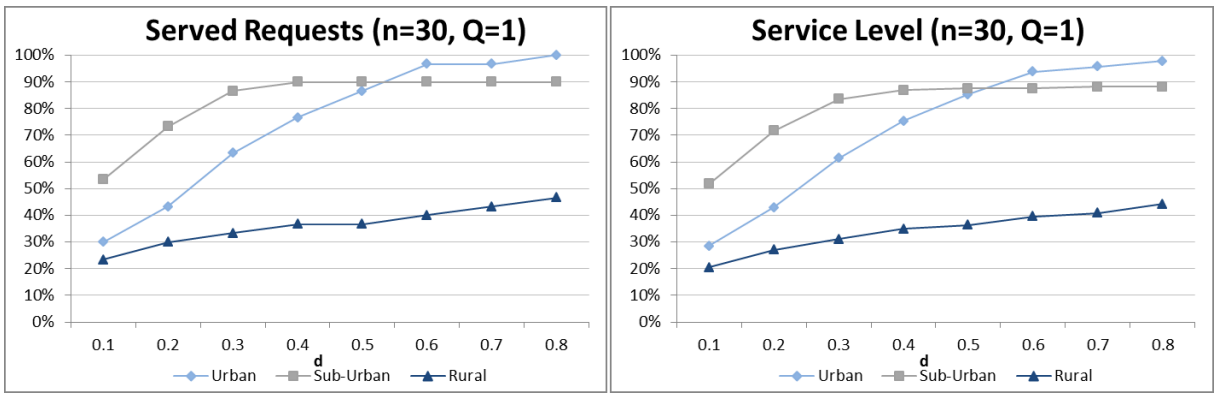


Figure 3.9: Service level for $n = 30$ and $Q = 1$ for the three different environments

Figure 3.10 illustrates three additional cases for the rural environment in which: (a) the value of $Q/(n/k)$ is doubled, (b) the value of ξ/f is doubled, and (c) both values are doubled. In all cases, the system performance is improved as expected. The improvement resulting from enhancing the paratransit capacity is more pronounced than the improvement resulting from relaxing the time windows of the paratransit requests. The system with enhanced capacity and relaxed time windows is clearly the best performing, as expected.

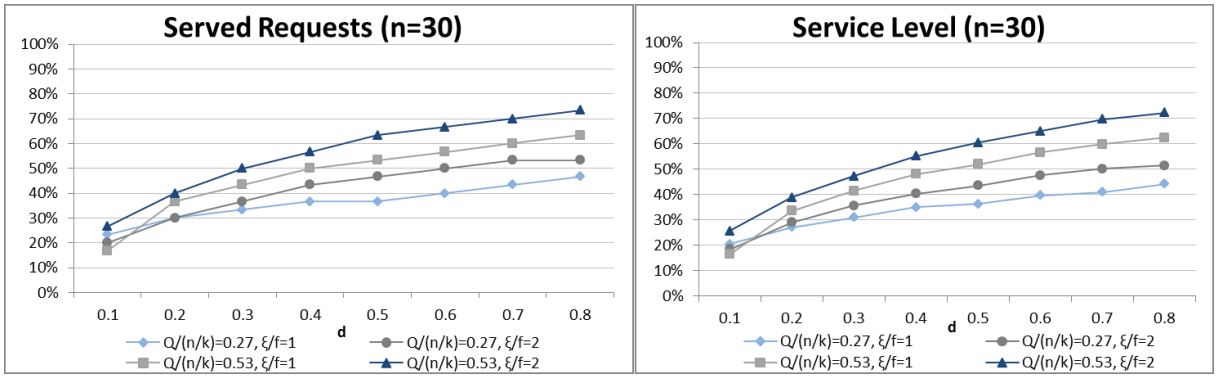


Figure 3.10: Service level for $n = 30$ and for different cases of the rural environment

3.4.4 System design considerations

Guided by the results of the above study, one may consider a systematic procedure for designing a scheduled paratransit transport system. In the first step of this procedure, the service headway f is determined in a straightforward manner. If the demand for non-paratransit passengers is h passengers per hour per direction and the vehicle stock's capacity is C passengers, then the number of trips per hour per direction is $k = h/C\ \mathrm{h}^{-1}$, or the headway is $f = 1/k$ h. Considering the ratio w of passengers that need assistance, then $n = h \cdot w$ is the hourly demand for paratransit passengers per direction, and n/k are the paratransit passengers per trip.

Subsequently, based on the desired service level and using an analysis similar to the one in subsection 3.4.3, the basic SPTS parameters could be defined; that is, (a) the maximum allowable delay of the nominal route d, (b) the number of available seats $Q/(n/k)$ per paratransit request per trip, and (c) the service level in terms of the width of the time window ξ/f over the service headway.

For example, Figure 3.11 illustrates the selection of the lower value of parameter d in order to serve more than 75% of the paratransit requests in a case within an urban environment with $h = 500\ \mathrm{pax} \cdot \mathrm{h}^{-1}$, $C = 50$ pax, $k = 10$ and $n/k = 2$. The four curves are related to four combinations of (Q, ξ) values. As expected, in all pair cases the performance of the system improves as parameter d increases. An interesting parameter setting is $Q = 2$, $\xi = 6 \cdot f$ and $d = 0.4$, which achieves the lowest disruption to the nominal route. Under the same paratransit capacity $Q = 2$, if one improves the waiting times for the paratransit clients to $\xi = 2 \cdot f$, then d should be increased to $d = 0.6$ to provide similar level of served requests.

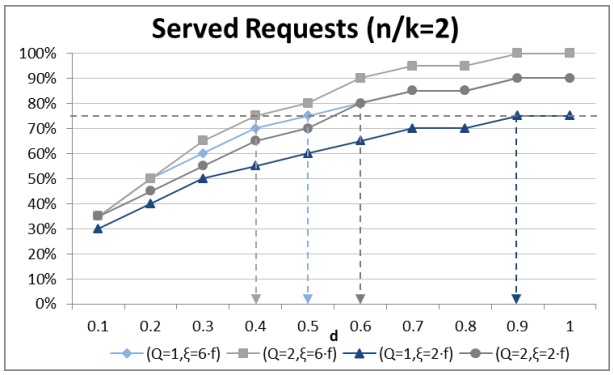


Figure 3.11: Ratio of served requests provided by SPTS for different parameters combinations assuming that $n/k = 2$

Chapter 4. Scheduled paratransit transport enhanced by accessible taxis (SPTS-T)

In Chapter 3 we studied the potential of public bus systems to provide paratransit, door-to-door transport services by deviating from their nominal public transport routes. Here we study an interesting enhancement of such SPTS services: In order to satisfy the total door-to-door paratransit demand, we supplement SPTS by accessible taxi (or private hire) services. In this manner, given sufficient number of accessible taxis, all paratransit requests may be satisfied, and the problem objective now becomes: minimize the cost of the combined public bus and taxi system, while offering a good level of service. We use the term SPTS-T to designate this enhanced system (T standing for Taxis).

In SPTS-T, similarly to SPTS, a paratransit request is booked prior to the start of the service period. Each paratransit request specifies the desired pick-up and drop-off locations and (a) the latest time of arrival at the destination in case of an outbound trip (i.e. from home to a certain destination), or (b) the earliest time of pick-up in case of an inbound trip (i.e. from the destination to home). For simplicity, and without loss of generality, in the rest of the Chapter we will refer to the case of outbound trips. All paratransit requests must be served either by the scheduled bus service, through an appropriate deviation from its nominal route, or by the accessible taxis. The same route restrictions as in SPTS apply in this case as well, i.e. (a) predefined regular bus stops must be visited once in a prescribed sequence, (b) the route is served by multiple bus trips that are offset in time (nominally by the headway of the service), (c) the earliest departure time from a regular bus stop is equal to the minimum bus travel time from the start of the route to the corresponding bus stop, (d) the latest bus departure time from each regular bus stop is set *a priori*, and (e) each bus in the system has a certain capacity for paratransit clients. Conditions (c) and (d) provide full transparency to the clients of the regular service, so that they are made aware of the time window within which the particular bus trip will serve their bus stop.

For the taxi (or private hire) vehicles we assume that: (a) a sufficient number of vehicles are available to serve the balance of paratransit requests, (b) all vehicles initiate from the same initial location (taxi rank) and return to same location, and (c) each vehicle has a certain capacity for paratransit clients (e.g. one or two wheel chairs). Note that assumption (a) is quite realistic in cities with significant fleets of accessible taxis.

In terms of costs, we consider that the cost of taxi services is related to the total distance travelled. For the buses, we only consider the additional distance travelled when the bus deviates from its nominal route. The cost of the nominal route is accounted as cost of the regular public transport service (scheduled service). While costing the two services, we also take into account (a) the differences between bus and taxi operational costs, and (b) the dissatisfaction of the passengers of the regular service caused by the route deviation. The latter may lead to some loss of revenue, which, in this case, may be modeled as additional cost.

An important issue in SPTS-T is the quality of service provided to the paratransit clients. The quality of service should be equivalent across the two modes, since paratransit clients book for SPTS-T services, without specifying the mode involved. Thus, we propose to regulate the service quality for the paratransit clients by setting appropriate time windows per mode as

presented in the Table 4.1 for both outbound (home to destination), or inbound (destination to home) trips. The text in parentheses refers to inbound trips.

Table 4.1: Time window settings per transportation mode and per trip type for outbound (inbound) trips

	Transportation mode	
Service parameter	**Bus**	**Taxi**
Latest time of arrival (earliest time of departure)	The client defines the latest time of arrival to his/her destination (the earliest time of departure from his/her origin).	The client defines the latest time of arrival to his/her destination (the earliest time of departure from his/her origin).
Earliest time of arrival (latest time of departure)	The earliest time of arrival to the destination is set to be the latest time of arrival defined by the client minus a predefined time threshold. (The latest time of departure from his/her origin is set to be the earliest time of departure defined by the client plus a predefined threshold).	The earliest time of arrival to the client's destination (the latest time of departure from his/her origin location) is defined by a predefined allowance (slack).
Ride time	The paratransit client enjoys the same service level quality in terms of ride time as the clients of the regular bus service.	The ride time quality is defined by the size of the time slack.

Based on the above, the time windows per mode in SPTS-T regulate the waiting and ride times for both client types. Table 4.2 presents a synopsis of how the quality of service is addressed in this setting.

Table 4.2: Quality of service per client type of SPTS-T (for the time window widths, please refer to Table 4.1)

Client Type	Maximum waiting time	Maximum excess ride time[1]
Regular bus service clients	The width of the time window at the bus stops	The width of the time window at the bus stops
Paratransit bus service clients	The width of the time window at the destination for outbound requests (origin for inbound requests)	The width of the time window at the bus stops
Paratransit taxi service clients	The size of the predefined time slack	The size of the predefined time slack

[1]In addition to the travel time required for a direct ride from the desired origin to the desired destination of the request

The remainder of this Chapter is structured as follows: Section 4.1 presents the mixed integer-linear formulation for the SPTS-T problem, which is a combination of the Dial-A-Ride problem and the Demand Responsive Bus Routing Problem, and hereafter we designate this combined problem as DRBRP-T. Section 4.2 discusses the proposed branch and price

approach to solve DRBRP-T to optimality, while in Section 4.3 we study the effects of realistic service parameters on the total cost of the system.

4.1 Mathematical formulation of DRBRP-T

In order to formulate the SPTS-T problem we enhance the notation already defined in Chapter 3. Specifically, we consider that:

- $G(N, A)$ is a directed graph where N is the set of all nodes related to the problem, and A is the set of arcs that connects the nodes
- B is the set of nodes related to the bus stops of the network; it comprises subsets B_k of nodes, each including all bus stops per bus trip $k \in K$, i.e. $B = \bigcup_{k \in K} B_k$, where K is the set of bus trips. Furthermore, similarly to Chapter 3, let $S \subseteq B$ contain the first bus stop of each bus trip and $E \subseteq B$ contain the last bus stop of each bus trip
- $C = C^+ \cup C^-$ is the set of nodes related to the paratransit requests, where C^+ contains the pick-up nodes of the paratransit requests and C^- contains the related delivery nodes. Moreover, we consider n be the number of paratransit requests. Note that for a paratransit request with pick-up node $i \in C^+$ the respective delivery node is $i + n \in C^-$.

For DRBRP-T we introduce set V to be the set of taxis, and for modeling purposes we assume one taxi to be available per paratransit request. Additionally, new nodes are introduced with respect to the starting and returning positions of the taxi vehicles, $\bar{S} = \{\bar{S}_k | k \in V\}$ and $\bar{E} = \{\bar{E}_k | k \in V\}$, respectively. Thus, set N is now defined as the union of all the aforementioned node sets, i.e. $N = B \cup C \cup \bar{S} \cup \bar{E}$.

Let the set of triplets $\bar{A}$ indicate which node may be served by which vehicle. Note that the first two elements of the triplets define arcs of A and the third is a bus trip of set K, or a taxi (trip) of set V. Set $\bar{A}$ is defined according to the problem needs as follows:

$$\bar{A} = \bigcup_{k \in K}\{(i,j,k) | i \in (S \cap B_k), j \in (\{i+1\} \cup C^+)\} \cup \bigcup_{k \in K}\{(i,j,k) | i \in B_k \backslash (E \cup S), j \in (\{i+1\} \cup C)\} \cup \bigcup_{k \in K}\{(i,j,k) | i \in C^+, j \in (C \cup B_k)\backslash(S \cup E \cup \{i\})\} \cup \bigcup_{k \in K}\{(i,j,k) | i \in C^-, j \in (C \cup B_k)\backslash(S \cup \{i-n, i\})\} \cup \bigcup_{k \in V}\{(i,j,k) | i \in \bar{S}_k, j \in (C^+ \cup \bar{E}_k\} \cup \bigcup_{k \in V}\{(i,j,k) | i \in C^+, j \in C \backslash \{i\}\} \cup \bigcup_{k \in V}\{(i,j,k) | i \in C^-, j \in (C \cup \bar{E}_k)\backslash\{i-n, i\}\}$$

Also consider that:

- $[a_i, b_i], i \in N \backslash \{\bar{S} \cup \bar{E}\}$ is the time interval within which each node (in $C \cup B$) must be served in case it is served by a bus, i.e. $k \in K$

- $[\hat{a}_i, \hat{b}_i], i \in C$ is the time interval within which each node i must be served in case it is served by taxi, i.e. $k \in V$.

Note that the following formulation is flexible and allows different time window settings depending on the modes and the type of nodes. Specifically, for the paratransit clients these settings depend on two factors: (a) the mode, bus or taxi, and (b) the direction of travel, inbound or outbound. For regular service clients these values are set according to the time windows defined in the bus schedule. How to define the time window settings in order to achieve the proposed service attributes presented in Tables 4.1 and 4.2 is further discussed in subsection 4.3.1. Furthermore consider that:

- $d_i, i \in N$ is the demand of paratransit passengers of each node. We assume that $d_i = 0$, for $i \in B \cup \bar{S} \cup \bar{E}$, $d_i = 1$ for $i \in C^+$, and $d_i = -1$ for $i \in C^-$. Note that if a paratransit client waits (or desires to be dropped-off) at a bus stop, then the respective nodes are replicated in C
- Q is the maximum number of paratransit passengers that any bus (or precisely bus trip) $k \in K$ may carry. As in Chapter 3, only the capacity of the bus in paratransit passengers is of concern
- $\bar{Q}$ is the maximum number of paratransit passengers that any taxi $k \in V$ may carry
- $T_{ij}, i, j \in N$ is the time required to travel directly from node i to node j considering that the travel times satisfy the triangle inequality.

The decision variables for DRBRP-T are as follows:

- $x_{ijk}, (i, j, k) \in \bar{A}$ is assigned the value 1 if the arc from node i to node j is traversed by a bus trip\taxi k, otherwise $x_{ijk} = 0$
- $s_i, i \in N$ is the time that each node i is served
- $q_i, i \in N$ is the number of paratransit passengers on board, either on the bus or on the taxi that serves node i, immediately after the node is served
- $w_i, i \in N \setminus (\bar{S} \cup \bar{E})$ is a binary variable that is used for linearization purposes (see below).

The objective of the problem is to minimize the relevant transportation cost. In the latter we include: (a) the total distance travelled by the taxis, and (b) the differential distance travelled by the bus trips in excess of the regular route; we multiply the latter by a scaling constant a. The following are noted with respect to this objective function:

- The cost functional relates only to the paratransit system costs, and, thus, excludes the distance of the regular bus routes.
- The scaling constant is used to account for the relatively higher cost of bus vs. taxi operation. It also accounts for the discomfort experienced by regular route passengers due to the deviations from the nominal route. As mentioned above, this discomfort may lead to loss of regular passengers, and thus revenue, which is accounted here as cost.

Thus, the objective function is defined as follows:

$$TC = min \sum_{\{(i,j,k)|k\in V,(i,j,k)\in \bar{A}\}} c_{ij}x_{ijk} + \sum_{\{(i,j,k)|k\in K,(i,j,k)\in \bar{A}\}} ac_{ij}x_{ijk} - a|K|F \tag{4.1}$$

where F is the cost of a bus to complete a single bus trip without deviating from its nominal route. Minimization of (4.1) is subject to (4.2) to (4.23), which are discussed in groups below according to their nature.

Path constraints

$$\sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \;\; i \in B_k \backslash E, k \in K \tag{4.2}$$

$$\sum_{\{i|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \;\; j \in (E \cap B_k), k \in K \tag{4.3}$$

$$\sum_{\{j|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \;\; i \in \bar{S}, k \in V \tag{4.4}$$

$$\sum_{\{i|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \;\; j \in \bar{E}, k \in V \tag{4.5}$$

$$\sum_{\{(j,k)|(i,j,k)\in \bar{A}\}} x_{ijk} = 1, \;\; i \in C^+ \tag{4.6}$$

$$\sum_{\{i|(i,h,k)\in \bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k)\in \bar{A}\}} x_{hjk} = 0, \;\; h \in B_k \backslash (S \cup E), k \in K \tag{4.7}$$

$$\sum_{\{i|(i,h,k)\in \bar{A}\}} x_{ihk} - \sum_{\{j|(h,j,k)\in \bar{A}\}} x_{hjk} = 0, \;\; h \in C, k \in K \cup V \tag{4.8}$$

$$\sum_{\{i|(i,h,k)\in \bar{A}\}} x_{ihk} - \sum_{\{j|(j,h+n,k)\in \bar{A}\}} x_{j,h+n,k} = 0, \;\; h \in C^+, k \in K \cup V \tag{4.9}$$

Service sequence and times constraints

$$s_{i-1} + T_{i-1,i} \leq s_i, \;\; i \in B_k \backslash S, k \in K \tag{4.10}$$

$$s_i + T_{i,i+n} \leq s_{i+n}, \quad i \in C^+ \tag{4.11}$$

$$s_i + T_{ij} - M(1 - x_{ijk}) \leq s_j, \quad (i,j,k) \in \bar{A} \tag{4.12}$$

$$s_j \leq a_j + (b_j - a_j)w_j + M(1 - x_{ijk}), \quad k \in K, (i,j,k) \in \bar{A} \tag{4.13}$$

$$s_j \leq s_i + t_{ij} + a_j(1 - w_j) + M(1 - x_{ijk}), \quad k \in K, (i,j,k) \in \bar{A} \tag{4.14}$$

Service level constraints

$$\sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}\,\hat{a}_i + \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk}\,a_i \leq s_i, \quad i \in C \tag{4.15}$$

$$s_i \leq \sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}\,\hat{b}_i + \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk}\,b_i, \quad i \in C \tag{4.16}$$

$$a_i \leq s_i \leq b_i, \quad i \in B \tag{4.17}$$

Vehicle capacity constraints

$$q_i + d_j - M(1 - x_{ijk}) \leq q_j, \quad (i,j,k) \in \bar{A} \tag{4.18}$$

$$q_i \leq Q \sum_{\{(j,k)|k\in K,(i,j,k)\in\bar{A}\}} x_{ijk} + \bar{Q} \sum_{\{(j,k)|k\in V,(i,j,k)\in\bar{A}\}} x_{ijk}, \quad i \in N\backslash\{S \cup E \cup \bar{S} \cup \bar{E}\} \tag{4.19}$$

$$\sum_{i\in S\cup E\cup\bar{S}\cup\bar{E}} q_i = 0 \tag{4.20}$$

Range of decision variables

$$x_{ijk} \in \{0,1\}, \quad (i,j,k) \in \bar{A} \tag{4.21}$$

$$w_i \in \{0,1\}, \quad i \in N(\bar{S} \cup \bar{E}) \tag{4.22}$$

$$q_i \in \mathbb{N}^0, \quad i \in N \tag{4.23}$$

Regarding the path constraints: Constraints (4.2) and (4.3) ensure that all regular bus stops will be served once per bus trip: Constraints (4.4) and (4.5) define that all available taxi vehicles will depart from the starting location and will return to the destination location, respectively. Constraint (4.6) ensures that the paratransit pick-up nodes (and thus the respective paratransit requests) must be served once. Equations (4.7) and (4.8) ensure continuity for all nodes except those corresponding to the starting and ending regular bus stops of each bus trip, and the starting and ending nodes of the taxi vehicles. Equation (4.9) ensures that the corresponding pick-up and delivery nodes of a paratransit request will be served during the same bus trip or by the same taxi vehicle.

Regarding the service sequence and times: Constraint (4.10) ensures that regular bus stops will be served with the predefined sequence: Constraint (4.11) ensures that paratransit delivery nodes will be served after their related pick-up nodes. Constraint (4.12) defines the time flow, assuming that $M \gg 1$. Constraints (4.13) and (4.14) ensure that if a node is served during a bus trip, then the service will be performed at the earliest possible time, similarly to Constraints (3.11) and (3.12) of Chapter 3.

Regarding the service level: Constraints (4.15) and (4.16) ensure service level (time window) compliance for paratransit requests: Specifically, Constraints (4.15) specify that service is provided to a paratransit request at a time greater than or equal to the lower limit of the corresponding time window (i.e. $\hat{a}_i$ if it is served by a taxi or a_i if it is served by a bus). Similarly, Constraints (4.16) specify that this time is less than or equal to the upper limit of the time window ($\hat{b}_i$ if it is served by a taxi or b_i if it is served by a bus). Constraint (4.17) defines the time windows for the bus stop locations. Note that within this formulation in order to achieve the desired service level one should predefine appropriately the values of the time windows for all locations (both pick-up $i \in C^+$ and drop-off $i + n \in C^-$). The latter may require the deactivation of the time windows for certain locations. For instance in case of an outbound trip the time windows of the origin node may not be used. As mentioned above, the definition of these values in the context of the service described by Tables 4.1 and 4.2 is further discussed in subsection 4.3.1.

Regarding the vehicle capacity: Constraint (4.18) defines the change in the paratransit vehicle load: Constraint (4.19) ensures that the paratransit vehicle load is greater or equal to zero and less than or equal to Q or $\bar{Q}$, depending if the node is served by a bus or a taxi, respectively. Constraint (4.20) defines that every vehicle starts and ends empty.

Finally, regarding to the range of decision variables: Constraints (4.21) and (4.22) define the binary nature of decision variables x_{ijk} and w_i, respectively, and Constraint (4.23) ensures that q_i is a positive integer or zero.

4.2 A branch and price approach for the DRBRP-T problem

As in Chapter 3, we introduce a Branch and Price framework to solve DRBRP-T to optimality. We will also leverage this framework in Section 4.3 in order to assess the implication of various design factors to the total cost of an SPTS-T system. In order to implement the Branch and Price framework for DRBRP-T, we first define a route-based

formulation for this problem, subsequently we describe how the respective column generation technique is applied to it, and finally we describe the overall Branch and Bound scheme.

4.2.1 The master problem for DRBRP-T

Let $\delta = \delta_1 \cup \delta_2 \cup \dots \cup \delta_{|K|} \cup \delta_{|K|+1}$ be the set of all possible feasible routes including both the routes of the bus trips, and those of the taxi vehicles. In this notation, set $\delta_k, k = \{1,2,\dots,|K|\}$ contains all feasible routes of bus trip k, which satisfy constraints (4.2) - (4.23) by setting $K = \{k\}$ and $V = \emptyset$. Each route in $\delta_{|K|+1}$ is a feasible route of a taxi vehicle, which must satisfy constraints (4.2)-(4.23) by setting $|V| = 1$ and $K = \emptyset$.

a_{ij} is a binary coefficient that takes the value 1 if paratransit request $i \in C^+$ is served by route $j \in \delta$ or 0 otherwise. Let also y_j be the binary variable which is assigned the value 1 if route j is used, and 0 if not. In terms of cost:

- For each route $j \in \delta_k, k = \{1,2,\dots,|K|\}$, let c_j be the differential cost of trip j multiplied by the scaling constant a. The differential cost (distance) is equal to the difference between the total cost of bus trip j minus F, the cost of a bus to complete a single bus trip without deviating from its nominal route
- For each route $j \in \delta_{|K|+1}$, c_j is the total distance travelled by the taxi vehicle traversing route j.

Then the problem may be formulated as a set partitioning problem. The objective function is the following:

$$TC = min \sum_{j \in \delta} c_j y_j \tag{4.24}$$

Subject to:

$$\sum_{j \in \delta} a_{ij} y_j = 1,\ i \in C^+ \tag{4.25}$$

$$\sum_{j \in \delta_k} y_j = 1,\ k = 1,\dots,|K| \tag{4.26}$$

$$y_j \in \{0,1\},\ j \in \delta \tag{4.27}$$

Constraint (4.25) ensures that all paratransit requests (pick-up nodes) may be served once. Constraint (4.26) ensures that one route must be used per bus trip; thus all regular bus stops will be served once during each bus trip.

4.2.2 Column generation for DRBRP-T

In column generation, the relaxed Linear Programming version of the Master Problem (MP) for the DRBRP-T is solved to optimality by sequentially solving restricted versions of it (Desaulniers *et al.*, 2005). The linear programming version of the problem (LMP) is achieved by replacing Constraint (4.27) with $0 \leq y_j \leq 1, j \in \delta$.

The solution of MP is achieved by iteratively solving a restricted version (RLMP) of LMP that considers a subset of variables (columns) $\delta' \subseteq \delta$, and by identifying new columns of δ to be added to δ'. The latter procedure is performed until no new columns can be identified. The identification of new columns is accomplished by the so-called sub-problem (SP). The SP is based on the dual information of the last solved RLMP and aims to identify columns (which in this case represent routes) with negative reduced cost to provide to the RLMP.

In the case under consideration one may define $|K|$+1 independent sub-problems (SP), one for each bus trip $k \in K$ and one that deals with all taxi vehicles $k \in V$, since we consider that the same restrictions apply for all taxi vehicles. Each SP will produce new routes to be added to the related route sets δ_k or $\delta_{|K|+1}$. For the sub-problems with respect to the bus trips $k \in K$, the new routes j are those satisfying the following condition:

$$c_j - \sum_{i \in C^+} a_{ij}\pi_i - \bar{\pi}_k < 0 \tag{4.28}$$

where π_i are the dual prices related to Constraint set (4.25), and $\bar{\pi}_k$ are the dual prices related to Constraints (4.26). As aforementioned, these routes should satisfy constraints (4.2) - (4.23) by setting $K = \{k\}$, and $V = \emptyset$.

These sub-problems are solved by the labeling algorithm presented in Chapter 3 for the SPTS service, excluding all information related to the ride time of each paratransit client, i.e. rt, (3.37) and (3.48).

For the taxi-related sub-problem $k \in V$, the routes j to be provided to RLMP satisfy condition (4.29) below, and constraints (4.2) - (4.23) by setting $K = \emptyset$, and $|V| = 1$.

$$c_j - \sum_{i \in C^+} a_{ij}\pi_i < 0 \tag{4.29}$$

This sub-problem is addressed as an Elementary Shortest Path Problem with Time Windows, Capacity, and Pick-up and Delivery, and it is solved by the labeling algorithm introduced in Ropke and Cordeau (2009).

4.2.3 Branch and bound

Given that column generation does not guarantee the integrality of the problem solution, in our approach we implement a typical branch and bound scheme similar to the one in Chapter 3. That is, we branch on flow variables based on the value of x_{ijk}. New branches will include constraints $x_{ijk} = 1$ and $x_{ijk} = 0$, respectively. The variable x_{ijk} that is closest to 0.5 is selected for branching. Interested readers may refer to Desaulniers *et al*. (2005) for alternative branching strategies.

4.3 Tests and results

In this Section, we first study the effect of the cost scaling factor a on SPTS-T performance. Note that a depends on the particular operational environment of SPTS-T as well as the related bus and taxi costs (see below). Subsequently, for a selected value of a, we study how key design parameters of the system affect overall performance. These parameters are selected by the system designer, based on service level considerations, and are first presented and discussed in subsection 4.3.1 below.

4.3.1 System design considerations

During the design of an SPTS-T system, the designer aims to achieve high quality transportation services while minimizing the respective costs. This goal should be achieved considering the particular environment of application. The latter is characterized by the following parameters:

- the relationship between bus and taxi costs designated by the cost scaling factor a
- the level of paratransit demand characterized by the metric $n/(k \cdot Q)$ defined in Chapter 3, where n is the total number of paratransit requests in the time period under consideration, k is the number of bus trips in the same period (defined with respect to f the headway between bus trips), and Q is the bus capacity per trip for paratransit passengers. Metric $n/(k \cdot Q)$ indicates how many requests are competing for the reserved seats per trip, i.e. high values of this metric correspond to higher demand.

As mentioned in the introduction of this Chapter and in Section 4.1, an important issue for the designer is to define the service level to be provided by SPTS-T. This is done by designating

appropriately the time windows at the various nodes. For these time windows we consider the following:

- For every bus stop $i \in B$ of a certain bus trip, as in Chapter 3, the opening of the time window a_i, is set to the minimum travel time of the bus between the start of the trip and the bus stop under consideration. The width $b_i - a_i$ of the time window is defined by setting $b_i = a_i + d \cdot r/u$, where d is a design parameter, r is the average distance between sequential regular bus stops and u is the average speed of a bus. Thus, d regulates the maximum time a passenger of the regular route may wait at a bus stop, in order not to miss the particular bus trip. Note that the time offset $d \cdot r/u$ is cumulative with the progression of the bus trip; thus, if a bus would deviate from its route at the start of its trip, it would result in less time allowed time for deviation for the latter part of the trip.
- For a paratransit client served by bus, and for an outbound trip, the time window is considered at the destination $i + n \in C^-$ (with $i \in C^+$). In this case b_{i+n} is the latest time of arrival designated by the client and $a_{i+n} = b_{i+n} - \xi$, where ξ is the width of the time window, which in our approach is defined w.r.t to f, the headway between bus trips. Thus, a paratransit client may be served by ξ/f different bus trips and dropped-off up to ξ earlier from the desired time at his/her destination. For inbound trips, the time window is considered at the origin node, and the opening of the time window is set by the client, while the time window width is identical to the outbound trip case. Note that a_i and b_i may be deactivated by setting $a_i = 0$ and $b_i = b_{i+n}$.
- For a paratransit client served by a taxi, the width of the time window concerns the entire trip, and, thus, by setting the latest time of arrival at the destination of an outbound trip (earliest time of departure at the origin of the inbound trip), the time windows are set with respect to both the origin and destination locations $i \in C^+$and $i + n \in C^-$, respectively. In this study we assume that $\hat{b}_{i+n} = b_{i+n}, i + n \in C^-$, i.e. the latest time of arrival for an outbound trip is the same for both modes (since the client does not specify the mode). In this case $\hat{a}_i = \hat{b}_{i+n} - (T_{i,i+n} + d \cdot \xi)$, where $T_{i,i+n}$ the time required to travel directly from node i to node j. This provides to taxi paratransit clients a similar service level in terms of combined ride and waiting time to the one provided by the bus system to these clients. Furthermore, $\hat{a}_{i+n}$ and $\hat{b}_i$ may be deactivated by setting $\hat{a}_{i+n} = \hat{a}_i$ and $\hat{b}_i = \hat{b}_{i+n}$.

A final issue related to the service is whether the taxis should carry more than one client on board, i.e. ride sharing. This is regulated by setting the value of the taxi capacity $\bar{Q}$.

Based on the above discussion, the design parameters to be set by the designer of SPTS-T are the following:

- parameter d that sets: (a) the maximum allowable time deviation of a bus trip from the nominal route, (b) the maximum time a passenger of the regular route may wait at a bus stop; it also affects the maximum ride time of a paratransit client to be served by a taxi
- parameter ξ that is the width of the time window for a paratransit request served by bus; it also affects the maximum ride time of paratransit requests served by taxis
- the taxi capacity $\bar{Q}$ for paratransit passengers
- the bus capacity Q for paratransit passengers.

The key design questions to be investigated are the following:

- how does the cost scaling factor (a) affect the utilization of the two available modes?
- what is the appropriate service level to be provided by the system?
- how does service level affect system costs?

4.3.2 Problem generation

To generate a wide range of test problems for the experimental investigation, we used the random problem generator presented in Appendix A and the parameters of Table 4.3. These parameters relate to an urban case environment, in which SPTS-T will operate.

Table 4.3: Parameters settings used for problem generation (urban environment)

Symbol	Description	Value
ρ (m)	Average distance between sequential regular bus stops	500
f (mins)	Headway between bus trips	10
m	Number of regular bus stops	20
u (km/h)	Average vehicle speed	25
Pick-up and Drop-off demand generation	Pick-up and drop-off locations of the paratransit requests along the service route. The allowable regions in which the pick-up and drop-off locations lie are ellipses. Each ellipsis has as foci two sequential regular bus stops with major axis as below	-
ω (m)	Ellipse major axis	500

Also we assume that one seat is reserved for paratransit clients per bus trip ($Q = 1$).

The Branch and Price framework was implemented using Matlab R2013 on a PC equipped with a 2.00 GHz Intel Xeon E5-2620 and 32 GB of RAM. For the solution of the Master Problem we used the Gurobi Matlab interface (Gurobi Optimization Inc., 2014). In terms of the computational times required to perform the tests described in the following sections (a) for over 90% of the cases the optimal solution was obtain in under 10 sec, and (b) for the rest 10% of the cases the optimal solution was obtained in under 300 sec.

4.3.3 The effect of the cost scaling factor on bus vs. taxi utilization and system costs

Selecting an appropriate value for the cost scaling factor α of objective function (4.1) affects significantly system performance. The appropriate choice is influenced partially by the relationship of acquisition and operating costs between buses and taxis. This relationship is not uniform across systems, and depends on the capacity, typical load factor and costs of the bus public system available for SPTS services, as well as the capacity and costs of the accessible taxi (or private hire) services, etc. Scaling factor α is also influenced by possible loss of revenues following the reduced attractiveness of the regular service caused by the route deviations necessary to pick-up paratransit clients.

In the current Section we investigate how α affects system performance, so that the system designer is aware of the impact of her/his selection (or estimation) of α. Conversely, in a possible tender for taxi or private hire outsourcing, the public transport agency may set a goal for the desired system performance, and thus set the value of α. Based on this value, and the known bus costs, the goal of outsourcing costs for the related taxi services may be set in the response to the tender.

We examine the effect of α on system performance for three cases of progressively improving service level (and thus tightening the related constraints). The values of parameters ξ and d related to these cases are provided in Table 4.4. Furthermore, the Table presents the corresponding service levels per case.

All cases of Table 4.4 were studied for two values of taxi capacity $\bar{Q} \in \{1,2\}$. Note that $\bar{Q}$ affects the service level, since for $\bar{Q} = 1$ the client receives strictly individualized service, and the ride time is the minimum possible. For $\bar{Q} = 2$ clients may share a ride, extending the

ride time of one or both clients as necessary. Of course, for $\bar{Q} = 2$ fewer resources and lower costs are expected.

Table 4.4: Problem cases with progressively increasing service levels

Case	Client Type	Maximum waiting time	Maximum excess ride time[3]
$\xi = k \cdot f$ [2], $d = 3$	Regular bus service	$3 \cdot (r/u)$ [1]	$3 \cdot (r/u)$
	Paratransit bus service	$k \cdot f$	$3 \cdot (r/u)$
	Paratransit taxi service	$3\ k \cdot f$	$3\ k \cdot f$
$\xi = k \cdot f$, $d = 0.3$	Regular bus service	$0.3 \cdot (r/u)$	$0.3 \cdot (r/u)$
	Paratransit bus service	$k \cdot f$	$0.3 \cdot (r/u)$
	Paratransit taxi service	$0.3 \cdot k \cdot f$	$0.3 \cdot k \cdot f$
$\xi = 2 \cdot f$, $d = 0.3$	Regular bus service	$0.3 \cdot (r/u)$	$0.3 \cdot (r/u)$
	Paratransit bus service	$2 \cdot f$	$0.3 \cdot (r/u)$
	Paratransit taxi service	$0.6 \cdot f$	$0.6 \cdot f$

[1] r/u is the average travelling time between consecutive bus stops, [2] $k \cdot f$ is the bus service operation period (e.g. period of daily service); k or $k \cdot f$ are used in order to examine the unconstraint cases, [3] In addition to the minimum ride time

One hundred (100) problem instances were generated and solved for each of the three cases presented in Table 4.4, two values of $\bar{Q}$ and eleven values of a for a total number of $100 \times 3 \times 2 \times 11 = 6{,}600$ instances. The solutions were obtained using the Branch and Price method of Section 4.2. The results are presented in Fig. 4.1 to 4.3; the values in the related graphs are the averages of the 100 instances per case.

For case 1, Fig. 4.1a presents the percentage of paratransit clients transported by bus or taxi for increasing values of the cost factor a and for $\bar{Q} = 1$. As expected, as the bus system becomes more expensive, more paratransit clients are transported by taxi vehicles. Although, it seems that even for very large values of a more than 10% of the passengers are still being transported by buses. This indicates that the SPTS-T system could be applied even for cases where the deviation of a bus from its nominal route is very expensive, if these deviations are limited. The two curves meet for large values of a (approximately equal to 50) which indicates that in this SPTS-T case, serving paratransit request by taxis costs about 50 times more than serving the requests by allowing buses to deviate from their nominal route.

Figure 4.1b presents the travel cost for bus (differential cost), taxi, and total SPTS-T cost. It seems that for low values of a the bus cost significantly increases as a increases. This is attributed to the fact that for these values of a the bus mode is still preferable in terms of cost. Specifically, for $a < 20$ the rate at which the bus modal split decreases (Fig. 4.1a) is low,

while the rate at which the bus costs increase is high. After a maximum at $a = 20$, the bus mode is not preferable in terms of cost which is obvious in Fig 4.1a in which the rate at which the bus modal split decreases becomes higher, and in Fig 4.1b, in which the bus mode cost gradually decreases. For large values of a, buses serve only those paratransit clients that correspond to low differential distances.

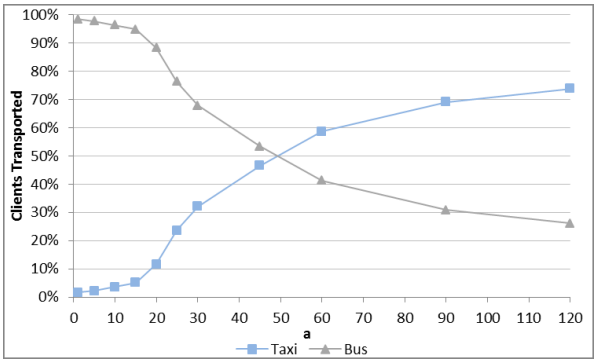


Figure 4.1a: Modal split of paratransit clients for increasing values of a and for ξ=high, d=high and $\bar{Q} = 1$

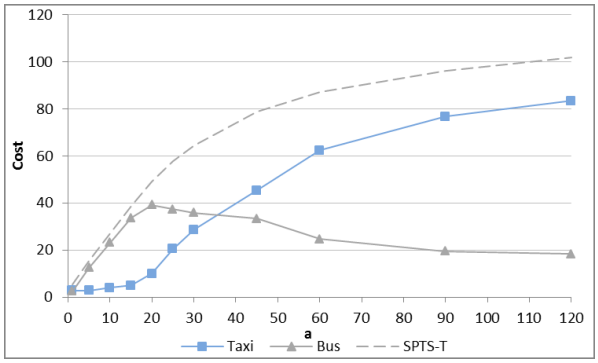


Figure 4.1b: Cost per mode for increasing values of a and for ξ=high, d=high and $\bar{Q} = 1$

Figure 4.2a presents the ratio of clients transported by bus for the three cases discussed above. It seems that for $a \leq 20$ the effects of the constraints are significant. Specifically, for $a = 1$: for $\xi = k \cdot f$, $d = 3$ (the most relaxed case 1) almost all paratransit clients are transported by bus (98%), while for $\xi = 2 \cdot f$, $d = 0.3$ (the most constrained case 3) the percentage of clients transported by bus drops to 78%. As expected, when the service level constraints are relaxed, the cost of the SPTS-T is lower throughout the range of the cost scaling factor, (see Fig. 4.2b). This indicates that for cases with expensive bus services (e.g. $a > 30$), the most relaxed case may be selected, offering lower total cost, without significantly compromising service levels.

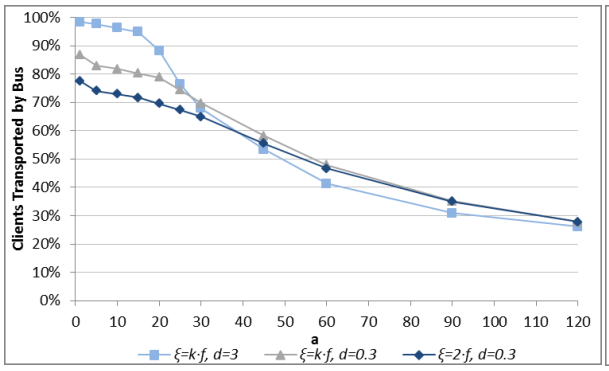


Figure 4.2a: Bus modal split for the three cases studied, $\bar{Q} = 1$, and for increasing values of a

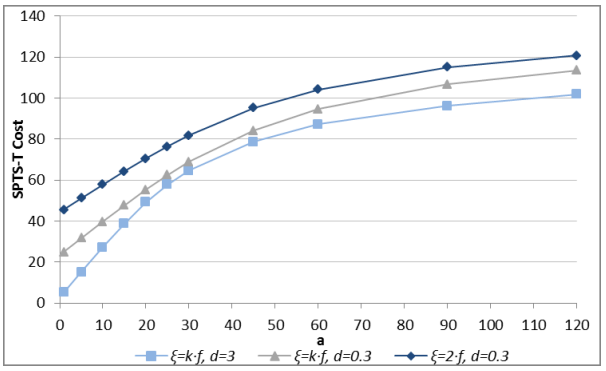


Figure 4.2b: Cost of the SPTS-T for and the three cases studied, $\bar{Q} = 1$, and for increasing values of a

Figures 4.3a and 4.3b present the variation of the bus modal split and the total cost as α increases for taxis with capacity $\bar{Q} = 2$. For more relaxed ride time constraints (case 1) and for higher values of a, taxis take advantage of the enlarged feasible space and transport more clients, lowering the total transportation cost. This is evident from the greater distances between the three curves if Fig. 4.3b, as compared to the case with $\bar{Q} = 1$. Note that in this case, clients may share the ride, which extends the ride time, but decreases the use of bus, and the distances travelled by taxis.

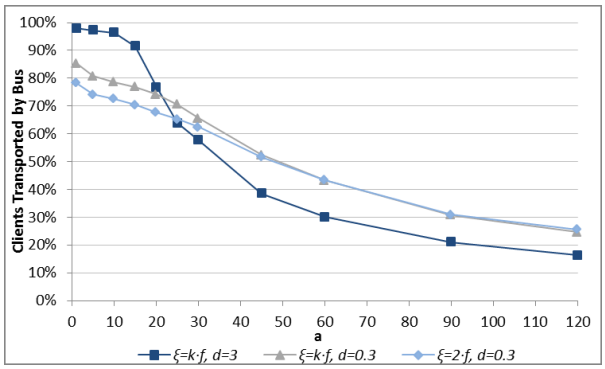


Figure 4.3a: Bus modal split for the three cases studied, $\bar{Q}$=2 and for increasing values of a

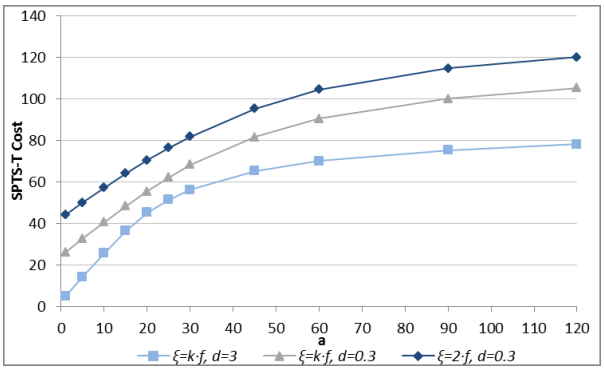


Figure 4.3b: Cost of SPTS-T for the three cases studied, $\bar{Q} = 2$, and for increasing values of a

4.3.4 Effect of service level parameters on costs

As discussed previously, an important decision concerning the design of the SPTS-T system concerns the selection of the service levels to be provided by the system. Obviously, this decision has cost implications. The relationships between the design parameters that define the service level and the system costs are explored below. For this study:

- the cost scaling factor a is set to 10. This value relates to practice considering that: (a) the operating expenses per vehicle revenue mile of a bus are about three times higher than those of taxis (National Transit Database, 2013), and (b) the latter ratio should be increased to include revenue losses resulting from the dissatisfaction of regular route passengers due to the bus route deviations
- the problems generated correspond to two levels of paratransit demand – high and low – characterized by the metric $n/(k \cdot Q)$ (see subsection 4.3.1) Specifically, for high demand we assumed that $n/(k \cdot Q) = 2.1$ and for low demand that $n/(k \cdot Q) = 0.75$.

Furthermore, in terms of the design parameters we examine different cases with respect to:

- $\bar{Q} \in \{1,2\}$, i.e. not allowing and allowing ride sharing, and

- $\xi \in \{3 \cdot f, k \cdot f\}$, i.e. relatively narrow time windows providing satisfying service level, and wider time windows (practically unconstrained) providing inferior service level.

As in the previous Section, one hundred (100) problem instances were generated for the eight (8) combinations resulting from the two demand levels ($n/(k \cdot Q)$) and the two levels each of the design parameters ($\bar{Q}$ and ξ), and for $d \in \{1,2,\ldots,10\}$. The total number of instances generated and solved is 2^3x10x100 = 8,000. Again, the solutions were obtained using the Branch and Price method of Section 4.2 and the results presented in Fig. 4.4 and 4.5 include the values of the averages of the 100 instances per problem case.

Figures 4.4a and 4.4b present the relationship between the modal split and the parameter d for high and low demand, respectively. Each Figure presents the combinations for $\bar{Q} \in \{1,2\}$ and $\xi \in \{3 \cdot f, k \cdot f\}$ for d values increasing from 0.1 to 1.

In terms of the variation of the modal split, in all cases as d increases from 0.1 to 0.4 the bus modal split increases significantly. This is attributed to the fact that as d increases, buses are given more time to deviate from their nominal route, and thus serve more paratransit requests. For d values from 0.5 to 1 it seems that the bus modal split reaches a plateau; this is attributed to the fact that the bus modal split may not increase further without increasing the system total cost. This is validated by Fig. 4.5 below.

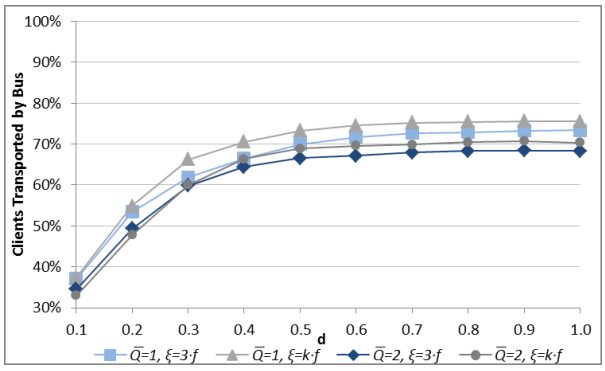


Figure 4.4a: Bus modal split for different cases of ξ and $\bar{Q}$ for increasing values of d and $a = 10$ and $n/(k \cdot Q) = 2.1$

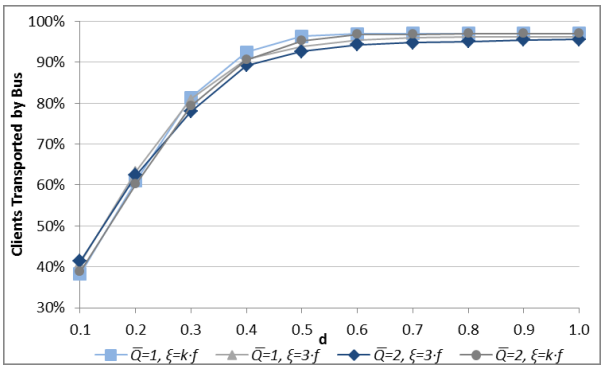


Figure 4.4b: Bus modal split for different cases of ξ and $\bar{Q}$ for increasing values of d and $a = 10$ and $n/(k \cdot Q) = 0.75$

Among the four cases, more clients are transported by taxis for the cases with $\bar{Q} = 2$, as expected. Furthermore, between the cases with the same value of $\bar{Q}$, more clients are transported by bus for those with $\xi = k \cdot f$. This indicates that for $a = 10$, the bus mode contributes more in the reduction of the total system cost. Focusing on the differences between the two demand cases (Fig. 4.4a vs. Fig. 4.4b), it seems that for low paratransit

demand, buses serve almost all paratransit clients and there are no significant differences between the cases of $\bar{Q} = 2$ and $\bar{Q} = 1$.

Figures 4.5a and 4.5b present the variation of the system costs with increasing values of d for the two levels of demand, respectively. For each demand level, the four combinations of parameters ξ and $\bar{Q}$ have been examined as before. As d increases the total cost declines, since a key service level constraint is relaxed, and thus the system may achieve improved cost. Furthermore, it seems that system cost improves significantly as d increases from 0.1 to 0.4, which corresponds to the range in which the bus modal split is significantly affected (see Fig. 4.4a and 4.4b). The system costs reach a plateau similarly to the bus modal split for $d \geq 0.5$.

Notably, for the $\bar{Q} = 2$, $\xi = k \cdot f$ case, with the most relaxed service related constraints, the lowest system cost is achieved for all cases as expected (refer to Fig 4.5a and 4.5b). Although according to Fig. 4.4a and 4.4b the total system cost is not correlated closely to the bus modal split; it seems that the bus modal split for the $\bar{Q} = 2$, $\xi = k \cdot f$ case is moderate compared to the other three cases, indicating that the lowest system cost is not achieved for the highest values of the bus modal split. This is evidence that SPTS-T may achieve better solutions in terms of system costs, comparing to a system that first assigns the maximum number of paratransit requests to the bus service (as in SPTS), and then assigns the remaining requests to taxis.

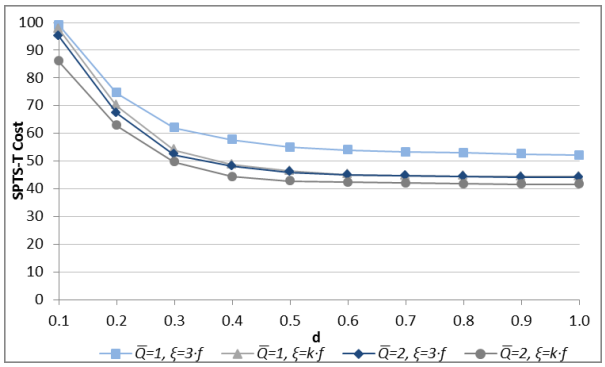


Figure 4.5a: Cost of SPTS-T for different cases of ξ and $\bar{Q}$ for increasing values of d, $a = 10$ and $n/(k \cdot Q) = 2.1$

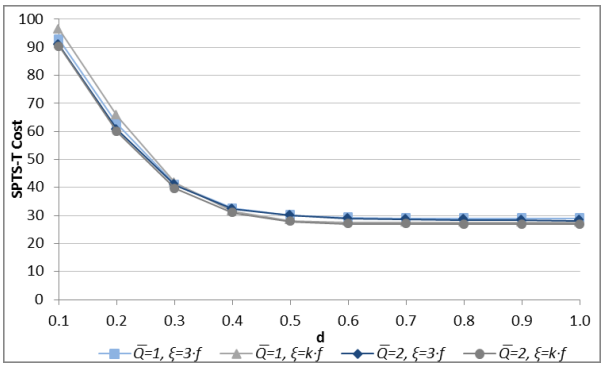


Figure 4.5b: Cost of SPTS-T for different cases of ξ and $\bar{Q}$ for increasing values of d, $a = 10$ and $n/(k \cdot Q) = 0.75$

Comparing the two demand cases (Fig. 4.5a vs. Fig. 4.5b), it seems that for low paratransit demand, in which buses serve almost all paratransit clients, there are no significant differences in terms of system cost among all cases, and the system costs are lower than for the cases of higher demand. This is due to the fact that in the low demand case, especially for

$d > 0.4$, more than 90% of the paratransit requests are served by bus, and thus the margin of cost improvement is low compared to the high demand case in which less than 80% of the requests are served by bus in all cases.

The study above validated expected trends including the following:

- as the service level constrains are relaxed, i.e. the values of d, $\xi$and $\bar{Q}$ are increased, the system cost improves and SPTS-T balances the transportation requests between the taxi and the bus modes
- d is the most important parameter; for values from 0.1 to 0.5 it affects significantly both the system cost and the respective modal split
- in high demand cases, more paratransit requests are served by taxis.

However, the study also indicated some interesting facts that may be useful to SPTS-T designers, such as:

- In terms of d, there seems to be a threshold, beyond which system performance no longer improves. Thus, d should not be increased beyond this threshold, avoiding unnecessary dissatisfaction to regular service passengers. In the cases of Fig. 4.5a and 4.5b, the appropriate threshold value is $d = 0.5$.
- In low demand cases, providing services of inferior quality with respect to ξ and $\bar{Q}$ does not improve significantly system costs; thus, the system may provide higher quality services for the same cost.
- In high demand cases, allowing clients to share rides in a taxi, improves the system cost for the same value of ξ. Note also that the cases $\bar{Q} = 1$, $\xi = k \cdot f$ and $\bar{Q} = 2$, $\xi = 3 \cdot f$ achieve similar values of cost for all values of d. The latter indicates that the service quality, in terms of waiting and ride times, may be significantly improved under the same costs, by allowing ride sharing in taxis.

4.3.5 SPTS-T design guidelines

The results of the current study suggest that the following issues should be considered for designing an integrated SPTS-T service.

First, it is important to estimate the life cycle costs of taxis vs. buses in the particular environment. The resulting comparison will be used in setting the appropriate value of the

cost scaling factor α. However, the value of α should consider the losses of bus revenues due to the route deviations as discussed above.

Subsequently, the designer should define the service parameters in a way that is similar to the design process of SPTS presented in the previous Chapter. To do so, the designer should consider the tradeoff between service level and system costs. Figures 4.4 and 4.5 indicate that the system cost, after a significant reduction, is stabilized for $d \geq 0.5$ and for all values of ξ and $\bar{Q}$. In this region of parameter values, it seems that allowing the taxi vehicles to carry more than one paratransit clients ($\bar{Q} = 2$) allows the system to achieve satisfactory system costs for low values of ξ, i.e. better service in terms of waiting and ride time. Thus, the system designer, having defined the value of d, needs to determine the values of ξ and $\bar{Q}$ in order to achieve the desired balance of system costs and service level.

Chapter 5. Solving the bus evacuation problem and its extensions

As stated in Chapter 2, paratransit transport also concerns cases in which transit-depended populations may need assistance in emergency situations. Characteristic examples include evacuations in response to an advance notice of an emergency, such as a natural disaster. In this Chapter we address two problems that fall in this type of paratransit transportation situations. First we address the problem of planning bus-based evacuation in case of an upcoming threat. Specifically, we deal with the Bus Evacuation Problem (BEP), which was presented, discussed and studied by Bish (2011).

In BEP, a number of homogenous buses, which are located initially at one or more origins (depots), are tasked to pick-up known numbers of evacuees from known pickup locations and transport them to available shelters. Each of the latter may accept up to a certain number of evacuees. The objective is to complete the evacuation in the minimum possible time period,

which begins when the first bus departs from its depot and ends when the last evacuee arrives to a shelter. Among solutions of equal evacuation time, the preferred one completes the evacuation operation at minimum cost (see also Bish, 2011).

Notwithstanding the fact that BEP addresses cases in which there is advance notice of threat, similar characteristics exist in circumstances after an event; as an example consider the transport of wounded military personnel to medical facilities. The latter may be hospitals able to handle a large number of incidents, or smaller units that are temporarily set near the incident area and have limited capacity. Typically, the medical rescue chain comprises the following phases: (a) Search and rescue, in which crews search for causalities in the area of interest, (b) triage, in which casualties are being sorted in different levels of medical urgency (see Mezzeti, 1998), (c) first aid on the spot, and (d) transportation to the available treatment facility. In such operations, the response time is vital and the available resources should be effectively used. Principles of casualty management after terrorist disasters have been discussed in Frykberg (2004). In this work we focus on the fourth phase of the medical rescue chain, and address the problem we have designated as the Casualty Evacuation Problem (CEP).

In both BEP and CEP the transportation operation bears similarities with the Vehicle Routing Problem with Satellite Facilities (VRPSF) studied by Bard *et al. (*1998). In VRPSF, a set of clients must be served within a maximum time limit by a fleet of homogeneous vehicles, which are initially located at the central depot. The vehicles may visit the central depot or any of the satellite facilities for stock replenishment more than once. All facilities, including the depot, hold unlimited commodity supplies. Besides the conceptual differences with VRPSF, the BEP/CEP cases studied here involve limited shelter capacity (or limited supplies in the satellite facilities in the VRPSF setting). This implies that the vehicles compete for the available capacity (inventory), and thus additional decisions are involved.

Other known problems that share common attributes with BEP, and its CEP extension, include the Multi-trip Vehicle Routing Problem (Taillard *et al.*, 1995; Brandão and Mercer, 1998; Petch and Salhi, 2003; Salhi and Petch, 2007), in which only one depot is available for vehicles to replenish their load between trips, and the VRP with Intermediate Facilities or with inter-depot routes (Crevier *et al.*, 2007; Tarantilis *et al.,* 2008), in which the vehicles may visit intermediate depots for load replenishment along their trips.

A related, but more general, class of problems includes the Location-Routing Problem (LRP) also described in Chapter 2. Prins *et al.* (2006) address an LRP that considers depots and vehicles with limited capacities, as well as fixed costs to establish a depot or to use a vehicle. The scope is (a) to determine which depots to enable, and (b) to design all vehicle routes. The latter start and return to the same (enabled) depots, and the problem's objective is to minimize the total cost comprising all fixed and route costs. The authors posed and solved a number of benchmark instances by applying a two phase meta-heuristic that combines a GRASP procedure with path relinking. We leverage some of their benchmark problems in the current work to test our models and approach. Coutinho-Rodrigues *et al. (*2012) proposed a multi-objective form of the LRP in order to identify paths and the location of shelters for the design of urban evacuation plans. Their approach was tested in a simulated fire situation in the city of Coimbra, Portugal.

To deal with BEP/CEP we have proposed new two-index, mixed integer-linear formulations, and we have developed a Hybrid Column Generation (HCG) framework to provide efficient solutions in reasonable computational times. Our proposed solution framework has been inspired by the findings of Parragh and Schmid (2013), in which an HCG framework and the Large Neighborhood Search (LNS) method have been combined to solve the Dial-a-Ride Problem. LNS is a stand-alone method, first introduced by Shaw (1998), in which in each iteration a complete solution is partially destroyed and reconstructed. Numerous destroy and repair operators have been introduced over the last decade for solving routing problems; see Pisinger and Ropke (2010). In the HCG framework of Parragh and Schmid (2013) the Variable Neighborhood Search (VNS) heuristic has been applied to provide additional routes (columns) into a binary program in a similar fashion to the column generation technique. For the VNS method we refer to Mladenovic and Hansen (1997), Hansen and Mladenovic (2001).

The remainder of the Chapter is structured as follows: Section 5.1 describes the two-index formulation for BEP and the appropriate enhancements to formulate CEP. Section 5.2 describes the proposed solution method. Section 5.3 presents an extensive experimental study on the efficiency of the hybrid method for the two evacuation cases studied.

5.1 Mathematical formulations for BEP and CEP

The two cases under investigation, BEP and CEP, share similar constraints and concepts, although their objectives and transportation policies are different. The objective of BEP aims

to determine the set of routes that minimize the total evacuation time; among the possibly multiple solutions with the minimum evacuation time, the one with the minimum operational cost (time) is selected. A more appropriate objective for CEP is to minimize the total time required to transport casualties to the medical facilities; this is equivalent to minimizing the average transportation time per casualty. This seemingly slight difference affects significantly routing decisions and route quality. In this Section, first we present a mixed integer programming model for BEP, and, subsequently, we enhance it to meet CEP's special characteristics.

5.1.1 A 2-index formulation for BEP

Let $M = \{1, \dots, \mu\}$ be the set of available shelters each of capacity $P^m, m \in M$, and let $K = \{1, \dots, v\}$ be the set of available vehicles (buses), all of capacity Q. All vehicles initiate and finish their routes from/to a single depot. However, in order to model the problem by a 2-index formulation, we define originating and ending vehicle depot locations for each bus through sets S and E, where $S = \{s^k | k \in K\}$ and $E = \{e^k | k \in K\}$. Each of those locations may be considered as a single parking space and are used in order to address the requirement for the buses to return to the ending depot, thus, allowing for the following:

- Separating the total operation time (or operational cost) from evacuation time; note that the end of the evacuation time is defined by the time the last evacuee arrives to a shelter, and the total operation time is the sum of the operation times of all vehicles (till they return to the ending depots).
- Ensuring that the parking spaces will be utilized appropriately to accommodate all vehicles.

Note that the current notation may also address the case studied by Bish (2011), by setting all traveling distances of the arcs ending at nodes of E to zero.

Let C be the set of all nodes representing the evacuee locations, called *pick-up* locations hereafter. Moreover, let $d_i, i \in C \cup S$ be the number of evacuees waiting at each node of the graph, assuming of course that $d_i = 0, i \in S$ and $d_i \in \mathbb{N}^+, i \in C$. Note that a vehicle visits a pick-up location only when it is able to pick-up all evacuees from this location. Thus, an appropriate number of multiple nodes need to be established for each location $i \in C$ with $d_i > Q$.

For modeling the possibly multiple visits of vehicles to shelters within the 2-index formulation, we define a set of new nodes $D = \{D_i^m | i \in C \cup E, m \in M\}$. For each pick-up or ending depot location $i \in C \cup E$, set D contains multiple *pseudonodes* D_i^m, one for each shelter $m \in M$. Under the above we assume that:

(a) Each real node $i \in C \cup E$ is visited exactly once.

(b) The nodes $m \in M$ are not visited. If a vehicle intends to visit shelter m to drop-off evacuees, then the vehicle will visit node D_i^m, where $i \in C \cup E$ is the location to be served by the vehicle after shelter m in order to pick-up evacuees, or end its trip, respectively.

(c) Each pseudonode may be visited at most once.

(d) A vehicle visiting pseudonode D_i^m may also visit another shelter $t \in M \backslash \{m\}$ to drop-off more evacuees prior to visiting location $i \in C \cup E$. In this case, the vehicle, after pseudonode D_i^m will visit pseudonode D_i^t.

Using the above notation, a shelter $m \in M$ may be visited by the vehicles at most $|C| + |E|$ times, that is, once prior to each pick-up and each ending depot location.

Figure 5.1 illustrates a case with $k = 1, |M| = 2$ and $|C| = 2$. In this case three pseudonodes are defined for each $m \in M$. Assume that the vehicle is planned to follow route: $S_1 - C_1 - M_2 - C_2 - M_2 - M_1 - E_1$. Given our formulation, this route will be realized as follows: $S_1 - C_1 - D_{C_2}^{M_2} - C_2 - D_{E_1}^{M_2} - D_{E_1}^{M_1} - E_1$.

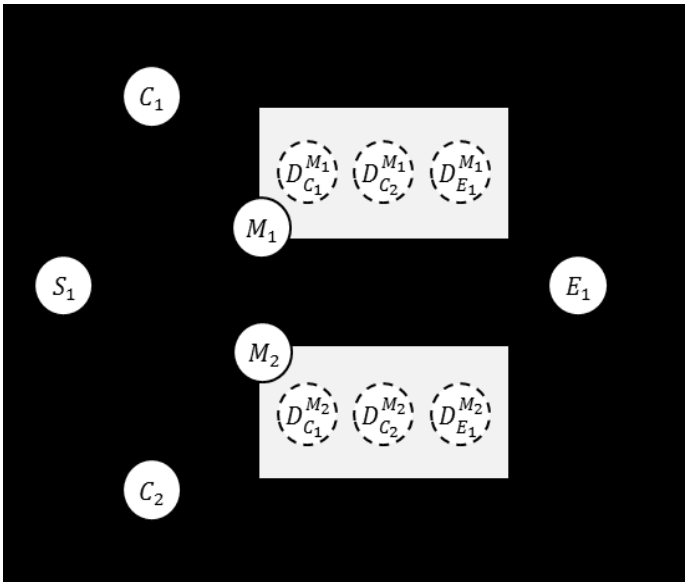


Figure 5.1: Illustration of the problem nodes in a simple example

Figure 5.2 displays all arcs of A for the simple case of Figure 5.1. The arrows of Figure 5.2 are in line with assumptions (b) and (d) above. For instance, considering the pseudonodes with respect to C_1, arrows (arcs) are directed: (a) from nodes $D_{C_1}^{M_1}$ and $D_{C_1}^{M_2}$ to C_1, and (b) from $D_{C_1}^{M_1}$ to $D_{C_1}^{M_2}$, as well as from $D_{C_1}^{M_2}$ to $D_{C_1}^{M_1}$.

We formalize now the definition of directed graph $G(V, A)$, in which $V = S \cup E \cup C \cup D$ and A is the arc set. The arcs in A connect:

(a) each node of set S to all pick-up and ending depot locations. Arcs from the starting to the ending depot locations are used to model idle vehicles. A direct connection to shelter pseudonodes may be excluded, since the vehicles initiate empty from their originating depot locations

(b) each pick-up location $i \in C$ to all other pick-up locations $j \in C\backslash\{i\}$ and to every other pseudo-node except those that are related to i, i.e. $\{D_j^m | j \in (C\backslash\{i\}) \cup E, m \in M\}$. Again, traveling from a pick-up location to an ending depot location is prohibited since all vehicles must conclude their trip empty

(c) pseudonodes D_i^m to their reference location $i \in C \cup E$ and to other pseudonodes of the same reference location, $\{D_i^t | t \in M\backslash\{m\}\}$.

Thus, A may be defined as follows:

$$A = \{(i,j) | i \in S, j \in C \cup E\} \cup \{(i,j) | i \in C, j \in (C\backslash\{i\}) \cup (D \backslash \{D_i^m | m \in M\})\} \cup \{(i,j) | i = D_h^t, h \in C \cup E, t \in M, j \in \{h\} \cup \{D_h^m | m \in M\backslash\{t\}\}\}$$

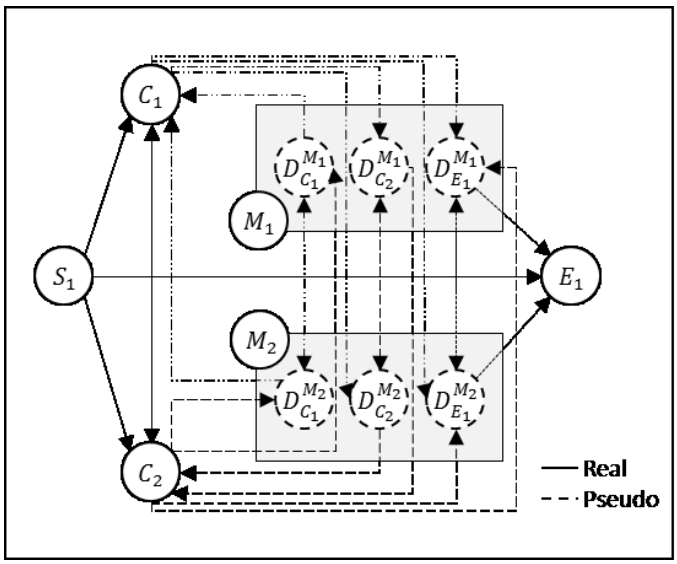


Figure 5.2: Illustration of the graph of the simple example of Fig. 5.1

Note that the proposed formulation reduces the numbers of arcs significantly.

Let t_{ij} be the time to traverse arc $(i,j) \in A$, assuming the travel times satisfy the triangular inequality. Let also:

- $w_i, i \in V$ be the time that the vehicle arrives to node i; note that within the 2-index formulation a node may be visited by any of the available vehicles, but only by one of them
- $q_i, i \in V$ be the number of evacuees onboard the vehicle just before its arrival to node i
- $d_i^m, i \in C \cup E, m \in M$ be the variables representing the number of the passengers dropped off at shelter m when the vehicle visits pseudo-node D_i^m, that is, just before it visits node i
- x_{ij} be assigned the value 1 if arc $(i,j) \in A$ is traversed, and 0 otherwise
- T_{evac} be the duration of the evacuation, i.e. the time span defined by the start of the evacuation until the time the last evacuee arrives to a shelter.

Then the objective function of BEP is defined as follows:

$$TC = \min\{T_{evac} + \frac{1}{L}\sum_{i \in E} w_i\} \tag{5.1}$$

where the second term is the total vehicle operation time (cost) and L ensures that the first term of objective function (5.1) dominates lexicographically the second term; i.e. in case of multiple solutions with the same evacuation time the one with the lowest total travel time will be preferred. To do so the value of L should be greater than the maximum possible value of the second term of (5.1) that is: $L > \sum_{(i,j) \in A} t_{ij}$. Optimization of (5.1) is subject to:

$$T_{evac} \geq w_j - \sum_{i \in V | (i,j) \in A} t_{ij} x_{ij},\ j \in E \tag{5.2}$$

$$\sum_{j \in V | (i,j) \in A} x_{ij} = 1,\ i \in C \cup S \tag{5.3}$$

$$\sum_{j \in V | (i,j) \in A} x_{ij} \leq 1,\ i \in D \tag{5.4}$$

$$\sum_{i \in V | (i,j) \in A} x_{ij} = 1,\ j \in E \tag{5.5}$$

$$\sum_{i \in V | (i,h) \in A} x_{ih} - \sum_{j \in V | (h,j) \in A} x_{hj} = 0,\ h \in V \backslash (S \cup E) \tag{5.6}$$

$$w_i + t_{ij} - B(1 - x_{ij}) \leq w_j,\ (i,j) \in A \tag{5.7}$$

$$q_i + d_i - B(1 - x_{ij}) \leq q_j, \ i \in C \cup S, \{j \in V | (i,j) \in A\} \quad (5.8)$$

$$q_{D_i^m} - d_i^m - B\left(1 - x_{D_i^m j}\right) \leq q_j, \ i \in C \cup E, m \in M, \{j \in V | (D_i^m, j) \in A\} \quad (5.9)$$

$$\sum_{i \in C \cup E} d_i^m \leq P^m, \ m \in M \quad (5.10)$$

$$\sum_{m \in M} \sum_{i \in C \cup E} d_i^m = \sum_{i \in C} d_i \quad (5.11)$$

$$\sum_{i \in S \cup E} q_i + \sum_{i \in S} w_i = 0 \quad (5.12)$$

$$0 \leq q_j \leq Q \sum_{i \in V | (i,j) \in A} x_{ij}, \ j \in V \backslash (S \cup E) \quad (5.13)$$

$$0 \leq w_j \leq B \sum_{i \in V | (i,j) \in A} x_{ij}, \ j \in V \backslash S \quad (5.14)$$

$$d_i^m \in \mathbb{N}_0, \ i \in C \cup E, m \in M \quad (5.15)$$

$$x_{ij} \in \{0,1\}, \ (i,j) \in A \quad (5.16)$$

Regarding the path constraints: Inequality (5.2) ensures that the duration of the evacuation is greater than the time of the latest visit of a vehicle to shelter, that is, the time that the last evacuee arrives to a shelter. Note that the time of the last visit to a shelter is defined through the time of arrival to the ending locations. Equation (5.3) ensures that all pick-up locations are visited once, and that all vehicles depart from the originating depot locations. Inequality (5.4) ensures that each pseudo-node may be visited once or not visited at all. Note that a vehicle may not be used (i.e. directed to the ending depot location without visiting any pick-up location), if this is favored by the objective function. According to Equation (5.5) all vehicles must arrive to their ending depot locations. Equation (5.6) ensures that if a vehicle arrives to a location, other than the originating and ending depot locations, it will also depart from it.

Regarding the service and resources constraints: Inequality (5.7) preserves the appropriate flow of time when traveling from location i to location j, assuming that $B \gg 1$. Inequality (5.8) defines the change of the vehicle's load after visiting node i. Inequality (5.9) defines the change of the vehicle's load when it departs from a shelter. Inequality (5.10) ensures that all pseudonodes of the same reference shelter will not accept more evacuees than the capacity of this shelter. According to Equation (5.11) the number of evacuees to be dropped-off at all shelters will equal the total evacuee demand. Equation (5.12) ensures that all vehicles initiate from the originating depot locations empty, at time equal to zero, and conclude their trips at the ending depot locations empty. Inequality (5.13) ensures that at any location, other than the originating and ending depot locations, the number of evacuees aboard the vehicles will

not exceed the vehicle's capacity nor it will be negative. Furthermore, if the location is not visited by any vehicle, then the number of evacuees on the vehicle will be equal to zero. Similarly, Constraint (5.14) ensures that the time of arrival to any node will be greater or equal to zero and, specifically, it will be equal to zero if the location is not visited. Constraint (5.15) ensures that the number of evacuees being dropped-off at the shelters will be equal to or greater than zero. Finally, Constraint (5.16) defines the binary nature of the arc variables.

Considerations regarding split service

An important issue discussed by Bish (2011) is the significance of performing split service, which in BEP is twofold: (a) the vehicle may drop-off the evacuees on board to different shelters (similar to the multi-depot case of VRP), and (b) during each visit of a vehicle to a physical evacuee pick-up location, and in order to preserve optimality, the number of evacuees to be picked up should be free to vary between the residual capacity of the vehicle and the number of evacuees present at this location during the vehicle's visit (similar to split delivery in case of VRP problems). Ergo evacuees waiting at the same pick-up location may be picked up by different vehicles or in different trips, if this favors the objective function.

The former issue (a) is addressed in the proposed model by decision variable d_i^m, which refers to the number of passengers dropped off at shelter m when the vehicle visits pseudo-node D_i^m.

A way to address issue (b) with the proposed two-index formulation, which requires single visits to nodes of C and D, is to define for each pick-up location multiple nodes, one for each evacuee present at this location. Of course this may increase significantly the number of nodes and the size of the network. A way to mitigate this increased complexity, is to discretize the demand at an appropriate level of granularity, and use a separate pick-up location for each resulting portion of the demand. To do so, assume a discretization step U, $1 \leq U \leq Q$ and $U \in \mathbb{N}^+$. For each pick-up location $i \in C$ with demand $d_i > U$, we define, $\lceil d_i/U \rceil$ replicate nodes, each with demand up to U and total demand of the replicate nodes equal to the demand of the original node. If $U = Q$, then replicate nodes are introduced only for cases with $d_i > Q$. If $U = 1$, then all possible split pick-ups are allowed. The impact of split service on solution quality and complexity is further studied in subsection 5.3.2.

5.1.2 Formulation for CEP

As mentioned in the introduction of this Chapter, the objective for the casualty transportation problem is to minimize the sum of arrival times of casualties at the medical facilities (and

thus, in effect, the average arrival time per casualty). For this objective, additional notation should be introduced to the model. Let:

- R be the set that contains each individual casualty h waiting to be transported
- $\tilde{d}_i^h, i \in V, h \in R$ receive the value 1 if casualty h waits at location i, otherwise the value 0
- $a_h, h \in R$ be the time that casualty h arrives at a medical facility
- $\tilde{d}_i^{mh}$ for $i \in C \cup E, m \in M, h \in R$ receive the value 1 if casualty h is dropped off at pseudo-node m with reference node i
- $\tilde{q}_i^h$ for $i \in V, h \in R$ receive the value 1 if casualty h is onboard a vehicle just before arriving at node i, otherwise the value 0.

The objective function of CEP is defined as follows:

$$TC = \min \sum_{h \in R} a_h \tag{5.17}$$

Subject to Constraints (5.3) – (5.7) and:

$$\tilde{q}_i^h + \tilde{d}_i^h - B\left(1 - x_{ij}\right) \le \tilde{q}_j^h, \;\; i \in C \cup S, \{j \in V | (i,j) \in A\}, h \in R \tag{5.18}$$

$$\tilde{q}_{D_i^m}^h - \tilde{d}_i^{mh} - B\left(1 - x_{D_i^m j}\right) \le \tilde{q}_j^h, \; i \in C \cup E, m \in M, \{j \in V | (D_i^m, j) \in A\}, h \in R \tag{5.19}$$

$$w_{D_i^m} - B\left(1 - \tilde{d}_i^{mh}\right) \le a_h, \; i \in C \cup E, m \in M, \; h \in R \tag{5.20}$$

$$\sum_{i \in C \cup E} \sum_{m \in M} \tilde{d}_i^{mh} = 1, \; h \in R \tag{5.21}$$

$$\sum_{h \in R} \sum_{i \in C \cup E} \tilde{d}_i^{mh} \le P^m, \; m \in M \tag{5.22}$$

$$\sum_{i \in S} w_i + \sum_{i \in S \cup E} \sum_{h \in R} \tilde{q}_i^h = 0, \tag{5.23}$$

$$0 \le \sum_{h \in R} \tilde{q}_i^h \le Q \sum_{j \in V | (i,j) \in A} x_{ij}, \; i \in V \backslash (S \cup E) \tag{5.24}$$

$$\tilde{d}_i^{mh}, \tilde{q}_i^h, x_{ij} \in \{0,1\} \tag{5.25}$$

Inequality (5.18) defines the change in the number of casualties aboard the vehicle as it departs from location i. Inequality (5.19) defines the change in the number of casualties when the vehicle visits a medical facility. Inequality (5.20) defines the time when each casualty arrives at a medical facility. Equation (5.21) ensures that all casualties will be transported to a medical facility. Inequality (5.22) ensures that a medical facility will not receive more casualties than its capacity. Equation (5.23) ensures that all vehicles initiate from the

originating depot locations empty, at time equal to zero, and conclude their trips at the ending depot locations empty. Inequality (5.24) ensures that the number of casualties aboard a vehicle may not exceed its capacity.

For CEP one may also consider a variation of the objective of BEP; this variation minimizes the evacuation time and, among solutions with the same evacuation time value, it selects the one that corresponds to the lowest sum of arrival times of the casualties to the medical facilities.

$$TC = \min T_{evac} + \frac{1}{\tilde{L}} \sum_{h \in R} a_h \tag{5.26}$$

In this case $\tilde{L} = |R| \cdot \sum_{(i,j) \in A} t_{ij}$. For the rest of the Chapter we will refer to this formulation as *CEP-MO*.

5.2 A hybrid solution framework for BEP

We have developed a solution framework to solve BEP, which employs an appropriate combination of methods to generate and refine feasible BEP solutions arriving to an efficient near-optimal one. This framework is presented in Figure 5.3. It initially generates a feasible solution through a Nearest Neighbor like heuristic (NN-BEP), and subsequently it applies a two phase improvement process. Phase I produces a good solution for the problem by employing:

- Large Neighborhood Search (LNS) to destroy and repair complete solutions and create a pool of feasible alternative routes (or column pool Φ') to be used in the next step to synthesize a good complete solution
- A mixed integer linear program for BEP, the Restricted Master Problem (RMP), that is solved to synthesize the best combination of feasible routes in column pool Φ' and generate a complete solution
- A Travelling Salesman Problem Improvement (TSPI) heuristic to improve one by one the vehicle routes of the complete solution produced by RMP.

Phase I runs until no better solution could be found. Phase II improves the solution of Phase I by diversifying the search. This is done by employing:

- A column generation-like solver that includes:
    - Linear relaxation of the RMP (LRMP) which is solved with respect to column pool Φ' to identify the best solution and produce the respective dual information

  - Variable Neighborhood Search (VNS) that exploits the dual information produced by the solution of LRMP to produce vehicle routes to be added to column pool Φ'.
- A similar to Phase I sequence of steps that includes:
  - LNS to destroy and repair complete solutions, starting from the best identified up to that point
  - RMP to identify a complete solution by selecting the best combination of all partial solutions produced during both LNS and VNS
  - TSPI to improve one by one the vehicle routes of the complete solution produced by the RMP.

Phase II is executed for predefined number of iterations N^{iter}, in which the search gradually widens along with T, a temperature-like parameter. The constitutional elements of the proposed framework are described in detail in the following sub-sections of this Section.

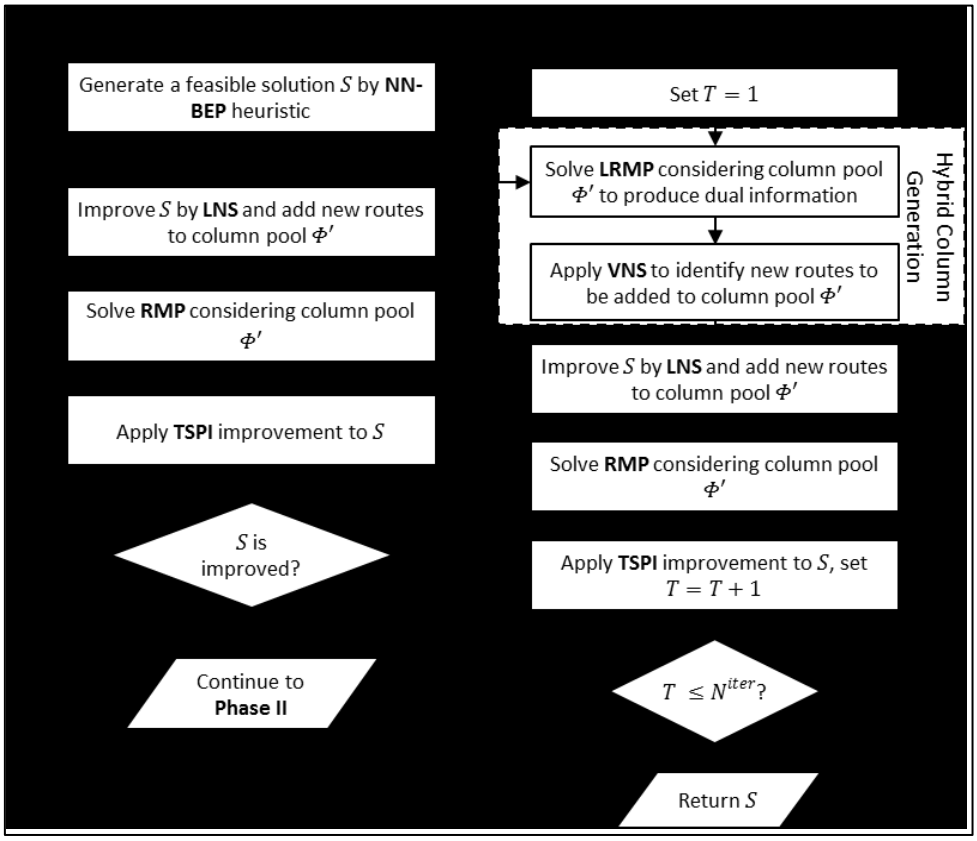


Figure 5.3: Outline of the hybrid framework

A similar method has been proposed by Parragh and Schmid (2013) to solve the Dial-a-Ride problem. The difference of their work with the current one is that Parragh and Schmid (2013) used parts of the algorithms of Phases I and II as stand-alone ones, and compared their results for a predefined number of iterations. We have integrated the two algorithms by using the

first to reach an efficient solution, which is subsequently refined it in the second phase. In this last Phase we enlarge the search variables considering, in addition to the vehicle routes, the number of evacuees to be dropped off to the shelters (see subsection 5.2.5). The efficacy of integrating the two algorithms is validated in the experimental Section below.

5.2.1 Initial solution by the nearest neighbor heuristic for BEP

In order to create an initial solution we apply a simple heuristic based on the nearest neighborhood procedure. The heuristic creates one route for each of the k available vehicles, in which the vehicles are restricted to visit at most $\left\lceil \frac{|C|}{k} \right\rceil$ pick-up locations. This restriction is imposed in order to share the workload among vehicles. Note that a solution of a single vehicle transferring all evacuees to the shelters would extent the evacuation time.

The procedure generates routes for the vehicles one by one. The first vehicle initiates from the starting position and travels to the nearest of those pick-up locations, for which the total number of waiting passengers may be served by the vehicle's residual capacity. As soon as the capacity of the vehicle is not adequate to serve any other pick-up locations, the vehicle returns to the nearest shelter that may accept all evacuees on-board to drop them off. If there is no shelter with enough capacity to accept all evacuees on board, the vehicle travels to the nearest available shelter to drop-off as many evacuees as possible, and then proceeds to the next available shelter. This is repeated until all on-board evacuees are dropped-off. When the vehicle visits $\left\lceil \frac{|C|}{k} \right\rceil$ pick-up locations, or no other unvisited pick-up locations are available, the vehicle returns to the nearest available shelter to drop-off its evacuees (or shelters as before) and travels to the ending position. Subsequently, the second vehicle initiates from the starting position following the same procedure. This continues until all k available vehicles return to the ending position.

5.2.2 Large neighborhood search (LNS) for BEP

The LNS, as introduced by Shaw (1998), is a heuristic method that improves a certain solution by iteratively destroying and repairing it. This is done by removing randomly a number of its elements and then repairing it by reinserting those elements in new positions aiming to improve the initial solution. We adapt this concept to meet the special characteristics of BEP. Our LNS approach is inspired by the approaches of Ropke and

Pisinger (2006) for the Pick-up and Delivery Problem with Time Windows, as well as the findings of Parragh and Schmid (2013) for the Dial-a-Ride Problem.

Given a solution, at each LNS iteration, l pick-up locations are removed from the routes of this solution and are re-inserted by the repair operations. As in Parragh and Schmid (2013): (a) the number l is randomly selected in the range $[0.1|C|, 0.5|C|]$, (b) the removal operations are the random removal and the worst removal, and (c) the repair operations are the random insertion, the greedy insertion and the k-regret insertion for $k \in \{2,3,4,5\}$. Interested readers are referred Parragh and Schmid (2013) for further details.

At each iteration of LNS, P solutions are destroyed and repaired simultaneously (i.e. in a parallel process). Up to N^{LNS} solutions are destroyed and repaired in total. At the first iteration of the heuristic the initial solution is replicated P times to build the set of current solutions (first generation) $s_p, p = 1, \dots, P$. Then, P new solutions s_p^* are produced for $p = 1,2, \dots, P$. For each of the N^{LNS}/P iterations following the first one, the next generation of current solutions s_p is created by selecting randomly P solutions (one solution may be selected more than once) from the latest generation, s_p^*. The selection process uses the probability F_p^* given below:

$$F_p^* = \frac{f(s_p^*)}{\sum_{p=1,\dots,P} f(s_p^*)}$$

where $f(s_p^*) = e^{[obj(s_{best}) - obj(s_p^*)]/obj(s_{best})}$, assuming that $obj(s_p^*)$ is the value of the objective function of solution s_p^* and $obj(s_{best})$ is the value of the objective function of the best identified solution so far. Thus, a solution with better value of the objective function is more likely to be selected for the next iteration of LNS. Algorithm 5.1 presents the pseudo-code of LNS applied for BEP.

To diversify our search, we implement two versions of the greedy and the k-regret insertion operators (see Parragh and Schmid, 2013) as discussed below. In the first version the best position is initially selected without considering any feasibility issues and then the route is modified by a proposed heuristic (PRS) to be feasible; in the second version feasibility is guaranteed for every insertion position. Note that only feasible solutions are produced at the end of all repair operations.

Algorithm 5.1: Large Neighborhood Search for the BEP

1. Set $N = 1$
2. Set 1st solution generation $s_p = s_{best}$ **for** $p = 1$ ***to*** P
3. **While** $N < N^{LNS}$
3.1. **for** p =1 **to** P
3.1.1. Set l by $unif(0.1|C|, 0.5|C|)$
3.1.2. **Remove** l pick-up locations from s_p
3.1.3. **Re-insert** the l pick-up locations to create s_p^*
3.1.4. **if** $obj(s_p^*) < obj(\mathrm{s_{best}})$ **then** set $s_{best} = s_p^*$
3.1.5. Add new unique routes of s_p^* in Φ'
3.2. Set new solution generation s_p from s_p^* by F_p^*
3.3. Set $N = N + P$
4. **Return** s_{best} and Φ'

The Probabilistic Return to Shelters Heuristic

For BEP, feasibility in both the greedy and the k-regret insertion operators is treated by the proposed Probabilistic Return to Shelters (PRS) heuristic. PRS is a constructive heuristic, which starting from any route (feasible or not) produces a new feasible route by maintaining the sequence of the route's pick-up locations and rearranging the existing, or inserting new, returns to shelters in order to drop-off evacuees, and, thus, ensure feasibility of the capacity constraint. Note that returning to a shelter only when the vehicle cannot pick-up any other evacuees is not always in favor of the objective function value, since preemptive visits to shelters that are close to the route being executed, even if the vehicle is partially loaded, may improve the total evacuation time.

To any given route (old route), the heuristic applies the following simple steps:

Step 1. Remove all shelters from the old route and update the residual capacity information of the corresponding shelters by adding to the number of evacuees those dropped off during the old route.

Step 2. Start a new route from the origin depot.

Step 3. Add the first (or next) pick-up location to the new route following the sequence of pick-up locations of the old route, and update all related information (travelling time, evacuees on board, etc.).

Step 4. If the probabilistic criterion described below is met, or the evacuees waiting at the next pick-up location of the sequence exceed the residual capacity of the vehicle, insert to the route the shelter which causes the minimum change to the value of the objective function and drop-off as many evacuees as possible to the inserted shelter.

Step 5. Repeat steps 3 and 4 until the last pick-up location is added to the new route.

Step 6. Visit the nearest shelter (or shelters) to drop-off all evacuees on board and add the destination depot to the new route.

Step 7. Replace the old route with the new one.

The probabilistic criterion assesses, at any pick-up location served by the route, preemptive returns to shelters if these favor the objective function, or returns that are required for the vehicle to continue its mission. It is the product $P_Q \cdot P_D$ of two terms that consider the residual capacity of the vehicle, and the relative distance of the vehicle from the nearest available shelter, respectively; it also takes into account the progress of the algorithm. Specifically:

- $P_Q = e^{\frac{q-Q}{T}}$: assuming that q is the number of evacuees on board and Q is the total vehicle capacity. T is a temperature parameter (inspired by simulated annealing), which, in our case, starts from the value 1 and increases by 1 at each iteration of the algorithm. This schedule for T was developed through tuning experiments
- $P_D = e^{\frac{-\Delta t/\bar{t}}{aT}}$: where $\bar{t}$ is the average travelling time and Δt is the time difference between two alternative routes that join two consecutive pick-up locations: the route through the nearest shelter, and the direct route. a is an appropriate scaling constant.

As described above, a shelter is always inserted if the vehicle is full and not able to continue its mission unless it alights some passengers. Additional discussion with respect to the probabilistic criterion $P_Q \cdot P_D$ and the values of the respective parameters is provided in Appendix A.

5.2.3 Restricted master problem

First consider the Master Problem (MP) for BEP formulated as a MILP that considers single feasible routes as variables (or columns): Let Φ denote the set of all feasible routes of a vehicle considering that it initiates (at time 0) from the origin and returns to the destination depot after visiting a number of pick-up locations. Let $\bar{c}_j$ be the time that the last evacuee (or the last group of evacuees) is dropped-off at a shelter during route $j \in \Phi$; also let $\bar{t}_j$ be the total travel time of route j, that is, $\bar{c}_j$ plus the time to travel to the ending depot location. Let u_{ij} be a constant that assumes the value 1 if location $i \in C$ is visited by route j (and 0 otherwise). Furthermore, let $\bar{d}_{mj}$ be the total number of evacuees being dropped-off to shelter $m \in M$ during route j; this number is known for each route and may include multiple visits to

m. Finally, let the binary variable $y_j, j \in \Phi$ receive the value 1 if route j participates in the solution of the problem, or 0 otherwise. Then the MP for BEP may be formulated by the following program:

$$TC = \min T_{evac} + \frac{1}{L}\sum_{j\in\Phi} \bar{t}_j y_j \tag{5.27}$$

subject to:

$$T_{evac} \geq \sum_{j\in\Phi} u_{ij}\bar{c}_j y_j, \;\; i \in C \tag{5.28}$$

$$\sum_{j\in\Phi} u_{ij} y_j = 1, \;\; i \in C \tag{5.29}$$

$$\sum_{j\in\Phi} \bar{d}_{mj} y_j \leq P^m, \;\; m \in M \tag{5.30}$$

$$\sum_{j\in\Phi} y_j = |K| \tag{5.31}$$

$$y_j \in \{0,1\}, \; j \in \Phi \tag{5.32}$$

Inequality (5.28) is the equivalent of Inequality (5.2) and defines the value of the maximum evacuation time. Inequality (5.29) ensures that all pick-up locations will be visited by one route. Inequality (5.30) safeguards that each shelter will receive less evacuees than its capacity. Finally, Inequality (5.31) insures that all vehicles will depart from the origin depot. Constraint (5.32) defines the binary nature of the decision variables.

The MILP program (5.27) to (5.32) may also describe the CEP-MO by replacing $\bar{t}_j$ with $\bar{a}_j$, the sum of arrival times of all casualties to the medical facilities during route j, and L with $\tilde{L}$. Thus, the CEP Master Problem is modeled by the following objective function (5.33) subject to constraints (5.28) to (5.32):

$$TC = \min \sum_{j\in\Phi} \bar{a}_j \, y_j \tag{5.33}$$

Let the Restricted Master Problem (RMP) be the restricted version of the problem corresponding to (5.27) to (5.32) that considers only a subset $\Phi' \subseteq \Phi$. RMP is solved in both Phases of the framework (see Fig. 5.3). Subset Φ' is constantly enhanced with new columns by LNS in Phase I, and both LNS and VNS in Phase II.

5.2.4 Travelling salesman problem improvement (TSPI) heuristic

The TSPI heuristic is applied in both Phases after the solution of the RMP. It is essentially a local search procedure that operates on a single route in order to improve the total traveling

time. To do so, it seeks the optimal visiting sequence for the sections of the route between successive visits to shelters. Consider such a route portion that departs from a shelter, visits a set of pick-up locations and then visits a shelter again. The visit sequence of the intermediate pick-up locations does not affect the feasibility of the path in terms of the vehicle capacity, or the decision of how many evacuees to drop-off at the visit to the later shelter. However, optimizing the travelling time between the two shelter visits will improve the overall routing cost of the solution, the respective evacuation time, and the sum of the arrival times of the evacuees (or casualties).

For a route j, the TSPI heuristic starts from the origin node (shelter), identifies the next shelter along the route (destination node), and considers all the intermediate pick-up locations to comprise the node set. Then problem is formulated as a TSP (Miller *et al.*, 1960) and solved by a MILP solver to identify the optimal node sequence that minimizes the total travelling time such that: (a) the first node of the sequence is the origin node path, (b) the last node of the sequence is the destination node, and (c) all nodes of the node set must be visited once. Subsequently, the current destination node is set as the origin node and the next shelter in the sequence as the destination node, and the procedure is repeated until the route is completed and refined.

5.2.5 Hybridization of column generation

The hybrid column generation is applied in Phase II of the framework, and follows the concept of the column generation technique. In column generation a linear program is solved to optimality by iteratively solving restricted versions of the problem. The restricted versions consider only a subset of the variables of the main problem. This subset is enhanced with new variables in each iteration of the method. The new variables are, in turn, generated by solving a sub-problem with respect to the dual information of the solution of the restricted version of the problem; interested readers may refer to Desaulniers *et al.* (2005). Integer and mixed integer-linear programing problems may be solved by combining the column generation with Branch and Bound. In the BEP case, the hybrid column generation comprises the linear relaxation of RMP, and employs VNS as the sub problem.

Linear relaxation of RMP and the enhancement of column pool Φ'

The linear relaxation of RMP (LRMP) may be defined by replacing constraint (5.32) with:

$$y_j \geq 0, \quad j \in \Phi' \tag{5.34}$$

The sub-problem, in our case the VNS, is applied to generate routes with negative reduced cost with respect to the LRMP and, thus, enhance Φ'. The reduced cost of a route is calculated with respect to the shadow (or dual) prices of the LRMP and corresponds to the improvement of the objective function value if the respective route is added to the solution. Specifically, the reduced cost of any route j with respect to BEP may be calculated as follows:

$$rc_j = \frac{\bar{t}_j}{L} - \sum_{i \in C} u_{ij} \bar{c}_j \gamma_i - \sum_{i \in C} u_{ij} \pi_i - \sum_{m \in M} \bar{d}_{mj} \sigma_m - \kappa \tag{5.35}$$

Where γ_i are the shadow (or dual) prices with respect to Constraint (5.28), π_i are the shadow prices with respect to Constraint (5.29), σ_m are the shadow prices with respect to Constraint (5.30) and κ is the shadow price with respect to Constraint (5.31). The reduced cost (5.35) may also describe the CEP-MO by replacing $\bar{t}_j$ with $\bar{a}_j$, and L with $\tilde{L}$. While for CEP the respective reduced costs of any route j may be calculated as follows:

$$rc_j = \bar{a}_j - \sum_{i \in C} u_{ij} \pi_i - \sum_{m \in M} \bar{d}_{mj} \sigma_m - \kappa \tag{5.36}$$

Note that in the hybrid framework, column pool Φ' is common for the RMP and the LRMP, since the corresponding route of each column is feasible for both RMP and LRMP. Furthermore, as already mentioned above, Φ' is also enhanced with new columns, in both Phases by LNS and VNS. Formally, we do not consider LNS as part of the hybrid column generation of Phase II, since LNS does not utilize the dual information of LRMP to identify new columns, which is the core part of the column generation technique.

Variable neighborhood search (VNS)

As introduced by Hansen and Mladenović (2001), VNS in a shaking phase that iteratively modifies an incumbent solution using pre-selected operations to generate neighbor solutions. These operations (within the VNS) are called neighborhood structures N_k for $k = 1, \dots, k_{max}$ and are discussed later in this Section.

In the current work, and similarly to Parragh and Schmid (2013), we use VNS to modify columns (routes) of the solution of LRMP to generate new feasible columns with negative reduced cost to be added to Φ'. This modification is performed to the corresponding routes one by one. However, in our approach, instead of targeting only routes with negative reduced cost $rc_j < 0$ to improve the value of LRMP, as in Parragh and Schmid (2013), we identify routes with $rc_j < UB - LB$, where UB is the value of the objective function of RMP (integer

solution) in the current iteration, and LB is the value of the objective function of LRMP (relaxed linear solution) again in the current iteration. In this way we aim for routes that will potentially improve the solution of RMP, providing an objective function value in the interval between LB and UB.

To generate the neighborhood structures N_k of VNS (see above), for each route we use similar settings as in Parragh and Schmid (2013). That is, in any VNS iteration:

- Up to $20 \cdot \lceil k/3 \rceil$% pick-up locations not in the route are being added to the route, for $k = 1,4,7$
- Up to $20 \cdot \lceil k/3 \rceil$% pick-up locations of the route are being removed, for $k = 3,6,9$, or
- Up to $20 \cdot \lceil k/3 \rceil$% pick-up locations of the route are exchanged with pick-up locations not in the route, for $k = 2,5,8$.

For the above we assume that $k_{max} = 9$. The selection of the pick-up locations is biased to the values of the dual variables; that is: (a) for the add operation, the selection probability is proportional to the shadow prices of the LRMP π_i; (b) for the remove operation, the selection probability is inversely proportional to the values of π_i; (c) in the exchange operation one pick-up location is being removed and another is being added again using the latter selection probabilities, respectively.

As stated above, VNS is executed separately for each route of the solution. If a neighborhood structure N_k results in a route that improves the best currently known route in terms of the $UB - LB$ criterion, the procedure starts over with this route as N_1, otherwise the route is discarded and the procedure continues with the next neighborhood $k + 1$. The procedure for a single route terminates (a) after the first N^{new} routes (columns) with reduced cost less than $UB - LB$ are identified, or (b) if no such route occurs until the last neighborhood structure (k_{max}). All new routes with reduced cost less than $UB - LB$ are inserted into column pool Φ'. Interested readers may refer to Parragh and Schmid (2013) for an analysis with respect to values of the parameters applied.

Note that in VNS during the add, remove and exchange operations we apply the PRS heuristic to produce feasible routes by inserting appropriately the returns of the vehicle to shelters to drop-off evacuees. In this case, Step 4 of the PRS heuristic is modified and the number of evacuees being dropped-off at the shelter is selected randomly in the range $[d^{min}, d^{max}]$ (with $d^{min} \leq d^{max}$) assuming that d^{min} is the minimum number of evacuees to drop-off in order to feasibly pick-up all evacuees waiting at the next pick-up location of the

sequence, and d^{max} is equal to the minimum value among the number of evacuees on board and the available capacity of the shelter. In case $d^{max} < d^{min}$ the number of evacuees being dropped-off at the shelter is set to d^{max}.

This change to Step 4 of the PRS implies that through VNS the algorithm explores a wider range of solution states, in terms of splitting the number of evacuees dropped off to the shelters, and, thus, may identify marginal solutions that could not reached in Phase I where new routes are generated only by LNS. Consequently, Phase I and LNS aim to identify solutions of sufficient quality, which are then refined during Phase II and VNS.

5.2.6 Hybrid framework

Algorithm 5.2 outlines the steps of the framework. An initial feasible solution is first constructed by the NN-BEP heuristic. This solution is then treated as the current solution for LNS in Phase 1 of the framework. All unique routes produced by LNS are inserted in column pool Φ' of RMP, and the routes of RMP's final solution are further improved by TSPI. Phase I is repeated until no improvement occurs after the solution of TSPI.

Subsequently, Phase II is applied for N^{iter} iterations. In each iteration the LRMP is solved to obtain the dual information used by VNS to identify columns with negative reduced cost to be added to column pool Φ'. RLMP and VNS are followed by LNS, RMP and TSPI in each iteration. After the N^{iter} iterations, the best solution identified (S_{best}) is returned as the solution of the problem.

Algorithm 5.2: Hybrid Column Generation

1. Produce S by **NN-BEP** heuristic and add routes $\varphi \in S$ in Φ'
2. Set $S = S'$, $obj(S_{best}) = \infty$ and $T = 1$
3. **while** $obj(S) < obj(S_{best})$
3.1. Set $S = S_{best}$
3.2. Improve S by **LNS** add new unique routes in Φ'
3.3. Solve **RMP** considering Φ' resulting to S_{best}
3.4. Run **TSPI** for all $\varphi \in S_{best}$, update S_{best} and Φ'
4. **while** $T \leq N^{iter}$
4.1. Solve **LRMP** considering Φ'
4.2. Run **VNS** for all $s \in S_{best}$ and add new routes in Φ'
4.3. Improve S_{best} by **LNS** add new unique routes in Φ'
4.4. Solve **RMP** considering Φ' resulting to S_{best}
4.5. Run **TSPI** for all $\varphi \in S_{best}$, update S_{best} and Φ'
4.6. Set $T = T + 1$
5. Solve **RMP** considering Φ' resulting to S_{best}
6. **Return** S_{best}

The inputs and outputs of each constituent element are provided in Table 5.1.

Table 5.1: Inputs and outputs of the constituent elements of the hybrid framework

Method	Inputs[1]	Outputs	Phases of application
NN-BEP	Basic Data[1]	Feasible solution	Before I
LNS	Feasible solution	Improved Feasible solution, New columns for Φ'	I and II
RMP	Column pool (routes) Φ'	Feasible solution	I and II
TSPI	Feasible solution	Improved Feasible solution	I and II
LRMP	Column pool (routes) Φ'	LB, Dual information	II
VNS	Columns of Φ' that participate in the relaxed solution	New columns for Φ' with reduce cost less thatn $UB - LB$	II

[1] The basic data of the problem is an input for all methods. The basic data include: the set of nodes, the set of arcs, the set of vehicles, the vehicles' capacity and the shelters' capacity

5.3 Tests and results

In this Section we study the efficiency of the hybrid framework (HCG) when applied to the two problem cases under investigation. First we examine the quality of the solutions by comparing those obtained by HCG with the optimal solutions obtained by solving the mathematical model of Section 5.1 by the gurobi MILP solver (Gurobi Optimization Inc., 2014) for problems of limited size. Then, we investigate the impact of the split pick-ups level (U) on the solution quality considering various demand scenarios. Finally, we examine the performance of the algorithm in large problems from the literature that resemble practical cases. Appendix A provides additional tests with respect to computational and convergence issues.

Thus, for the experimental investigation we used two types of problems with respect to size: (a) smaller scale problems to test the efficiency of HCG w.r.t. optimal solutions, and (b) larger instances provided by Prins *et al.* (2006) for the LRP, suitably modified to meet the characteristics of BEP and CEP.

For the studies of subsection 5.3.1 to 5.3.5 we used a random generator to produce the instances of smaller scale. The generator produces nodes uniformly in a square of [100,100] that represent the locations of the origin and destination depots, the pick-up locations and the shelters (or medical facilities). The number of vehicles, their capacity, the demand at every pick-up location, the number and the availability of the shelters are defined appropriately for every instance, as described in the respective Sections below. All tests and algorithms were implemented using Matlab R2013 on a PC equipped with a 2.00 GHz Intel Xeon E5-2620 and 32 GB of RAM.

5.3.1 BEP: Solution quality

In order to evaluate HCG in terms of solution quality we generated problems with $n \in \{3, 4, 5, 6, 7, 8, 15\}$ pick-up locations using the problem generator mentioned above. For each value of n, we generated 10 instances for a total number of 70 instances. Furthermore, for each instance the following holds: (a) the number of evacuees waiting at each pick-up location is produced by selecting a value from $\{1,2\}$ with equal probability, (b) two vehicles are considered, each capable of carrying three evacuees ($Q = 3$), and (c) there are two shelters each able to accommodate 62.5% of the casualties (P^1, P^2).

Each instance was solved both by the MILP solver and by HCG. We applied the latter 5 times for each instance in order to verify the consistency of the solutions produced by the proposed hybrid framework. A time limit of 6 hours was set for the MILP solver for $n \in \{3,4,5,6,7,8\}$, and 24 hours for $n = 15$.

Table 5.2 compares the solutions obtained from the MILP solver with those from the HCG framework.

- The first column presents the number of pick-up locations per problem case.
- The next two columns refer to the average value of the computational time to obtain the solution using the MILP solver and HCG, respectively. For the MILP solver the average per problem case was computed over the 10 instances. For the HCG, this average was computed considering the results of all five runs per instance for all 10 instances (i.e. 5x10 = 50 values). Note that the time limits imposed for the MILP solver were reached for all instances for $n \geq 7$, indicating the high complexity of the problem.
- The following three columns present the number of times HCG identified the optimal solution produced by the MILP solver (out of 50 runs for each n value) in terms of: (a) the objective value, (b) the evacuation time T_{evac} (lexicographically first objective), and (c) the total operation time T_{time} (lexicographically second objective).
- The following three columns present the average deviation from the optimal solutions of the MILP solver (i.e. for $n = 3, 4, 5, 6$) and only for the cases that HCG reached a different solution than that of the MILP solver. Again, each of the three columns refer to the objective, T_{evac} and T_{time}, respectively.
- The last three columns present the average deviation from the near-optimal solutions reached by the MILP solver within the time limits, i.e. for $n \in \{7, 8, 15\}$. Again, each of the three columns refer to the T_{evac} and T_{time}, respectively.

Table 5.2: Comparison of HCG performance with MILP for BEP

	CPU time (sec.)		Times best was reached[4]			Average deviation from optimal[5]			Average deviation from non-optimal[6]		
n	MILP	HCG[1]	Objective	T_{evac}	T_{time}	Objective	T_{evac}	T_{time}	Objective	T_{evac}	T_{time}
3	0.74	25.40	50/50	50/50	50/50	-	-	-	-	-	-
4	5.71	26.44	45/50	50/50	45/50	0.00%	-	1.91%	-	-	-
5	96.60	29.69	45/50	45/50	45/50	0.36%	0.36%	1.63%	-	-	-
6	1562.13	30.16	35/50	45/50	35/50	0.53%	1.70%	2.84%	-	-	-
7	21600.00[2]	34.65	-	-	-	-	-	-	0.64%	0.64%	1.05%
8	21600.00[2]	43.77	-	-	-	-	-	-	1.67%	1.67%	1.69%
15	86400.00[3]	123.35	-	-	-	-	-	-	0.68%	0.68%	3.72%

[1] Average computational time for the five runs of HCG, [2] Interrupted after 6 h limit,[3] Interrupted after 24 h limit, [4] Out of 50 runs, [5] For those cases for which the optimal was reached by the MILP and was not reached by HCG, [6]Compared to the solutions where MILP interrupted after the time limit

The results of Table 5.2 indicate that for these small scale problems HCG identifies the optimal solution for most of the cases. Even in the larger ones $n \in \{7, 8, 15\}$ the average deviation from the MILP solution is less than 2%, and for some cases HCG reaches better solutions that the MILP solver within the imposed limits (6 and 15 h).

5.3.2 BEP: Impact of split pick-ups on solution quality

In this Section we study the impact of the granularity of demand discretization, prescribed by the value of U, on the efficiency of the solution. As discussed in Section 5.1, the demand at each pick-up location is split in groups of size U (plus a group with the residual demand), and each vehicle may pick-up one or more of these groups from a pick-up location, depending on its residual capacity. Of course if $U = 1$, then the vehicle may pick-up as many evacuees as appropriate in order for the overall solution to be optimal. However, this increases significantly the complexity of the problem, preventing the efficient solution of large scale problems.

For this part of the study, we generated and solved five *basic* examples of relatively limited scale. Each comprises: (a) 10 pick-up locations, (b) two shelters with combined capacity of 120% of the total number of evacuees, and (c) two vehicles, each of capacity $Q = 10$. All locations were generated randomly. Subsequently, for each *basic* example we generated 9 instances with different number of evacuees $\bar{d}$ waiting per pick-up location; i.e. $\bar{d} \in \{2, \dots, 10\}$. Each of the 45 instances (5 basic problems x 9 instances per basic problem) was then considered under various U levels, $U \in \{1,2,4,6,8,10\}$. This was done as follows:

(1) If $\bar{d} > U$, then for every pick-up location, replicate locations were produced such that: (i) the number of evacuees per replicate location is equal to U (or less for one replicate), and (ii) the total number of evacuees of all replicate locations is equal to $\bar{d}$.

(2) If $\bar{d} \leq U$ no modification is required.

The case $U = 1$, in which every evacuee is considered as an individual pick-up location and all possible split pick-ups are considered, is used as the benchmark. For each problem (characterized by $\bar{d}$), and each case of U, five runs of the HCG were performed and the best solution was compared to the best solution identified among the five runs of the benchmark case with $U = 1$.

Table 5.3 presents the average value of the deviation from the benchmark case for each level of U. We refer to this average deviation as $\delta(U)$. Figure 5.4, presents the ratio of the computational time of the $U = 1$ case over the average computational time per level of U; we refer to this ratio as $\tau(U)$. Note that in Figure 5.4 only bars for cases in which $\bar{d} > U$ are shown, since for $\bar{d} \leq U$ there is no demand split. However, if $\bar{d}$ is odd, and since there is no corresponding U level, in the analysis and the Figure we have included the next higher level of U in order to include the case for which the demand is taken as a whole.

Table 5.3: Effect of U on the quality of the solution; i.e. the deviation $\delta(U)$ for $U \in \{2, 4, 6, 8, 10\}$ from the solution with $U = 1$

$\bar{d}\backslash U$	2	4	6	8	10
2	0%	-	-	-	-
3	0%	0%	-	-	-
4	**1%**	**4%**	-	-	-
5	0%	0%	0%	-	-
6	0%	0%	**27%**	-	-
7	0%	**2%**	**11%**	**11%**	-
8	0%	0%	0%	0%	-
9	0%	0%	0%	0%	0%
10	0%	0%	0%	0%	0%

We propose that the quality of the solution with $U > 1$ depends on how well the related packets of demand are packed in the bus capacity Q, without leaving significant residual capacity that may not be used. In fact, to test this observation for each $(\bar{d}, U)$ pair we solved a simple knapsack problem to identify the maximum utilization of the vehicle capacity. The results are shown in Table 5.4 and illustrate the theoretical empty capacity of the vehicles as a percent of the total capacity Q, if the demand $\bar{d}$ is split into packets of U.

By comparing the Tables 5.3 and 5.4 it is observed that for those $(\bar{d}, U)$ pairs for which the theoretical empty capacity is low, the quality of the solution is superior, as expected. Inversely in many cases for which the empty capacity is high (>10%) the quality if the solution tends to be inferior. Of course an exact match is not expected due to the routing constraints and the constraints related to the capacity of the shelters of BEP.

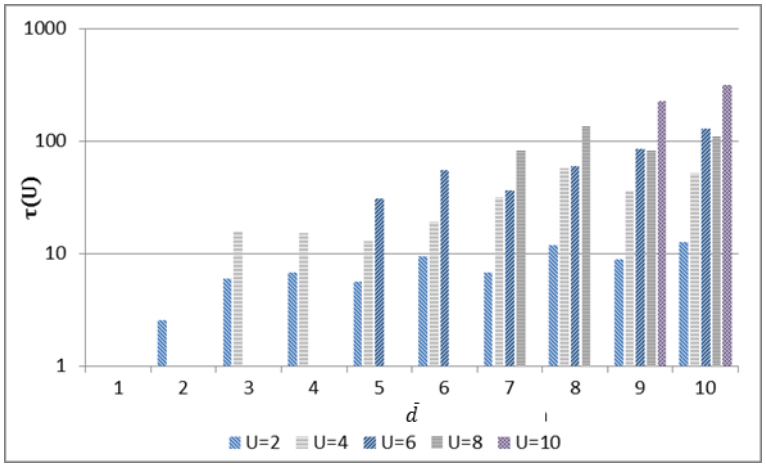


Figure 5.4: Effect of U on computational time: Ratio $\tau(U)$ of time required when $U = 1$ with time required when $U \in \{2, 4, 6, 8, 10\}$

Focusing to Table 5.3 we observe:

For $\bar{d} = 10$ all levels of U result in similar solution efficiency to that obtained under $U = 1$. This is attributed to the fact that an empty bus may pick-up the entire demand from a certain location. As expected, in terms of computational times the solution under $U = 10$ is obtained much faster (over 300 times faster) than the solution under $U = 1$.

For $\bar{d} = \{2,3,4,5\}$ and $\bar{d} = \{8,9\}$ the solutions under all levels of U perform similarly, that is under 5% deviation from the solution of the case with $U = 1$. The related computational time ratios are gradually increasing with increasing $\bar{d}$, ranging from under an order of magnitude to two orders of magnitude faster than the time required to solve the problem with $U = 1$. In these cases, fine discretization of the demand does not appear to be an efficient policy, since the improvement of the solution quality is minor at the expense of significant computational time.

Finally, for $\bar{d} = \{6,7\}$ it seems that the efficiency of the solutions for $U \geq 6$ [1] are significantly inferior to those with $U = 1$ or with $U = \{2,4\}$. That is, for cases $U \geq 6$ one bus may not

[1] Of course if $U \geq d$ the demand at the node is not split

serve two pick-up locations without an intermediate visit to a shelter. Specifically, for the case of $\bar{d} = 6$ and $U \geq 6$ the utilization of the vehicles in the best case is 60%.

Table 5.4: Theoretical empty capacity of the vehicles if the demand $\bar{d}$ is broken into groups of U

$\bar{d}\backslash U$	2	4	6	8	10
2	0%	-	-	-	-
3	0%	**10%**	-	-	-
4	0%	**20%**	-	-	-
5	0%	0%	0%	-	-
6	0%	0%	**40%**	-	-
7	0%	0%	0%	**30%**	-
8	0%	**20%**	0%	**20%**	-
9	0%	0%	**10%**	0%	**10%**
10	0%	0%	0%	0%	0%

Based on the above analysis, when selecting the appropriate discretization level U, one should aim to high vehicle utilization. To do so:

(1) If $\bar{d}$ is less than half of the bus capacity, i.e. $\bar{d} < 0.5Q$, then $U = Q$ results in high vehicle utilization and solutions of high quality, along with fast computational times.

(2) If $0.5Q \leq \bar{d} < 0.8Q$ then U may be set to approximately $0.4Q$ to result in high utilization, high quality solutions and reasonable computational times.

(3) If $\bar{d} \geq 0.8Q$ then U may be set to $0.8Q$ to achieve the above.

5.3.3 BEP: Efficiency on large scale problems

Based on the tests problems of Prins *et al.* (2006), we created eight large scale problems involving on average the transportation of 2,000 or 4,000 evacuees. Table 5.5 provides more information regarding these problems: The first three columns of the Table present the problem designation, the number of pick-up locations, and the number of shelters, as per Prins *et al.* (2006). The fourth column refers to the number of vehicles available, which is set to a value lower than the best published value of the Prins *et al.* LRP problems. This ensures that the vehicles will perform more than one trip from/to shelters. The fifth column is the bus capacity, which is reduced to values appropriate for a typical bus. The sixth column refers to the average total demand to be evacuated. We assume that at each pick-up location the number of waiting evacuees is generated with equal probability among the numbers $\{4,8,20,50\}$. The number of evacuees each shelter may accommodate is equal to 120% of the total demand divided by the number of shelters.

Five runs of HCG were performed to solve each problem of Table 5.5, assuming that $U = Q$ since $\bar{d} = 20.5$, $Q = 50$ and, thus, $\bar{d} < 0.5Q$. The results obtained are included in the last three columns of Table 5.5, which for the five runs present the average computational time, the best value of the objective function, and the average absolute deviation of each solution from the best identified.

Table 5.5: Parameters of the example problems and performance of HCG for the pre-disaster evacuation case

Problem Name	Pick-up Locations (n)	Shelters ($\|M\|$)	Vehicles ($\|K\|$)	Vehicle Capacity (Q)	Average total demand	Average CPU time (hr)	Best Obj. Value	Average Deviation
coord100-5-1	100	5	20	50	2000	2.17	113.10	3.77%
coord100-5-2	100	5	20	50	2000	2.68	86.00	3.37%
coord100-5-3	100	5	20	50	2000	2.13	98.70	0.97%
coord100-10-1	100	10	20	50	2000	5.45	178.00	2.92%
coord100-10-2	100	10	20	50	2000	3.00	102.30	5.73%
coord100-10-3	100	10	20	50	2000	8.91	246.20	3.95%
coord200-10-1	200	10	20	50	4000	9.21	207.30	2.91%
coord200-10-2	200	10	20	50	4000	7.73	211.80	1.62%

The relatively low value of the average deviation of the solutions per problem indicates that the HCG framework produces consistent results in large scale problems. Furthermore, all solutions are generated in reasonable times, if one considers that many evacuation plans are developed based on forecasts in advance of the upcoming emergency.

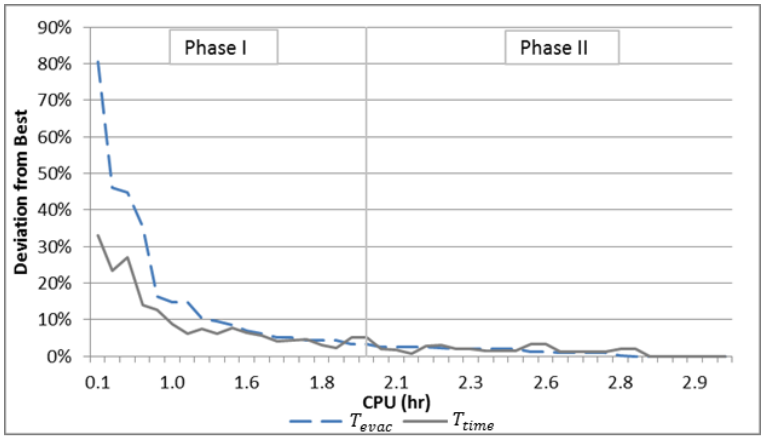


Figure 5.5: HCG convergence during the two framework phases for problem "coord100-10-2"

Figure 5.5 illustrates the way the proposed method converges to the solution for problem coord100-10-2; it appears that efficient solutions are reached early in the solution cycle (Phase I). Similarly, Fig. 5.6 shows the way the proposed method converges to the solution

for instance coord100-5-1, for which Phase II contributes significantly to the solution refinement. We have noticed that Phase II, during which the search widens, contributes to an improvement of the solution on average of about 5%, which is significant in terms of the overall efficiency of the method.

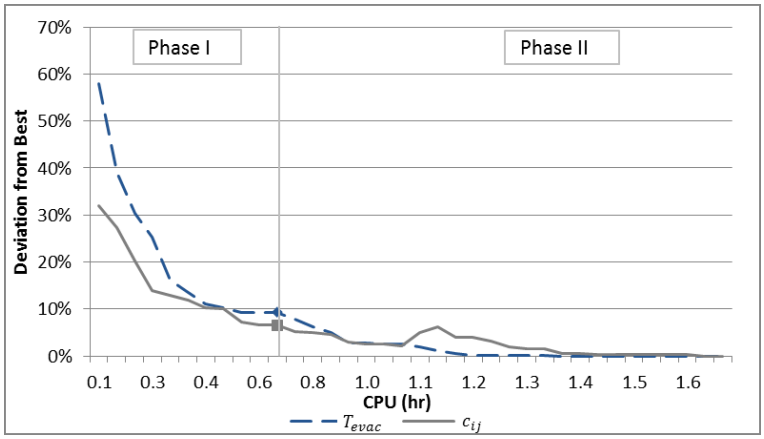


Figure 5.6: HCG convergence during the two framework phases for problem "coord100-5-1"

5.3.4 CEP: Solution quality

In order to test the quality of the HCG solutions, similarly to the study presented in subsection 5.3.1, we generated 70 random instances, 10 for each value of $n \in \{3,4,5,6,7,8,15\}$. The following parameters were used by the random problem generator: (a) one or two casualties per pick-up location chosen with equal probability, (b) two vehicles with $Q = 3$, and (c) two shelters each able to accommodate 62.5% of the casualties. In this case also, each instance was solved by HCG 5 times. For the MILP solver a time limit of 1 hour was set for $n \in \{3,4,5,6,7,8\}$ and 24 hours for $n = 15$.

Tables 5.6 and 5.7 compare the solutions obtained by the HCG framework to those obtained by the MILP solver for the problem cases CEP and CEP-MO, respectively. Table 5.6, similarly to Table 5.2, presents the average deviation values for each component of the CEP-MO objective function separately: T_{evac} is the total evacuation time and T_{a_h} is the sum of arrival times of the casualties to the medical facilities.

The results in Tables 5.6 and 5.7 indicate that for those cases for which MILP identified the optimal solution prior to the imposed time limit, i.e. $n \leq 4$ for CEP and $n \leq 5$ for CEP-MO, the HCG framework identified the optimal solution in the majority of runs for both CEP and CEP-MO. For those runs in which the optimal solution was not found, the average deviation from the optimal solution is less than 3%. For the cases that MILP was stopped due to the

time limit imposed, i.e. all CEP instances with $n \geq 5$, and all CEP-MO instances with $n \geq 6$, HCG identified more efficient solutions (in terms of the objective function), for all cases except for the case $n = 15$ of CEP-MO (in which the average deviation from the MILP solutions is less than 1%).

Table 5.6: Comparison of HCG and MILP performance for CEP

	CPU time (sec.)		Times optimal reached[4]	Average deviation from optimal[5]	Average deviation from non-optimal[6]
n	MILP	HCG[1]			
3	6.07	37.92	45/50	2.80%	-
4	122.72	49.91	50/50	0.00%	-
5	3600.00[2]	26.97	-	-	-0.23%
6	3600.00[2]	35.96	-	-	-1.35%
7	3600.00[2]	52.46	-	-	-0.93%
8	3600.00[2]	64.82	-	-	-1.99%
15	86400.00[3]	275.96	-	-	-0.78%

[1] Average computational time for the five runs of HCG, [2] Interrupted after 1 h limit, [3] Interrupted after 24 h limit, [4]For cases for which MILP reached the optimal solution, [5] For those cases for which MILP reached the optimal solution and HCG did not, [6]Compared to the solutions in which MILP was interrupted due to the time limit

Furthermore, it appears that solving the problems under the CEP-MO objective function requires significantly less computational time than under the CEP objective function. From all tests it seems that HCG deals with CEP (and CEP-MO) problems in a very effective manner.

Table 5.7: Comparison of HCG and MILP performance for CEP-MO

	CPU time (sec.)		Times best was reached[4]			Average deviation from optimal[5]			Average deviation from non-optimal[6]		
n	MILP	HCG	Objective	T_{evac}	T_{a_h}	Objective	T_{evac}	T_{a_h}	Objective	T_{evac}	T_{a_h}
3	3.37	39.07	50/50	50/50	50/50	0.00%	0.00%	0.00%	-	-	-
4	24.14	52.15	50//50	50/50	50/50	0.00%	0.00%	0.00%	-	-	-
5	508.36	29.79	45/50	45/50	45/50	3.03%	3.3%	1.40%	-	-	-
6	3600.00	40.16	-	-	-	-	-	-	-0.91%	-2.30%	-6.08%
7	3600.00	57.03	-	-	-	-	-	-	-1.33%	-2.36%	2.23%
8	3600.00	72.15	-	-	-	-	-	-	-5.47%	-5.47%	-5.48%
15	86400.00	337.73	-	-	-	-	-	-	0.52%	0.52%	10.70%

[1] Average computational time for the five runs of HCG, [2] Interrupted after 1 h limit, [3] Interrupted after 24 h limit, [4]Out of 50 runs, [5] For those cases for which the MILP reached the optimal solution and HCG did not, [6]Compared to the solutions in which MILP was interrupted due to the time limit

5.3.5 CEP: Efficiency of the algorithm on large scale problems

In order to further analyze the efficiency of HCG in instances of practical scale, as well as investigate the impact of the two alternative objective functions, we employed the problems presented in Table 5.8. These instances vary w.r.t. the number of pick-up locations from 10 to 50, and w.r.t. the number of casualties from 20 to 100. The smaller size cases (designated as "*R*" instances) were generated by the random problem generator described at the start of Section 5.3. For problems of moderate and larger sizes (designated as "*coord"*), we used the location data of the Prins *et al.* (2006); however, as before, in order to force the vehicles to perform multiple trips we reduced their number and the number of medical facilities (depots). For both *R* and *coord* instances, the demand at each pick-up location was generated by a random selection from $\{1,2,3\}$ with equal probability among the three values, the capacity of the vehicle was set to four, $Q = 4$, and the number of casualties each medical facility may accommodate is equal to 120% of the total number of casualties divided by the number of medical facilities.

Table 5.8: Parameters of casualties evacuation instances

Instance Name	n	Medical Facilities	Vehicles	Q	Average Total Casualties
R10-2-2	10	2	2	4	20
R10-2-2	10	2	2	4	20
coord20-5-1	20	2	5	4	40
coord20-5-2	20	2	5	4	40
coord50-5-1	50	2	5	4	100
coord50-5-2	50	3	5	4	100
coord50-5-3	50	3	5	4	100

Table 5.9 presents the results of HCG when applied on the CEP instances of Table 5.8. Columns two and three provide the average computational times of HCG. Columns four and five refer to the average value of the sum of casualty arrival times to the medical facilities. The CEP column corresponds to the instances under the CEP objective, while the CEP-MO column corresponds to the instances under the objective of CEP-MO. As expected, with respect to the sum of arrival times, in all cases the CEP results are superior. The corresponding difference is significant, since the total arrival times of CEP are on average 13% less than those of CEP-MO.

Similarly, columns six and seven of Table 5.9 refer to the total evacuation time obtained under the two objective functions. As expected, the CEP-MO results are superior. In this case, in the solutions of CEP the evacuation operation is completed on average 22% slower than the solutions obtained by CEP-MO.

Finally, column eight of Table 5.9 refers to the average deviation of the HCG framework from the best identified solution for the CEP objective, while the last column refers to the average deviation of the HCG framework from the best identified solution for the CEP-MO objective. The average deviation values of the solutions per problem and per objective case are relatively low, similarly to the BEP cases, indicating the consistency of the solutions obtained by the HCG framework.

Table 5.9: HCG performance on casualty evacuation problems

	CPU (sec)		Total casualties' arrival times		Total evacuation time		Average deviation from best	
Instance Name	CEP	CEP-MO	CEP	CEP-MO	CEP	CEP-MO	CEP	CEP-MO
R10-2-1	90.65	110.60	3158.14	3588.66	269.58	243.18	0.11%	0.70%
R10-2-2	93.39	123.57	3756.56	3930.16	367.44	304.50	0.28%	0.46%
coord20-5-1	364.66	294.54	1812.86	1994.28	111.88	77.54	0.38%	2.43%
coord20-5-2	343.98	327.11	2283.50	2642.36	121.72	97.80	0.90%	0.41%
coord50-5-1	4214.12	4183.68	13126.94	15807.10	311.26	273.34	1.25%	3.85%
coord50-5-2	4027.81	3823.13	14024.06	16513.62	336.50	292.92	1.48%	2.10%
coord50-5-3	4256.29	3885.86	8747.80	10656.48	225.92	184.10	3.11%	4.13%

In order to further investigate the solution characteristics under the two objective functions, we examined the progress of transporting the casualties to the medical facilities. Figure 5.7 presents the percent of casualties transported to the medical facilities during the operation as a function of normalized operation time. The basis of normalization is defined by the maximum evacuation time per instance, i.e. the solution of the instance with respect to the CEP objective.

The result of Fig. 5.7 is the average of all instances of Table 5.9. According to the Figure, under the CEP objective, most casualties are transported to the medical facilities early during the operation, while CEP-MO manages to complete the entire operation earlier. Note that the difference displayed in the graph (broken line) is the percentage difference of the number of

casualties transported to medical facilities using the CEP objective versus the respective number obtained by the CEP-MO objective.

This difference is important, and one should consider the emergency of the transportation process under consideration before selecting between the CEP or CEP-MO objectives. Considering cases where there is a definite and tight deadline (e.g. the expiration of an agreed short term truce or ceasefire), then CEP-MO seems more appropriate. On the contrary, if there is no such a deadline, then CEP may be more appropriate, since it shortens the average transport time of the causalities.

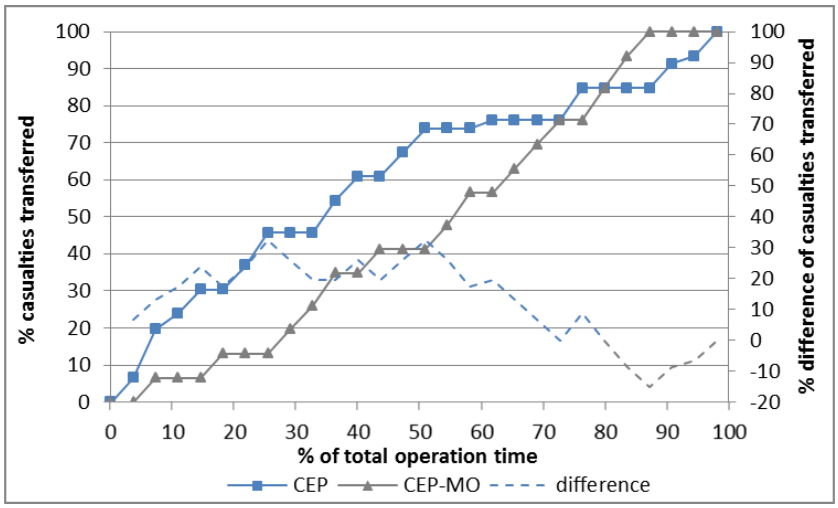


Figure 5.7: Percentage of casualties transported to the medical facilities during the transport operation

Chapter 6. Application of paratransit transport aspects to logistics

Some features of the problems studied in Chapters 3 to 5, as well as the related solution approaches, may be leveraged upon in special logistics applications. In particular the features under consideration concern: (a) predefined, strict route sequences similar to those of public bus services, and (b) limitations on the capacity of warehouses or inventory replenishment facilities, similar to the capacity limitations of shelters during evacuation operations.

Specifically in this Chapter we focus on the case of a single vehicle distributing a single product and using *two* warehouses (or depots) for load replenishment. Three cases with respect to the inventory level of the two warehouses are investigated, and efficient solutions for problems of large size are developed to efficiently address them.

To do so, we extend previous work on the problem of Vehicle Routing with Warehouse Returns (Tsirimpas *et al*., 2007) in three ways: (i) We consider two warehouses instead of one. (ii) We consider limited warehouse inventory instead of infinite. (iii) New solution methods are developed, since dynamic programming in no longer an efficient method to address the enhanced state space of the problems.

To deal with the progressively increasing state space of the above problem cases, we develop new efficient Labeling Algorithms. Furthermore, for the most complex of the three cases we have developed a new heuristic that enhances the performance of the labeling algorithm (see also Chapter 2), by decomposing the monolithic problem into an appropriate number of smaller sub-problems, the solution of which is coordinated by a suitable integer program. We study the efficiency of these algorithms by solving a wide range of problem instances, and by comparing the results with those obtained from a state-of-the-art MILP solver.

6.1 Description of the problem variants under consideration

The Vehicle Routing with Warehouse Returns Problem, as treated by Tsirimpas *et al*. (2007), considers the case of a single vehicle serving clients with a predefined sequence and known product demand. Due to capacity constraints the vehicle must return one or multiple times to the (single) warehouse for replenishment. The problem is to find the optimal visits to the warehouse so that the total distance traveled is the minimum possible and all clients are served. Tsirimpas *et al*. (2007) have studied three related cases and used dynamic programming to identify the optimal route for each. Yang *et al*. (2000), Tatarakis and Minis (2009), and Minis and Tatarakis (2011) consider the stochastic case of this problem with random client demand. In the first work, Yang *et al*. (2000) investigate single and multiple-vehicle problems distributing a single product, and minimize the total routing cost and the probability of route failure. In their first paper, Tatarakis and Minis (2009) consider the delivery of multiple products by a single vehicle to clients with random demand. In their second paper, they studied the pick-up and delivery case with random demand using a single vehicle. Bianchi *et al*. (2006) studied the performance of metaheuristics for solving the Yang *et al*. (2000) problem.

As mentioned above, in this Chapter we focus on the case of a single vehicle distributing a single product, and using *two* warehouses (or depots) for load replenishment, which is the new element in the problem under study. We consider three variants of increasing state space.

In the first variant (UU) both warehouses have Unlimited product availability (inventory); in the second variant (UL), one warehouse has Unlimited and the other Limited inventory, while in the third variant (LL) both warehouses have Limited product inventory. Note that all variants refer to a certain operating horizon (e.g. one day), and thus we may assume unlimited (or limited) product availability within this horizon, depending on the product and the related inventory turns. The limited warehouse inventory is another new element of the current problem. Both these elements are significant departures from previous literature necessitating new models and solution approaches.

Note that in UU, after each visit to a client, the decision is whether to return -or not- to the nearest warehouse facility for full replenishment. In UL or LL, since there exists at least one warehouse with limited inventory, the decisions to be made after each visit to a client are (i) return -or not- to a warehouse facility, (ii) which facility to return to, and (iii) how much product to load, in case the selected facility is the one with the limited inventory.

A similar problem in the literature is the Vehicle Routing Problem with Satellite Facilities (VRPSF) that has already been described in Chapter 5. The problems addressed here have interesting industrial applications. For instance, material handling systems in a manufacturing shop often operate along a fixed pathway that connects material warehouses with workcenters. In addition to the main pathway connecting the workcenters in a predefined sequence, there are spurs connecting each workcenter with material warehouses, allowing direct return of the material handling device (e.g. Automated Guided Vehicle, AGV). This setting is exactly the one addressed in this Chapter, and the objective is to minimize routing time in order to maximize productivity.

Related applications may be also found in the following cases:

- Transportation of bulk routine materials in a healthcare facility e.g. meals, linen, medical supplies and waste. In this case, staff (or AGVs) move along predefined sequences, usually room after room, and must return to the warehouse to reload or unload the transportation carts. Improving this procedure may offer cost benefits and on time delivery improvements
- Breakdown of a vehicle during delivery operations. In this case an active vehicle is tasked to complete the delivery plan of the incapacitated vehicle, in addition to its own delivery plan. The active vehicle may use both the warehouse and the incapacitated vehicle as replenishment warehouses, much like in problem UL above; see Mamasis *et al*. (2013).

Problems involving the management of distribution systems in the case of vehicle breakdowns have been also studied by Li *et al*. (2009a) and Li *et al*. (2009b) and Mu *et al*. (2011) although in this studies

- Battlefield situations, in which vehicles may collect crews or functional equipment from incapacitated military vehicles. In this case, the sequence of visits may be pre-determined, and the collection vehicle may need to drop the crews or equipment at more than one locations depending on the load collected from the incapacitated vehicles. In this case as well, minimizing the distance travelled, increases system efficiency with related benefits to battlefield effectiveness.

In some practical situations the number of clients is large, resulting to problems of increased complexity. For example, in the health care industry case mentioned above, planning of *multiple* delivery cycles may be necessary, thus, increasing the number of clients. Large-scale instances may also occur in component placement during product assembly (e.g. in the case of printed circuit boards). In this case, the products (clients) move on a conveyor and stop at a location related for the component placement machine. The machine may place (or assemble) several parts in one product. Moreover, the number of products in a batch may be quite large (e.g. 100 to 1000). The machines may pick-up several components at once from more than one picking points (warehouses) and place them in one or more products waiting at the appropriate locations. In this case, the aim is to minimize the total distance (total operation time), and, thus, increase the machine's productivity.

6.2 Modeling the three problem cases

Consider the following notation:

- The vertices in set $C = \{1, \dots, n\}$ correspond to the clients of the network
- Vertex 0 corresponds to the starting position of the vehicle, and vertex $n + 1$ corresponds to the ending position of the vehicle
- Vertices D^1 and D^2 correspond to the two available warehouses.

Consider now the directed graph $G(V, A)$, where $V = C \cup \{0, D^1, D^2, n + 1\}$ is the vertex set and $A = \{(i, i + 1), (i, D^1), (i, D^2), (D^1, D^2), (D^2, D^1), (D^1, i + 1), (D^2, i + 1): i \in V \setminus \{n + 1\}\}$ is the arc set. Figure 6.1 illustrates the above network. The vehicle must serve all clients with a predefined sequence $1, \dots, n$. Each client has a known positive demand $r_i, \forall i \in C$ and $c_{ij} \geq 0$ denotes the cost (distance) of traversing arc (i, j). The initial product inventories are

Π^1 and Π^2, respectively, and the vehicle capacity is Q product units. The demand of vertex 0 is $r_0 = 0$, and the load on the vehicle is assumed to be known.

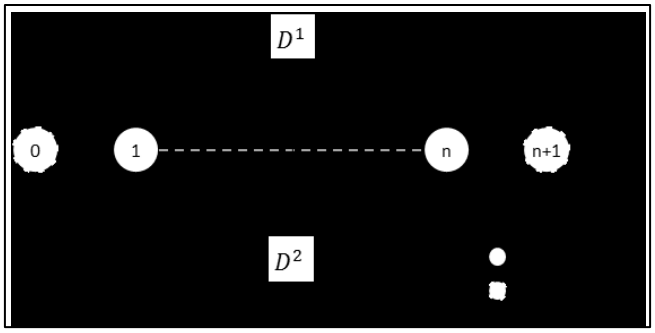


Figure 6.1: The network of the problem

The solution seeks a vehicle route serving all clients at the pre-defined sequence with minimal total distance.

6.2.1 Modeling problem case UU

In this case the two warehouses have unlimited product availability ($\Pi^1, \Pi^2 \to \infty$) and, as a result, visiting two warehouse facilities in sequence is not appropriate, since the vehicle can be fully loaded upon each visit to a single warehouse facility. Thus, the decision to be made after serving each client i is whether to continue to the next client $i+1$ or visit one of the two warehouses. The decision variables of the model are related to: (a) which arcs participate in the solution, (b) the product quantities loaded onto the vehicle at the two warehouses, and (c) the load of the vehicle leaving client i. Based on the above, the set of arcs (A') is a subset of A as follows:

$$A' = \{(i, i+1), (i, D^1), (i, D^2), (D^1, i+1), (D^2, i+1)\}: i \in V \setminus \{n+1\}\}, A' \subseteq A$$

The decision variables related to the network arcs are the following:

- $x_{i,i+1}, i = 0, \dots, n$, is related to arc $(i, i+1)$. It corresponds to a direct trip from client i to client $i+1$
- $x_i^D, i = 0, \dots, n$, is related to arc set $\{(i, D^1), (D^1, i+1)\}$ if $c_{iD^1} + c_{D^1 i+1} \leq c_{iD^2} + c_{D^2 i+1}$, or to arc set $\{(i, D^2), (D^2, i+1)\}$ otherwise. This corresponds to a trip from client i to client $i+1$ through the warehouse that causes the minimum route deviation.

Each of the above variables assumes the value 1 if the corresponding arc is traversed and 0 otherwise. As mentioned above the cost related to $x_{i,i+1}$ is $c_{i,i+1}$, while the cost related to related costs to x_i^D is:

$$\tilde{c}_i^D = \begin{cases} c_{iD^1} + c_{D^1 i+1}, & c_{iD^1} + c_{D^1 i+1} \leq c_{iD^2} + c_{D^2 i+1} \\ c_{iD^2} + c_{D^2 i+1}, & otherwise \end{cases}$$

Variables L_i define the product quantity loaded onto the vehicle during a visit to a warehouse and the load of the vehicle leaving client i is denoted as q_i. The mathematical model for **problem case UU** is as follows:

$$minTC = \sum_{i=0}^{n-1} \left(c_{i,i+1} x_{i,i+1} + \tilde{c}_i^D x_i^D \right) + c_{n,n+1} \tag{6.1}$$

Subject to:

$$x_{i,i+1} + x_i^D = 1, \; i = 0, \dots, n \tag{6.2}$$

$$q_i = q_{i-1} - r_i + L_{i-1}, \; i = 1, \dots, n \tag{6.3}$$

$$q_i + r_i \leq Q, \; i = 0, \dots, n \tag{6.4}$$

$$L_0 \geq Q - M \times (1 - x_0^D) \tag{6.5}$$

$$L_i \geq Q - q_{i-1} - M \times (1 - x_i^D), \; i = 1, \dots, n \tag{6.6}$$

$$L_i \leq M \times x_i^D, \; i = 0, \dots, n \tag{6.7}$$

$$q_i, L_i \in \mathbb{N}_0, \; i = 0, \dots, n \tag{6.8}$$

$$x_{i,i+1}, x_i^D \in \{0,1\}, \; i = 0, \dots, n \tag{6.9}$$

Equation (6.2) ensures that the vehicle departing from a client will drive to the next client directly, or through visiting a warehouse. Equation (6.3) provides the change of the load carried by the vehicle upon visiting client i. This load reduces by the demand of client i, and increases by the product units transferred onto the vehicle during its possible visit to one of the warehouses. If the vehicle does not visit a warehouse, then the respective quantity becomes zero through Constraints (6.7) – see also below. Constraint (6.4) ensures that the vehicle capacity cannot be exceeded at any point along the route. Constraints (6.4) to (6.7) ensure that every time the vehicle visits one of the two available warehouses, it will be filled up to its capacity. The large $M \gg 1$ is used to linearize the model by avoiding multiplication between the arc variables and the quantity to be loaded onto the vehicle. Equation (6.8) defines that all variables relating to load quantities belong to the set of non-negative integers. Finally, Equation (6.9) states that the arc set decision variables are binary.

6.2.2 Modeling problem case LL

In this problem case, the decision variables of the model are related to: (a) the arcs sets, (b) the product quantities loaded onto the vehicle at the two warehouses, (c) the load of the vehicle leaving client i, and (d) the inventory levels of the two warehouses:

- $x_{i,i+1}, i = 0, \dots, n$, is related to arc set $\{(i, i+1)\}$
- $x_i^1, i = 0, \dots, n$, is related to arc set $\{(i, D^1), (D^1, i+1)\}$
- $x_i^2, i = 0, \dots, n$, is related to arc set $\{(i, D^2), (D^2, i+1)\}$
- $x_i^{12}, i = 0, \dots, n$, is related to arc set $\{(i, D^1), (D^1, D^2), (D^2, i+1)\}$ if $c_{iD^1} + c_{D^1D^2} + c_{D^2i+1} \leq c_{iD^2} + c_{D^2D^1} + c_{D^1i+1}$ otherwise to arc set $\{(i, D^2), (D^2, D^1), (D^1, i+1)\}$

Each of the above variables assumes the value 1 if the corresponding arc is traversed and 0 otherwise.

The costs related to these variables are $c_{i,i+1}$ and:

- $\tilde{c}_i^1 = c_{iD^1} + c_{D^1i+1}$ and $\tilde{c}_i^2 = c_{iD^2} + c_{D^2i+1}$
- $\tilde{c}_i^{12} = \begin{cases} c_{iD^1} + c_{D^1D^2} + c_{D^2i+1}, & c_{iD^1} + c_{D^1D^2} + c_{D_i^2i+1} \leq c_{iD^2} + c_{D^2D^1} + c_{D^1i+1} \\ c_{iD^2} + c_{D^2D^1} + c_{D^1i+1}, & otherwise \end{cases}$

Variables $d_i^1, d_i^2 \in \mathbb{N}_0$ represent the quantity of products transferred onto the vehicle while the latter visits one or both warehouses D^1 and D^2. The inventory levels of the two warehouses, denoted as P_i^1 and P_i^2 respectively, are updated just before serving client $i+1$. That is, after serving client i, in case the vehicle moves directly from i to $i+1$, or after visiting the warehouse between client i and $i+1$.

The objective function of **problem LL** is defined as follows:

$$minTC = \sum_{i=0}^{n-1} \left(c_{i,i+1} x_{i,i+1} + \tilde{c}_i^1 x_i^1 + \tilde{c}_i^2 x_i^2 + \tilde{c}_i^{12} x_i^{12} \right) + c_{n,n+1} \tag{6.10}$$

Subject to:

$$x_{i,i+1} + x_i^1 + x_i^2 + x_i^{12} = 1, \quad i = 0, \dots, n \tag{6.11}$$

$$q_i = q_{i-1} - r_i + d_{i-1}^1 + d_{i-1}^2, \quad i = 1, \dots, n \tag{6.12}$$

$$q_i + r_i \leq Q, \quad i = 0, \dots, n \tag{6.13}$$

$$P_0^1 + d_0^1 = \Pi^1 \tag{6.14}$$

$$P_i^1 + d_i^1 = P_{i-1}^1, \quad i = 1, \dots, n \tag{6.15}$$

$$P_0^2 + d_0^2 = \Pi^2 \tag{6.16}$$

$$P_i^2 + d_i^2 = P_{i-1}^2, \quad i = 1, \dots, n \tag{6.17}$$

$$d_i^1 \leq M(x_i^1 + x_i^{12}), \quad i = 0, \dots, n \tag{6.18}$$

$$d_i^2 \leq M(x_i^2 + x_i^{12}), \quad i = 0, \dots, n \tag{6.19}$$

$$x_{i,i+1}, x_i^1, x_i^2, x_i^{12} \in \{0,1\}, \quad i = 0, \dots, n \tag{6.20}$$

$$P_i^1, P_i^2, q_i, d_i^1, d_i^2 \in \mathbb{N}_0, \quad i = 0, \dots, n \tag{6.21}$$

Equation (6.11) ensures that the vehicle departing from a client will be directed towards the next client directly, or through visiting either warehouse 1, or warehouse 2, or both warehouses (covering the minimum distance). Equation (6.12) provides the change of the load carried by the vehicle upon visiting client i. In this problem case, this load reduces by the demand of client i, and increases by the product units transferred onto the vehicle during its possible visit(s) to the warehouse(s). If the vehicle does not visit a warehouse, then the respective quantity becomes zero through Constraints (6.18) and (6.19) – see also below. Constraint (6.13) ensures that the vehicle capacity cannot be exceeded at any point along the route. Equations (6.14) and (6.16) define the inventory of the two warehouses after the vehicle visits one (or both) of them, immediately following its departure from the starting position. Note that the vehicle starts empty and it needs to be loaded prior to serving the first client. Equations (6.15) and (6.17) define the change in inventory each time a vehicle visits the warehouse facility. Since $M \gg 1$, Constraints (6.18) and (6.19) ensure that if the vehicle does not visit a warehouse, the quantity "loaded" onto the vehicle is zero. Equation (6.20) states that the arc set decision variables are binary. Finally, all variables relating to load quantities belong to the set of non-negative integers.

6.2.3 Modeling problem case UL

We assume that the second warehouse (D^2) is the one with the finite product inventory. The latter increases the complexity compared to the UU case. All decision variables associated with D^1 can be excluded from the model. Similarly to the UU case the vehicle will never visit sequentially the two warehouses, thus the arc set for this case also is A'. The detailed model is presented in Dikas *et al.* (2012).

6.3 Solution methods

We propose the following methods to solve the above problems: (a) A dynamic programming algorithm, *dp(UU)*, which provides the optimal solution for problem UU, but it is not efficient for UL and LL. (b) A new labeling algorithm, which solves the UU and UL problems to optimality, and provides good, near-optimal solutions for LL. Hereafter, as the variations of the labeling algorithm to the three problem cases are designated as $la(UU)$, $la(LU)$and $la(LL)$. (c) A new partitioning heuristic, which we have developed for the most complex case LL in order to enhance the efficiency of $la(LL)$. The time efficiency of all methods is benchmarked against the efficiency of a state-of-the-art MILP solver. Below we present *dp(UU)*, the labeling algorithm for all three problems, and the partitioning heuristic we developed for LL.

6.3.1 Dynamic programming algorithm – *dp(UU)*

Let's assume that the minimum distance $C_i(q_i)$ from client i until the end point $(n+1)$ is a function of the load q_i of the vehicle as it departs from client i. Then the *dp(UU)* Equations are formulated as follows:

If $i = n$, then:

$$C_i(q_i) = c_{n,n+1}, \qquad q_i = 0, \dots, Q \tag{6.22}$$

For clients $i = n-1, n-2, \dots, 1$ and for vehicle load $q_i = 0, \dots, Q$:

$$C_i(q_i) = min\begin{cases} c_{i,i+1} + C_{i+1}(q_i - r_{i+1}) \\ c_{i,D^1} + c_{D^1,i+1} + C_{i+1}(Q - r_{i+1}), \\ c_{i,D^2} + c_{D^2,i+1} + C_{i+1}(Q - r_{i+1}) \end{cases} \qquad q_i \geq r_{i+1} \tag{6.23 (a)}$$

$$C_i(q_i) = min\begin{cases} c_{i,D^1} + c_{D^1,i+1} + C_{i+1}(Q - r_{i+1}) \\ c_{i,D^2} + c_{D^2,i+1} + C_{i+1}(Q - r_{i+1})' \end{cases} \qquad q_i < r_{i+1} \tag{6.23 (b)}$$

For the starting position $i = 0$, and $q_0 = 0$:

$$C_0(q_0) = min\begin{cases} c_{0,D^1} + c_{D^1,1} + C_1(Q - r_1) \\ c_{0,D^2} + c_{D^2,1} + C_1(Q - r_1) \end{cases} \tag{6.24}$$

Equation (6.23a) ensures that if the vehicle load is adequate to satisfy the demand of the next client, then $C_i(q_i)$ will be the minimum distance considering all three possible paths the vehicle may follow. In Eq. (6.23b), the path that represents the direct visit from client i to

$i+1$ is excluded due to Eq. (6.4). The minimum distance $minTC = C_0(q_0)$ is provided by Eq. (6.24).

6.3.2 Labeling algorithm - *la(UU)*

The labeling algorithm creates partial paths iteratively; that is, paths that begin from the starting position and reach each client according to the predefined client sequence. Each partial path incident to a client i can be extended to the next client $i+1$ by: (a) direct visit to the next client, (b) visit through D^1, or (c) visit through D^2. The algorithm extends all partial paths reaching each client, determines the *quality* of these paths, and discards those of inferior *quality*. The paths of inferior *quality* are identified by applying the, so called, dominance rules. The procedure continues until the ending position $n+1$. The complete path (from the starting to the ending position) with the minimum total cost is selected as the optimal one. As described in Chapter 2 the labeling algorithm is an extension of the classic Dijkstra's algorithm for the directed single-source shortest path problem (Dijkstra, 1959). It has been applied to a multitude of problems, such as the VRP with time windows (Desrochers *et al.,* 1992), in airline crew pairing, Lavoie *et al*. (1988), and in flight crew scheduling, Graves *et al*. (1993). More recent works for the Elementary Shortest Path Problem with Resource Constraints (ESPPRC) include those of Feillet *et al.* (2004) and Chabrier (2006), who focus on the importance of the dominance rules.

Let $K = (0,1,2,\dots,n+1)$ be the predefined sequence of serving the clients. Also let Z_i be the set that contains all partial paths reaching client (node) $i \in K$. Each partial path $z \in Z_i$ is characterized by:

- $\bar{C}_z$: the total distance from the starting position (0) to client i, the last node of partial path z
- $\bar{q}_z$: the load of the vehicle after serving client i.

The steps of the algorithm are as follows:

Step 1: Start from the first node of the sequence ($i = 0$) with the initial, trivial, partial path $z \in Z_0$, where:

$$\bar{C}_z = 0, \bar{q}_z = 0, \qquad z \in Z_0$$

Step 2: Set $i = i + 1$

Step 3: [*Partial path extension*]: For each partial path $z, z \in Z_{i-1}$, connect node (client) $i-1$ to i, by the two possible ways and create, if feasible, two new partial paths $z'_1, z'_2 \in Z_i$. Overall, $2 \times |Z_{i-1}|$ new paths will be created. In case a partial

path exists for which the load of the vehicle is adequate to satisfy all clients until the end of the client sequence, $\bar{q}_z \geq \sum_{y=i}^{n+1} r_y$, then only partial path z'_1 is constructed by visiting node i directly, since any visit to a warehouse would unnecessarily increase the distance travelled by the vehicle.

Step 4: [*Computation of the path variables*]: For each newly created partial path ($z' \in Z_i$), the following are computed based on the corresponding variables of the parent partial path $z \in Z_{i-1}$ (note that index $k = 1,2$ of z'_k has been dropped for convenience).

i. The total distance:

$$\bar{C}_{z'} = \bar{C}_z + c_{i-1,i} x_{i-1,i} + \tilde{c}^D_{i-1} x^D_{i-1}$$

ii. The load $\bar{q}_{z'}$ of the vehicle after serving client i is calculated based on the visits of the vehicle to a warehouse facility as follows:

$$\bar{q}_{z'} = \begin{cases} \bar{q}_z - r_i, & x^D_{i-1} = 0 \quad \text{(6.25) (a)} \\ Q - r_i, & x^D_{i-1} = 1 \quad \text{(6.25) (b)} \end{cases}$$

According to these Equations, during the direct visit to client i the vehicle load decreases by the demand of client i, but when a vehicle visits a warehouse facility (D^1 or D^2), then it is loaded up to its capacity.

Step 5: [Discarding of partial paths]: If any of the partial paths in set Z_i is infeasible due to Constraints (6.3) to (6.7), then it is discarded (pruned). For the remaining paths, we discard dominated paths according to the following relationships:

i. $\bar{q}_{z'_1} \geq \bar{q}_{z'_2}$ (6.26)

ii. $\bar{C}_{z'_1} \leq \bar{C}_{z'_2}$ (6.27)

Partial path z'_1 dominates partial path z'_2 if both Inequalities (6.26) and (6.27) hold. If Inequality (6.26) is satisfied, then the vehicle may visit a warehouse fewer -or at most the same- number of times for replenishment, and, thus, may travel a shorter (or at most equal) distance from client i to the end of the route. If Inequality (6.27) is satisfied as well, then at least one complete path created from partial path z'_1 will correspond to a lower -or at least equal- distance as compared to *all* complete paths created from partial path z'_2. Thus, partial path z'_2 can be discarded.

Step 6: Steps 2, 3, 4 and 5 are repeated until $i = n + 1$

Step 7: From the set of paths Z_{n+1}, the one with the minimum distance is the optimal route to satisfy all clients with $minTC = \min_{z \in Z_{n+1}} \bar{C}_z$.

6.3.3 Labeling algorithm - *la(UL)*

As in the UU case, let Z_i be the set of all partial paths reaching client (node) $i \in K$. As before, each partial path $z \in Z_i$ is characterized by $\bar{C}_z$, and $\bar{q}_z$; additionally it is also characterized by the inventory level of the warehouse with limited inventory, that is:

- $\bar{P}_z^2$: the inventory in D^2 after serving the last client of partial path z

At each visit to D^2, the appropriate product quantity d_i^2 should be loaded onto the vehicle. Note that d_i^2 is a decision variable of problem UL, and its value is not known at the time of the visit to D^2. This difficulty is handled in dynamic programming by examining all possible values of d_i^2 in order to obtain the optimal solution. In *la(UL)*, in order to reduce the number of paths, we first consider that the maximum feasible quantity of products is loaded onto the vehicle when it visits D^2. As the path(s) extend downstream, we update this quantity as well as the inventory of D^2 by returning any unused quantity back to the inventory of this facility. That is, each time after the visit to D^2 the vehicle visits D^1 (obviously after serving some intermediate clients), no extra load is needed on the vehicle upon the visit to D^1, since the vehicle can load up to capacity in D^1. Thus, any extra load remaining on the vehicle upon its visit to D^1 should be returned to D^2 by updating its inventory.

In addition, we use the following variable in the dominance criteria:

- H_z^2: the maximum possible inventory in D^2 after serving the last client of partial path z.

This quantity is updated at each extension of a partial path by computing the exact quantity to be loaded onto the vehicle at D^2 to serve all clients hereafter till the last client of the extended partial path.

The steps of the algorithm are similar to the steps of *la(UU)*. The differences are as follows:

1. In Step 1, quantity H_z^2 is set to: $H_z^2 = \Pi^2, z \in Z_0$
2. In Step 4, P_z^2 and H_z^2 are updated based on what was discussed above:

   i. The level of the remaining inventory $\bar{P}_{z\prime}^2$ in D^2 after serving client i is computed according to the previous visits of the vehicle to the two warehouse facilities. Consider the last visit to D^2 for stock replenishment prior to client i. Let j be the client after which this visit occurred. Also let p_j be the remaining inventory in D^2 after serving client j by path $z \in Z_{i-1}$. Then, $\bar{P}_{z\prime}^2$ ($z' \in Z_i$ and is the extension of z) is calculated as follows:

$$\bar{P}_{z'}^2 = \begin{cases} \bar{P}_z^2 - \min\{\bar{P}_z^2, Q - \bar{q}_z\}, & x_{i-1}^2 = 1 \\ \min\{p_j, \bar{P}_z^2 + \bar{q}_z\}, & x_j^2 \times x_{i-1}^1 \times \left(1 - \min\left\{1, \sum_{m=j+1}^{i-2} (x_m^1 + x_m^2)\right\}\right) = 1 \\ \bar{P}_z^2, & otherwise \end{cases} \quad \begin{matrix} (6.28)\ (a) \\ (6.28)\ (b) \\ (6.28)\ (c) \end{matrix}$$

Equation (6.28a) provides the remaining inventory in D^2 if the vehicle reached client i from $i-1$ through D^2. In this case, the maximum possible quantity is loaded (perhaps temporarily) onto the vehicle, as discussed above. Equation (6.28b), concerns the case in which the vehicle reached client i from $i-1$ through D_1 and the vehicle has visited D^2 prior to i; in this case the inventory level of D^2 is updated by adding the remaining quantity of products on board ($\bar{q}_z$). This Equation ensures that the updated inventory level of D^2 does not exceed the level p_j of the inventory prior to the last visit to D^2. Equation (6.28c) states that there is no change in the inventory level of D^2, in any other case; that is: (a) if the vehicle does not visit any warehouse between clients $i-1$ and i, (b) between two sequential visits to D^1, (c) if the current visit is the first one to D^2.

ii. The maximum inventory level in D^2 after serving client i by partial path $z'(z' \in Z_i$ and is the extension of $z \in Z_{i-1})$ is calculated as follows:

$$\bar{H}_{z'} = \begin{cases} \min\{p_j, \bar{P}_z^2 + \bar{q}_z\}, & x_j^2 \times \left(1 - \min\left\{1, \sum_{m=j+1}^{i-1} (x_m^1 + x_m^2)\right\}\right) = 1 \\ \bar{H}_z^2, & otherwise \end{cases} \quad \begin{matrix} (6.29)\ (a) \\ (6.29)\ (b) \end{matrix}$$

In Eq. (6.29a) the maximum inventory level is defined as in Equation (6.28c), by adding (returning) the remaining quantity of products on-board, but not exceeding the previous inventory level (p_j) of D^2. Equation (6.29b) states that in any other case there is no change to the maximum inventory level of D^2.

3. In step 4, q_z' is updated as follows:

$$\bar{q}_{z'} = \begin{cases} \bar{q}_z + \min\{\bar{P}_z^2, Q - \bar{q}_z\}, & x_{i-1}^2 = 1 \\ \bar{q}_z - r_i, & x_{i-1}^1 + x_{i-1}^2 = 0 \\ Q - r_i, & x_{D_{i-1}^1} = 1 \end{cases} \quad \begin{matrix} (6.30)\ (a) \\ (6.30)\ (b) \\ (6.30)\ (c) \end{matrix}$$

- Equation (6.30a) states that when the vehicle visits D^2, then it loads the maximum feasible quantity of products. Equation (6.30b) states that during the direct visit to client i

the vehicle load is decreased by exactly the amount corresponding to the demand of client i. Equation (6.30c) states that when the vehicle visits D^1, then it is loaded up to its capacity.

4. In step 6, consider the two paths $z'_1, z'_2 \in Z_i$. Then partial path z'_1 dominates partial path z'_2 (and, thus, z'_2 can be discarded) if the following holds additionally to (6.26) and (6.27):

$$\bar{H}_{z'_1} \geq \bar{H}_{z'_2} \tag{6.31}$$

If (6.31) is satisfied, then in case of partial path z'_1, the inventory level in D^2 is equal or higher, and, thus, the vehicle may visit D^2 more (or the same number of) times for replenishment, if appropriate, in order to minimize the total distance travelled. As a result, partial path z'_2 can be discarded without further extending it.

6.3.4 Labeling algorithm *la(LL)*

This is again an extension of *la(UU)*, but it does not guarantee optimality. As in the previous cases, let Z_i be the set of all partial paths reaching client (node) $i \in K$. Each partial path $z \in Z_i$ is characterized by $\bar{C}_z$, and $\bar{q}_z$. In this case both warehouse facilities have limited inventory levels; we define, $\bar{P}_z^1, \bar{P}_z^2$ to be the inventory of the facilities D^1 and D^2, respectively, after serving the last client of partial path z. Finally, note that in this case the variables H_z^1 and H_z^2 for the maximum inventory level are not used.

The differences from the labeling algorithms of the previous cases are summarized as follows:

1. In Step 3 each partial path incident to a client i can be extended to the next client $i+1$ by: (a) Direct trip to the next client, (b) trip through D^1, (c) trip through D^2, and (d) trip through both warehouse facilities. In this case as well, the appropriate quantity to be loaded onto the vehicle during a warehouse visit is not known at the time of extending the path. Thus, the vehicle (initially) receives the maximum possible load (up to its capacity, or up to the available inventory).
2. A significant variation is necessary for cases (b) or (c) above if a different warehouse was visited prior to the current visit; i.e. D^2 for (b) and D^1 for (c). In such an event, two path extensions are created, differing only by the inventory level left in these latter warehouses. In the first extension, it is considered that the maximum possible quantity was loaded onto the vehicle during its previous (to the current) visit to the warehouse. In the second extension, it is considered that (during that previous visit) the vehicle was

loaded just to satisfy the demand of the clients served between the previous and the current visit. In both extensions it is assumed that during the current visit the vehicle receives the maximum possible load. By creating these two extensions we ensure that we retain path variations in which either warehouse (one warehouse per extension) holds the maximum possible inventory, to support future possible visits appropriately. Either case may result in lower distance travelled by the vehicle in the future.

3. Case (d) above is only considered if there isn't adequate quantity of products in one warehouse to load the vehicle up to its capacity; in that case it is necessary to create one more extension, in which the vehicle also visits the other depot before serving the next client. Again, this case may eventually lead to a minimum cost route. Overall, at most $5 \times |Z_{i-1}|$ new paths will be created.
4. In Step 4, P_z^1, P_z^2, q_z, are updated according to Eqs. (6.28a) - (6.28c) and (6.30a)-(6.30c), respectively.
5. In Step 6, in order to prune partial paths of inferior quality, the second criterion Eq. (6.27) is modified to:

$$\bar{C}_{z'_1} < \bar{C}_{z'_2} \tag{6.32}$$

This modification secures that the two extensions created as described in 1 above will not be discarded immediately. This modification may lead to unnecessary memory consumption due to a large number of accumulated labels. To avoid the latter problem the dominance rule described by Eq. (6.27) may be applied every R iterations. Note that these pruning criteria are strong, and, as consequence, some good quality paths may be discarded. Therefore, reaching the optimal solution of the problem is not guaranteed.

6.3.5 Partitioning heuristic for problem LL – *pa(LL)*

We developed a partitioning heuristic to enhance the performance of the *la(LL)* algorithm. The heuristic decomposes this problem as follows: (a) Initially the client sequence to be served is divided into N consecutive client subsequences (legs). (b) For each leg, *la(LL)* is applied to create several paths. (c) The best feasible combination among these paths is obtained by a binary integer programming problem, denoted as the matching problem.

Consider two consecutive legs, $l-1$, and l. The last client of leg $l-1$ is the starting point of leg l. In order to maximize the potential to reach a feasible solution by the matching problem, from all paths of leg $l-1$ only three are selected for expansion to leg l: (a) The one that reaches the last client of leg $l-1$ with the maximum vehicle load, (b) the one reaching this

client with the maximum inventory in warehouse 1, and (c) the one reaching the last client with maximum inventory in warehouse 2. In each leg, the algorithm creates the corresponding paths by applying the *la(LL)* algorithm.

Note that, based on the way paths are created, joining two randomly selected paths from two consecutive legs may not create a feasible sequence due to load incompatibilities. This is addressed by the matching algorithm. To formulate the matching problem, consider $\delta = \delta_1 \cup \delta_2 \ldots \cup \delta_N$ to be the set of all paths created for all legs (δ_1 is the set of paths created for leg 1, and so on). $R_l, l = 1, \ldots, N$ is the total demand of clients associated with leg l. Also, let y_j be the binary variable which is assigned the value 1, if a path j is used and 0 if not and c_j be the respective cost of path j. π_j^1 and π_j^2 are the total product quantities loaded onto the vehicle during path j from warehouses 1 and 2, respectively. Let Q_j^{out} be the vehicle load at the end of path j, and Q_j^{in} the minimum demand required at the start of path j in order to complete feasibly this path. The latter amount is calculated by summing the demand of the clients until the first visit to a replenishment point in path j. Then the matching problem may be formulated as follows:

$$min \sum_{j \in \delta} c_j y_j \tag{6.33}$$

Subject to:

$$\sum_{j \in \delta_l} y_j = 1, \quad l = 1, \ldots, N \tag{6.34}$$

$$\sum_{j \in \delta} y_j \, \pi_j^1 \leq \Pi_1 \tag{6.35}$$

$$\sum_{j \in \delta} y_j \, \pi_j^2 \leq \Pi_2 \tag{6.36}$$

$$\sum_{j \in \delta_l} y_j \, Q_j^{in} - \sum_{j \in \delta_{l-1}} y_j \, Q_j^{out} \leq 0, \quad l = 2, \ldots, N \tag{6.37}$$

$$\sum_{j \in \delta_l} y_j \, Q_j^{out} + \sum_{j \in \delta_l} y_j \, R_l \leq \sum_{j \in \delta_l} y_j \, \pi_j^1 + \sum_{j \in \delta_l} y_j \, \pi_j^2 + \sum_{j \in \delta_{l-1}} y_j \, Q_j^{out}, \quad l = 2, \ldots, N \tag{6.38}$$

$$y_j \in \{0,1\}, \quad j \in \delta \tag{6.39}$$

Constraint (6.34) ensures that all legs of the client sequence are served once. Constraints (6.35) and (6.36) ensure that the consumption from each warehouse does not exceed its initial inventory. Constraint (6.37) ensures that the load of the vehicle at the end of leg $l-1$ is enough to satisfy the minimum load required to serve leg l. Constraint (6.38) ensures that the

inflow product quantity of path j is less than the *outflow* product quantity of path j. The inflow quantity of path j of leg l is the sum of the load of the vehicle at the end of leg $l-1$ and the quantities loaded during possible warehouse visits. The outflow quantity is the sum of products delivered to clients and the vehicle load at the end of leg l. The matching algorithm comprises the following steps:

Step 1: Create $l = 1, \dots, N$ consecutive client sequences (legs)

Step 2: Create paths for the first leg $l = 1$ applying *la(L)*

Step 3: Create all feasible paths for leg $l+1$ by extending those paths of leg l that ended with: (a) the maximum vehicle load, (b) the maximum inventory in warehouse 1, and (c) the maximum inventory in warehouse 2

Step 4: $l = l + 1$

Step 5: Repeat Steps 3 and 4 until $l = N$

Step 6: Solve the matching problem using a binary programing solver.

6.4 Experimental analysis

A wide range of experiments were conducted in order to evaluate the efficiency of the methods developed to solve the above three problem variants. All methods were implemented in Mathworks' Matlab R2011b and run on a PC with an Intel Xeon E5335 CPU, and 8GB of RAM. The problems were also solved using the commercial MILP solver TOMLAB/CPLEX Version 7.5 (R7.5.0) to compare and assess the efficiency of the proposed algorithms.

To create the test problems we used a random generator, which distributed uniformly $n+4$ points in a 24×24 square space. The first n points represent the locations of the clients to be served; the next two points represent the starting and ending positions of the vehicle, while the last two points represent the locations of the two warehouses. The capacity of the vehicle was set to Q. The client demand was generated randomly using a uniform distribution in the interval $[1, Q/3]$. It is also assumed that the inventory of the warehouses was either unlimited or equal to half the mean cumulative demand of all clients ($\Pi = {}^{1}/_{2} \sum_{i=1}^{n} r_i$). In all problems, without loss of generality, we assume that the vehicle departs empty from its starting position ($q_0 = 0$). Therefore, prior to serving the first client, the vehicle should visit one of the warehouses to load product.

For each of the three problems cases (UU, UL, and LL) we generated problems for increasing values of n and Q. Ten (10) problem instances were generated and solved for each (n, Q) pair, and the computational time and total cost were logged.

6.4.1 Problem case UU

As mentioned in subsection 6.3.1 both *dp(UU)* and *la(UU)* algorithms solve the UU problem to optimality. In terms of computational efficiency, Figure 6.2 displays of the three solution methods applied (*dp(UU)*, *la(UU)*, and MILP) the average computational time over the 10 problem instances for the number of clients n increasing from 100 to 10000 by a step of 100. Three graphs are presented for different values of the vehicle's capacity $Q \in \{12,192,1000\}$. As it is expected the computational time of *dp(UU)* increases with n and Q.

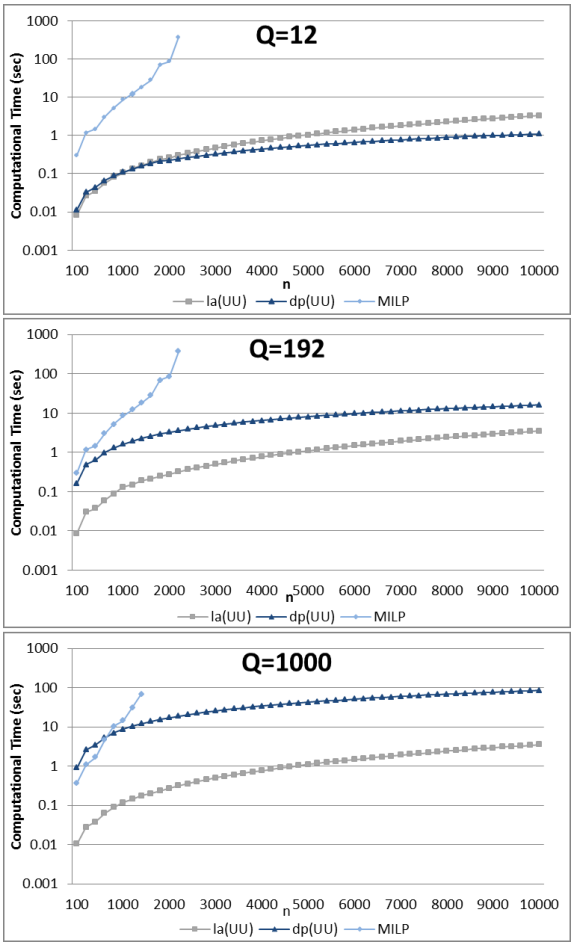


Figure 6.2: Problem UU - Computational time as a function of the number of clients n for three values of the capacity Q

The computational time of *la(UU)* increases with the number of clients n, but it appears to be less dependent on Q, since in this case paths that correspond to unfavorable loads are pruned early in the algorithm. Furthermore, since the ratio of the vehicle capacity to the average

demand is constant, the number of partial paths created by the algorithm is less dependent on Q. For $Q = 12$, there is no substantial difference in the efficiency of *dp(UU)* and *la(UU)*. For the larger values of Q, however, *la(UU)* is significantly more efficient (as expected). Note that both algorithms solve problems of up to 10^7 states at reasonable computational times. Finally, MILP appears to be significantly less efficient than both *dp(UU)* and *la(UU)*.

6.4.2 Problem case UL

Figure 6.3 shows the average computational time for solving UL by the following algorithms: (a) the labeling algorithm, denoted as *la(UL-opt)*, which solves the problem to optimality; (b) the labeling algorithm with the stronger dominance rules of Eqs. (6.26) and (6.27), denoted as *la(UL-non-opt)*, which reaches efficient but not optimal solutions; (c) the MILP solver. The dynamic programing algorithm for this problem case requires much higher computational times, and is deemed non-competitive (in this as well as in the LL case). The three solution methods are tested for n increasing from 100 to 1500 by a step of 100. Each graph corresponds to a different value of $Q \in \{12,192,1000\}$.

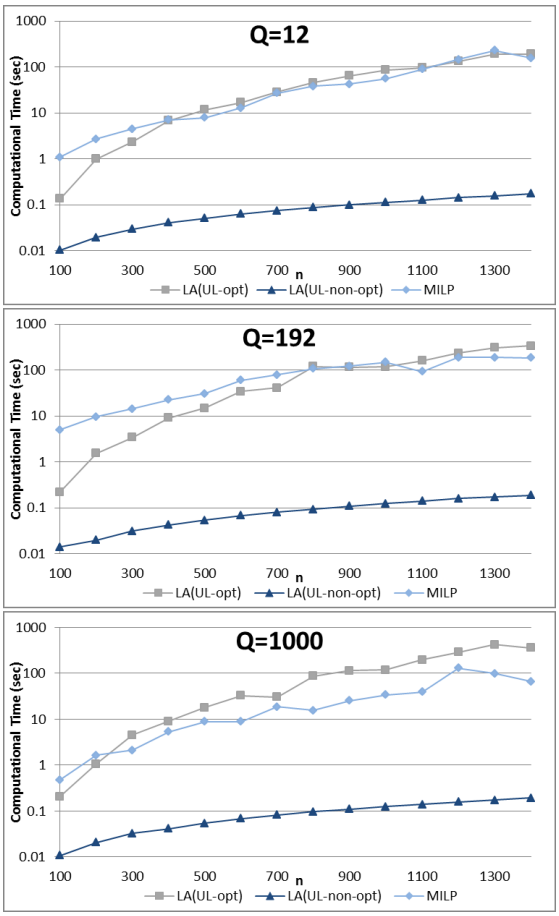


Figure 6.3: Problem UL - Computational time as a function of the number of clients n for three values of the capacity Q

In this case *la(UL-opt)* displays computational times similar to those of MILP for $Q = 12$ and inferior for the other two cases. However, *la(UL-non-opt)* results in significant computational savings, due to the stricter dominance criteria. Figure 6.4 displays the average deviation of the solutions obtained by the *la(UL-non-opt)* w.r.t. the optimal solutions. It appears that *la(UL-non-opt)* is a very effective heuristic for the UL problem.

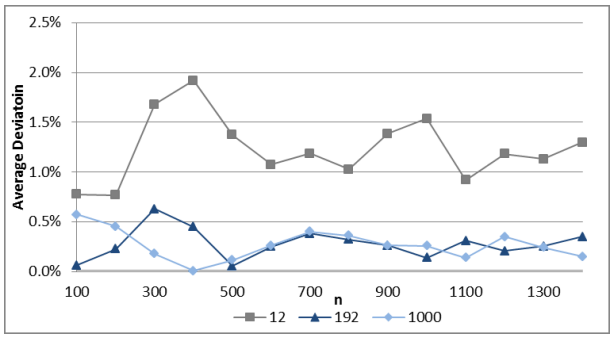


Figure 6.4: Average deviation of *la(UL-non-opt)* solutions from the optimal w.r.t. n for $Q = \{12, 192, 1000\}$

6.4.3 Problem case LL

For this problem, we initially conducted preliminary experiments in order to fine tune algorithm *pa(LL)*. Specifically, we investigated the effect of the ratio of clients per leg on computational time and solution quality. We used ratio values of 2, 5, 10, 50, and 100 (if applicable). For all ratio values except the value of 2, the solutions obtained were very close to the optimal solutions provided by the MILP. The average deviation from the optimal solution was found to be 0.7%. The results of the experiments have not been included here.

In order to examine the performance of *la(LL)* and *pa(LL)*, the MILP is also set to terminate its search if the gap between the best solution of the linear relaxation and the best integer solution at any point of the search is less than 1.5%. The latter gap is set to be approximately twice the observed average deviation of *la(LL)* from the optimum in order to get solutions of comparable quality by the two solution methods.

Figure 6.5 shows the average computational times of *la(LL)*, *pa(LL)* with 100 clients per leg, MILP(opt) and MILP(1.5%) (i.e. the optimal and suboptimal MILP solution methods, respectively) for n increasing from 100 to 1000 by a step of 100. In this range of n MILP(opt) reaches the optimal solution. For this n range, *la(LL)* and *pa(LL*-100) perform almost the same and they are about an order of magnitude faster than MILP(1.5%).

Figure 6.6 displays the average deviation of *la(LL)*, *pa(LL*-100) and MILP(1.5%) from the optimal solution provided by MILP(opt). It seems that the other three methods provide

solutions of the same quality, while MILP(1.5%) performs slightly better than *la(LL)* and *pa(LL*-100).

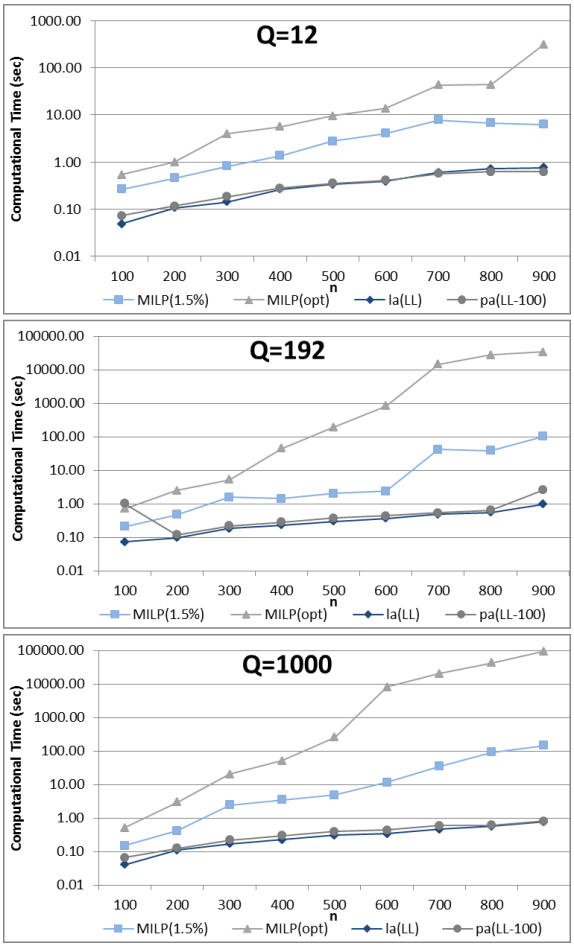


Figure 6.5: Problem LL - Computational time as a function of n for three values of the capacity Q

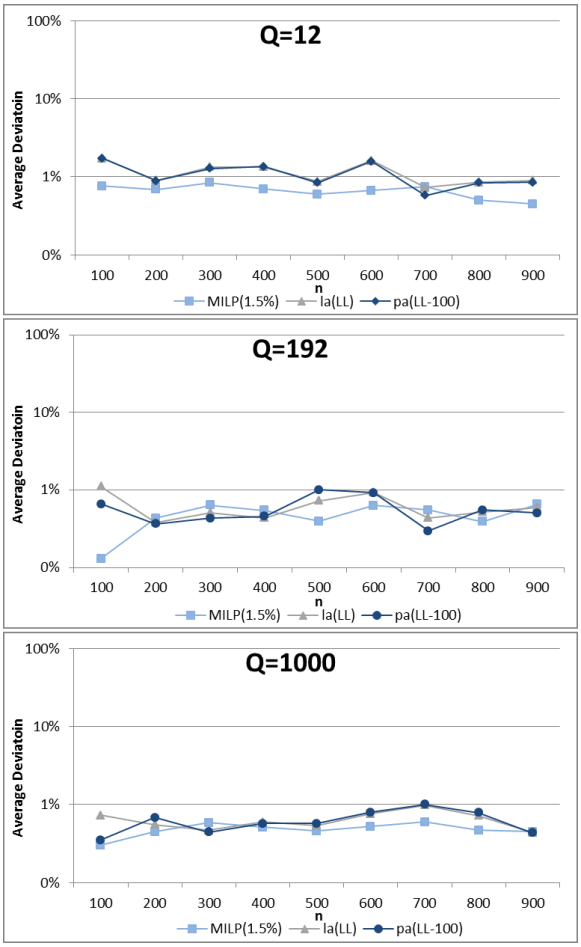


Figure 6.6: Problem LL - Average deviation of the solution from the optimal as a function of n for three values of the capacity Q

Figure 6.7 shows the average computational times of *la(LL)* and *pa(LL*-100) for n increasing from 100 to 10000 by a step of 500. It seems that between $n = 1000$ and 1500 *pa(LL*-100) gradually surpasses the performance of *la(LL)*; for $n = 10000$, *pa(LL*-100) is one order of magnitude faster than *la(LL)*. This indicates the superiority of *pa(LL*-100) for large scale problems.

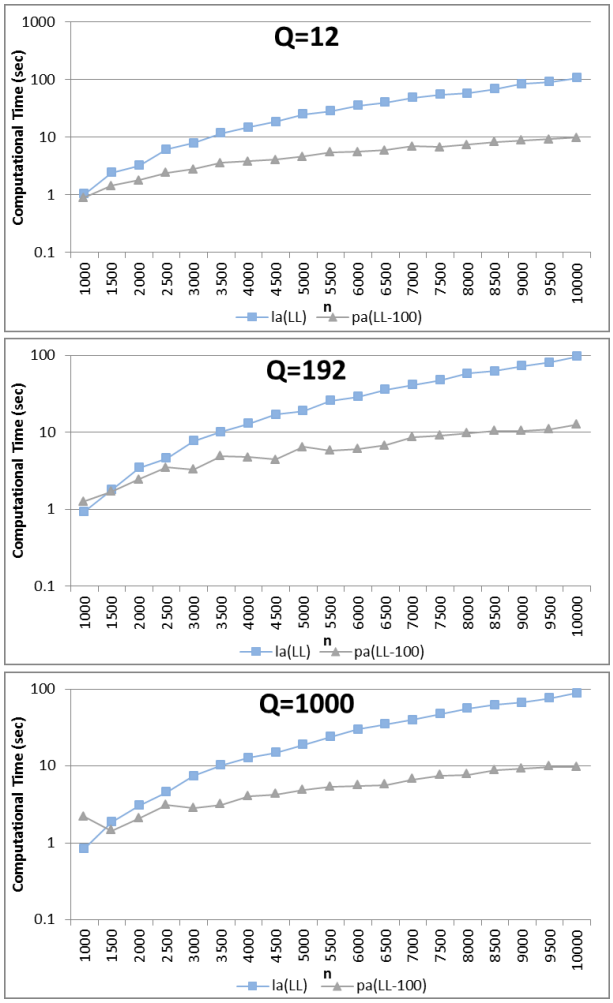


Figure 6.7: Problem LL - Computational time as a function of the number of clients n for three values of the capacity Q

Chapter 7. Conclusions and recommendations for future research

7.1 Conclusions

In this dissertation we investigated paratransit services under normal and emergency conditions. To do so, we examined the design aspects and analyzed the special characteristics of novel paratransit transport systems. Furthermore, we introduced, formulated and proposed exact and heuristic solution approaches to solve the related routing problems.

In the area of paratransit services under normal conditions, SPTS, a flexible surface transport service, was studied in which a bus may diverge from its nominal path to pick-up passengers with limited mobility and drop them off at their destination. Such a system may offer higher quality of transport to clients with a disability without increasing significantly public transport expenditure.

A new routing problem (DRBRP) was defined for the SPTS that aims to maximize the number of served paratransit requests, reduce the total ride time of the paratransit clients, while serving the nominal route at a pre-defined service level. To solve this problem we proposed an exact Branch and Price framework, which includes a labeling algorithm with appropriate dominance rules in order to reach the optimal solution within reasonable computational times. The proposed algorithm improves the computational time by up to 50% with respect to existing approaches for similar problems. We used this algorithm to: (a) study the effect of key problem constraints on the service level, and (b) support the design of SPTS systems taking into account the characteristics of different transport environments.

This study indicated that three system parameters should be considered carefully: (a) the allowable level d of disruption of the nominal route, (b) the available seats per paratransit request per trip $Q/(n/k)$, and (c) the ratio of the time window related to the requests over the service headway ξ/f. The values of these parameters must be high enough to ensure a good service level, but within reasonable limits in order to avoid major disruptions of the regular service, or high system costs. Furthermore, different considerations apply to different environments due to their intrinsic characteristics. For example, in an urban environment, which is characterized by high service rates, the values of $Q/(n/k)$, and ξ/f are relatively high, even for one reserved seat per trip. In this case, the system could be applied efficiently by allowing relatively low disruption to the nominal route (low values of d). In a suburban environment, with moderate service rates and longer distances between regular bus stops, the system could perform sufficiently well for a lower value of the allowable delay (d). Finally, in a rural environment, in which ξ/f is relatively low, the values of $Q/(n/k)$ and d need to be higher.

Subsequently, we examined SPTS-T, an interesting enhancement of SPTS, which allows the system to fully satisfy the door-to-door paratransit demand. SPTS-T considers two modes of transport: (a) the public bus system (similarly to SPTS), and (b) accessible taxi (or private hire) services. The problem objective is to minimize the cost of the combined transportation system, while serving all paratransit requests and offering sufficient level of service.

For SPTS-T we extended the SPTS mathematical model to include the combined attributes of DRBRP and DARP. This extension includes two types of vehicles, and appropriate mode-dependent time windows to safeguard uniform level of service. To solve this problem, we developed an exact Branch and Price framework, which combines two types of sub-problems, the first related to the bus mode, and the second related to the taxi mode. The partial solutions

generated by these sub-problems are treated in a column generation scheme to obtain the optimal solution. An extensive experimental study was performed to provide design guidelines for SPTS-T.

This study indicated that SPTS-T may be applicable even for cases for which the deviation of a bus from its nominal route is relatively expensive, i.e. for large values of the related cost scaling factor a. Having defined a, the designer should focus on setting the values of: (a) parameter d, which regulates the maximum allowable time deviation of a bus trip from its nominal route, as well as the maximum ride time of the paratransit clients to be served by taxi; (b) parameter ξ that is the width of the drop-off (or pick-up) time windows for the paratransit requests to be served by the bus system; (c) the taxi capacity $\bar{Q}$ for paratransit passengers. A suitable set of parameter values that balance SPTS-T costs without compromising the service level may be found. Increasing the values of the system parameters beyond these thresholds would deteriorate service quality without significantly improving the respective operational costs.

In the area of paratransit services under emergency conditions we investigated the problem of evacuating the residents of a region in case of an emergency using transit resources - the bus evacuation problem. We also studied an interesting variation of this problem, which deals with the evacuation of casualties from several triage points. This case may be encountered in a battlefield or after disasters (natural catastrophes, terrorist attacks, etc.), in which a significant number of casualties should be transported to available medical facilities. We developed a hybrid algorithm (HCG) that incorporates the column generation framework with the well-known meta-heuristic methods LNS and VNS. The efficiency of the algorithm has been validated using small test instances by comparing the solutions obtained with the optimal solutions of a MILP solver.

The results indicate that the algorithm generates near optimal solutions in very efficient computational times. In an attempt to reduce the problem complexity, we introduced discretization of the demand of the pick-up locations. Using HCG we studied the impact of the granularity of demand discretization on the quality of the solution. The results indicate that as the granularity of demand discretization increases, the computational times are significantly increased, although the impact on the solution quality is appreciable only for cases in which the demand discretization level restricts the efficient utilization of the vehicle's capacity. The proposed method was also tested with instances of practical size involving the evacuation of up to 4,000 evacuees for the BEP and up to 100 casualties for the

CEP. The results indicated that the hybrid method generates consistent solutions within reasonable computational times. Thus, HCG is appropriate to address efficient planning at short notice prior to population evacuations.

Finally, we studied insights for logistics cases, in which limitations similar to those encountered in the aforementioned transportation problems are of relevance; i.e. routes with strict sequences, or capacity limitations in collection/gathering centers. Specifically, we investigated the single vehicle routing problem with a predefined client sequence and two replenishment warehouse facilities. We defined three new cases of progressively increasing state space, based on the inventory of each warehouse: Unlimited – Unlimited (UU), Unlimited – Limited (UL), and Limited – Limited (LL). These cases were modeled, solved and analyzed. For solving them, we developed algorithms ranging from standard dynamic programming for the simplest UU case, to labeling algorithms and a new partitioning heuristic. We also used a commercial MILP solver for benchmarking purposes. The proposed solution methods were evaluated using a large number of randomly generated problem instances.

The labeling algorithms reach the optimum for UU and UL, and provide suboptimal, but efficient, solutions for LL. In terms of computational efficiency, the labeling algorithms were found to be significantly more efficient than dynamic programming. They were also found to be more efficient than MILP (running for a long but limited period) for the UU case. For problem LL, the *la(LL)* and *pa(LL*) heuristics are considerably more efficient than MILP (1-3 orders of magnitude faster) and obtain efficient solutions (within 1% away of the optimal). Finally, *pa(LL*) was found to be the most efficient method for large scale problems (up to one magnitude faster than *la(LL)*).

7.2 Recommendations for future research

Immediate directions for future research in SPTS-like systems include the following.

Dynamic SPTS systems: In such systems the paratransit requests arrive in real time during actual operation. These dynamic requests need to be incorporated in the evolving system state. One way to do so is through periodic re-optimization. In this case, the re-optimization problem should to be solved in real time. Thus, intelligent heuristics, metaheuristics or hybrid approaches may be developed to address the aforementioned problems, in an attempt to achieve efficient solutions in much shorter computational times. Furthermore the frequency

(or time sequence) of the re-optimization events, and the process of implementing the revised plans are areas of interest. For the former, the frequency of re-planning should be investigated based either on the number of new requests or the time elapsed between consecutive re-panning epochs. Concerning the process of implementing the revised plans, the investigation may consider whether during re-planning one should consider only the new requests for service, or also the ones not executed yet; see also Ninikas and Minis (2014) for related concepts.

Stochastic SPTS systems: Including stochastic information, concerning inputs and parameters other than the paratransit requests, and addressing the related issues appropriately would result in a better service quality by ensuring either that fewer failures will occur during implementation, or that less unscheduled disruptions will be realized. This disruptions may involve uncertainty with respect to: (a) the travel times of the vehicles, e.g. due to traffic congestion, (b) the service, time windows or service times, e.g. a booked request may delay its arrival to the designated pick-up location, and (c) the demand, e.g. cancelations of a request. We refer to Gendreau *et al.* (1996) for detailed review on stochastic vehicle routing problems.

Dynamic Stochastic SPTS systems: Stochastic information may be available regarding the arrival of new requests. Such information may involve probability distributions, perhaps time varying, derived from historical data. Addressing appropriately this information will result to better exploitation of the available resources, and thus to better quality of service. Scheduling for Dynamic Stochastic SPTS cases may be addressed by multiple scenario approaches (see Bent and Van Hentenryck, 2004) or by introducing dummy requests based on forecasts about the demand (Ichoua *et al.*, 2006).

Consider the *m*-SPTS, or *m*-SPTS-T problems, in which resources from multiple (*m)* bus routes compete for the same paratransit requests, thus enhancing the capability of the public bus system to serve these requests. The current formulation of Chapters 3 and 4 may address this case by appropriately expanding the set of bus stops B, and the set of bus trips K. Thus, the new sets will be of the form $\bar{B} = \bigcup_{m \in M} B^m$ and $\bar{K} = \bigcup_{m \in M} K^m$, where m is a bus route belonging to the set of all bus routes M. Furthermore, the branch and price approaches may be adapted appropriately, since they treat each bus trip independently.

Another interesting enhancement may consider multiple vehicle types in SPTS-T. This could be achieved by expanding (a) set V related to taxis to set $\bar{V} = \bigcup_{t \in T} V^T$ of vehicle modes T, and (b) $\bar{Q}$ the capacity of taxis to $\bar{Q}^t, t \in T$ to be the capacity of the corresponding vehicle mode in T. These

modifications may be incorporated to the solution method by defining a subproblem for each vehicle mode.

In the area of evacuation planning, further research could investigate

- Evacuation planning using multiple modes of transport. Following recent trends, multiple modes may incorporate public buses and private cars of volunteers. The latter case combines the location routing with the car sharing and carpooling problems
- Modeling the behavior of the evacuees and with respect to the manner they reach the pick-up locations. Zhang and Chang (2014) assumed a rate of arrival of the evacuees which they model by setting appropriate time windows. One may consider this aspect within the models and methods proposed in Chapter 5 of this dissertation through time window limitations related to the pick-up locations, and related modifications to the solution approach.
- Combining BEP and CEP, that is the evacuees of BEP and the casualties of CEP, in a single problem under a multi objective perspective.

Finally, applying the dissertation proposals to appropriate case studies related to real practical problems would further validate the models and methods proposed here, perhaps leading to appropriate extensions to address issues of significant practical value.

References

Abdelgawad, H., Abdulhai, B., & Wahba, M. (2010). Optimizing mass transit utilization in emergency evacuation of congested urban areas. *Proceedings, the 89th Annual Meeting of The Transportation Research Board.* Washington DC.

Agarwal, Y., Mathur, K., & Salkin, H. M. (1989). A set-partitioning-based exact algorithm for the vehicle routing problem. *Networks, 19*(7), 731-749.

Alsnih, R., & Hensher, D. A. (2003). The mobility and accessibility expectations of seniors in an aging population. *Transportation Research Part A: Policy and Practice, 37*(10), 903-916.

Altay, N., & Green, W. G. (2006). OR/MS research in disaster operations management. *European Journal of Operational Research, 175*(1), 475-493.

American Association of Retired Persons. (2006, AARP). *We Can Do Better: Lessons Learned for Protecting Older Persons in Disasters.* Washington, D.C: AARP.

An, S., Cui, N., Li, X., & Ouyang, Y. (2013). Location planning for transit-based evacuation under the risk of service disruptions. *Transportation Research Part B: Methodological, 54*, 1-16.

Anily, S., & Hassin, R. (1992). The swapping problem. *Networks, 22*(4), 419-433.

Anily, S., & Mosheiov, G. (1994). The traveling salesman problem with delivery and backhauls. *Operations Research Letters, 16*(1), 11-18.

Applegate, D. L., Bixby, E. R., Chvátal, V., & Cook, J. W. (2007). *The Traveling Salesman Problem: The Traveling Salesman Problem: A Computational Study.* Princeton: Princeton University Press.

Archetti, C., Feillet, D., Hertz, A., & Speranza, M. (2009). The capacitated team orienteering and profitable tour problems. *Journal of the Operational Research Society, 60*(6), 831-842.

Avendano, M., Glymour, M. M., Banks, J., & Mackenbach, J. P. (2009). Health Disadvantage in US Adults Aged 50 to 74 Years: A Comparison of the Health of Rich and Poor Americans With That of Europeans. *American Journal of Public Health, 99*(3), 540-548.

Bard, J., Huang, L., Dror, M., & Jaillet, P. (1998). A Branch and Cut Algorithm for the VRP with Satellite Facilities. *IIE Transactions on Operations Engineering, 30*(9), 821-834.

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W., & Vance, P. (1998). Branch-and-Price: Column Generation for Solving Huge Integer Programs. *Operations Research, 46*(3), 316-329.

Bent, R. W., & Van Hentenryck, P. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, 52(6), 977-987.

Berbeglia, G., Cordeau, J. -F., & Laporte, G. (2010). Dynamic pickup and delivery problems. *European Journal of Operational Research, 202*(1), 8-15.

Berbeglia, G., Cordeau, J., Gribkovskaia, I., & Laporte, G. (2007). Static pickup and delivery problems: a classification scheme and survey. *TOP, 15*(1), 1-31.

Bianchi, L., Birattari, M., Chiarandini, M., Manfrin, M., Mastrolilli, M., Paquete, L., et al. (2006). Hybrid metaheuristics for the vehicle routing problem with stochastic demands. *Journal of Mathematical Modelling and Algorithms, 5*(1), 91-110.

Bianco, L., Mingozzi, A., & Ricciardelli, S. (1994). A set partitioning approach to the multiple depot vehicle scheduling problem. *Optimization Methods and Software, 3*(1-3), 163-194.

Bish, D. (2011). Planning for a bus-based evacuation. *OR Spectrum, 33*(3), 629-654.

Blum, C., Roli, A., & Sampels, M. (2008). Hybrid metaheuristics. In *Studies in Computational Intelligence, Vol. 114*. Berlin: Springer.

Boland, N., Dethridge, J., & Dumitrescu, I. (2006). Accelerated label setting algorithms for the elementary resource constrained shortest path problem. *Operations Research Letters, 34*(1), 58-68.

Borndörfer, R., Grötschel, M., Klostermeier, F., & Küttner, C. (1999). *Telebus Berlin: Vehicle scheduling in a dial-a-ride system.* Berlin Heidelberg: Springer.

Boussier, S., Feillet, D., & Gendreau, M. (2007). An exact algorithm for the team orienteering problem. *4OR, 5*(3), 211–230.

Brake, J., Mulley, C., Nelson, J. D., & Wright, S. (2007). Key lessons learned from recent experience with Flexible Transport Services. *Transport Policy, 14*(6), 458-566.

Brandão, J., & Mercer, A. (1998). The multi-trip vehicle routing problem. *The Journal of the Operational Research Society, 49*(8), 799-825.

Bretschneider, S. (2013). *Mathematical Models for Evacuation Planning in Urban Areas* (Vol. Lecture Notes in Economics and Mathematical Systems 659). Berlin Heidelberg: Springer Verlag.

Broome, K., Worrall, L., Fleming, J., & Boldy, D. (2012). Evaluation of flexible route bus transport for older people. *Transport Policy, 21*, 85-91.

Brunkard, J., Namulanda, G., & Ratard, R. (2008). Hurricane Katrina Deaths, Louisiana, 2005. *Disaster Medicine and Public Health Preparedness, 2*(4), 15-23.

Burns, P. C. (1999). Navigation and the mobility of older drivers. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 54*(1), S49-S55.

Butt, S. E., & Cavalier, T. M. (1994). A heuristic for the multiple tour maximum collection problem. *Computers & Operations Research, 21*(1), 101-111.

Chabrier, A. (2006). Vehicle Routing Problem with elementary shortest path based column generation. *Computers & Operations Research, 33*(10), 2972-2990.

Chandra, S., & Quadrifoglio, L. (2013). A model for estimating the optimal cycle length of demand responsive feeder transit services. *Transportation Research Part B: Methodological, 51*, 1-16.

Chao, I., Golden, B. L., & Wasil, E. A. (1996). The team orienteering problem. *European Journal of Operational Research, 88*(3), 464-474.

Chao, I., Golden, B. L., & Wasil, E. A. (1996a). A fast and effective heuristic for the orienteering problem. *European Journal of Operational Research, 88*(3), 475-489.

Christofides, N. (1979). Traveling salesman problem. In N. M. Christofides, *Combinatorial Optimization.* New York: Wiley.

Christofides, N., & Beasley, J. E. (1984). The period routing problem. *Networks, 14*(2), 237-256.

Christofides, N., Mingozzi, A., & Toth, P. (1981). Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations. *Mathematical Programming, 20*(1), 255-282.

Church, A., Frost, M., & Sullivan, K. (2000). Transport and social exclusion in London. *Transport Policy, 7*(3), 195-205.

Clarke, C. C., & Habib, M. A. (2010). Evaluation of Multi-Modal Transportation Strategies for Emergency Evacuations. *Annual Conference and Exhibition of theTransportation Association of Canada-2010*, (pp. 26-29). Halifax, Nova Scotia.

Clarke, G., & Wright, J. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research, 12*(4), 568-581.

Cordeau, J. -F. (2006). A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research, 54*(3), 573-586.

Cordeau, J. -F., Desaulniers, G., Desrosiers, J., Solomon, M. M., & Soumis, F. (2002). The VRP with Time Windows. In P. Toth, & D. Vigo, *The Vehicle Routing Problem* (pp. 157-193). Philadelphia: SIAM.

Cordeau, J.-F., & Laporte, G. (2003). The Dial-a-Ride Problem (DARP): Variants, modeling issues and algorithms. *4OR: A Quarterly Journal of Operations Research, 1*(2), 89-101.

Cordeau, J.-F., & Laporte, G. (2007). The dial-a-ride problem: models and algorithms. *Annals of Operations Research, 153*(1), 29-46.

Coutinho-Rodrigues, J., Tralhão, L., & Alçada-Almeida, L. (2012). Solving a location-routing problem with a multiobjective approach: the design of urban evacuation plans. *Journal of Transport Geography, 22*, 206-218.

Crainic, T. G., Malucelli, F., & Nonato, M. (2000). A demand responsive feeder bus system. *7th World Congress Intelligent Transport Systems, 7WC-ITS.* Torino, Italy.

Crainic, T. G., Malucelli, F., Nonato, M., & Guertin, F. (2005). Meta-Heuristics for a Class of Demand-Responsive Transit Systems. *INFORMS Journal on Computing, 17*(1), 10-24.

Crevier, B., Cordeau, J.-F., & Laporte, G. (2007). The multi-depot vehicle routing problem with inter-depot routes. *European Journal of Operational Research, 176*(2), 756-773.

Currie, G., & Delbosc, A. (2010). Exploring public transport usage trends in an ageing population. *Transportation, 37*(1), 151-164.

Daganzo, C. F. (1984). Checkpoint dial-a-ride systems. *Transportation Research B: Methodological, 18*, 315-327.

Daganzo, C. F., & Hall, R. W. (1993). A routing model for pickups and deliveries: No capacity restrictions on the secondary items. *Transportation Science, 27*(4), 315-329.

Danna, E., & Le Pape, C. (2005). Branch-and-Price heuristics: A case study on the vehicle routing problem with time windows. In G. Desaulniers, J. Desrosiers, & M. M. Solomon, *Column Generation.* Springer Science and Business Media, Inc.

Dantzig, G. B., & Wolfe, P. (1960). Decomposition principle for linear programs. *Operations research, 8*(1), 101-11.

Dantzig, G. B., Fulkerson, R., & Johnson, S. (1954). Solution of a large scale traveling-salesman problem. *Journal of the Operations Research Society of America, 2*(4), 393-410.

Dantzig, G., & Ramser, J. (1959). The Truck Dispatching Problem. *Management Science, 6*(1), 80-91.

Deif, I., & Bodin, L. (1984). *Extension of the Clarke and Wright algorithm for solving the vehicle routing problem with backhauling.* Babson Park: A. Kidder.

Dellinger, A. M., Langlois, J. A., & Li, G. (2002). Fatal crashes among older drivers: decomposition of rates into contributing factors. *American Journal of Epidemiology, 155*(3), 234-241.

Desaulniers, G., Desrosiers, J., & Solomon, M. (2005). *Column Generation.* New York: Springer Science + Business Media Inc.

Desaulniers, G., Desrosiers, J., Erdmann, A., Solomon, M., & Soumis, F. (2002). VRP with Pickup and Delivery. In P. Toth, & D. Vigo, *The Vehicle Routing* (pp. 225-242). Philadelphia: SIAM Monographs on Discrete Mathematics and Applications.

Desrochers, M., Desrosiers, J., & Solomon, M. (1992). A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research, 40*, 342-354.

Desrosiers, J., Soumis, F., & Desrochers, M. (1984). Routing with time windows by column generation. *Networks, 14*(4), 545-565.

Diana, M., & Dessouky, M. M. (2004). A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. *Transportation Research Part B: Methodological, 38*(6), 539-557.

Dijkstra, W. E. (1959). A note on two problems in Connexion with Graphs. *Numerische Mathematik, 1*(1), 269-271.

Dikas, G., Minis, I., & Mamassis, K. (2012). *Single vehicle routing with predefined client sequence and multiple warehouse returns: The case of two warehouses. Mathematical formulations and suitable dynamic programming algorithms*. (DeOPSys Lab, University of the Aegean) Retrieved from http://labs.fme.aegean.gr/deopsys/?q=node/167

Dumas, Y., Desroiers, J., & Soumis, F. (1991). The pickup and delivery problem with time windows. *European Journal of Operational Research, 54*, 7-22.

Farquhar, M. (1995). Elderly people's definitions of quality of life. *Social Science & Medicine, 41*(10), 1439-1446.

Federal Transit Administration. (2007, FTA). *Transportation Equity in Emergencies: A Review of the Practices of State Departments of Transportation, Metropolitan Planning Organizations, and Transit Agencies in 20 Metropolitan Areas.* Washington, DC: U.S. Department of Transportation (FTA-PA-26-8001-2007).

Feillet, D., Dejax, P., Gendreau, M., & Gueguen, C. (2004). An Exact Algorithm for the Elementary Shortest Path Problem with Resource Constraints: Application to Some Vehicle Routing Problems. *Networks, 44*, 216-229.

Fillenbaum, G. G. (1985). Screening the elderly: a brief instrumental activities of daily living measure. *Journal of the American Geriatrics Society, 33*(10), 698-706.

Fischetti, M. (2001). Drowning new orleans. *Scientific American, 285*(4), 76-85.

Fischetti, M., Lodi, A., & Toth, P. (2003). Solving Real-World ATSP Instances by Branch-and-Cut. In *Combinatorial Optimization—Eureka, You Shrink!* (pp. 64-77). Berlin Heidelberg: Springer.

FLIPPER. (2011). *C4 - Feasibility studies and pilot applications, Deliverable D6 - Feasibility studies report.* INTERREG IVC.

Fonda, S. J., Wallace, R. B., & Herzog, R. A. (2001). Changes in Driving Patterns and Worsening Depressive Symptoms Among Older Adults. *Journals of Gerontology: Series B, 56*(6), S343-S351.

Ford Jr., L. R., & Fulkerson, D. R. (1958). A suggested computation for maximal multi-commodity network flows. *Management Science, 5*(1), 97-101.

Frykberg, E. (2004). Principles of Mass Casualty Management Following Terrorist Disasters. *Annals of Surgery, 239*(3), 319-321.

Gabriel, Z., & Bowling, A. (2004). Quality of life from the perspectives of older people. *Ageing and Society, 24*(5), 675-691.

Gendreau, M., Laporte, G., & Séguin, R. (1996). Stochastic vehicle routing. *European Journal of Operational Research*, 88(1), 3-12.

Garey, R. M., & Johnson, S. D. (1979). *Computers and intractability: a guide to the theory of NP-completeness.* San Francisco: WH Freeman & Co.

Gerber, J. B., Norwood, F., & Zakour, M. (2010). *Disasters, Evacuations and Persons with Disabilities: An Assessment of Key Issues Facing Individuals and Households.* U.S. Department of Education, National Institute on Disability and Rehabilitation.

Giuliano, G., Hu, H.-H., & Lee, K. (2003). *Travel patterns of the elderly: the role of land use.* (No. FHWA/CA/OR-2003/06): METRANS Transportation Center.

Glasgow, N., & Blakely, R. M. (2000). Older Nonmetropolitan Residents' Evaluations of Their Transportation Arrangements. *Journal of Applied Gerontology, 19*(1), 95-116.

Glover, F. (1989). Tabu search-part I. *ORSA Journal on computing, 1*(3), 190-206.

Glover, F. (1990). Tabu search—part II. *ORSA Journal on computing, 2*(1), 4-32.

Glover, F. (1997). Tabu search and adaptive memory programming—advances, applications and challenges. In *Interfaces in computer science and operations research* (pp. 1-75). US: Springer.

Goerigk, M., & Grün, B. (2014). A robust bus evacuation model with delayed scenario information. *OR Spectrum, 36*(4), 923-948.

Goerigk, M., Grün, B., & Heßler, P. (2013). Branch and bound algorithms for the bus evacuation problem. *Computers & Operations Research, 40*(12), 3010-3020.

Goerigk, M., Grün, B., & Heßler, P. (2014). Combining Bus Evacuation with Location Decisions: A Branch-and-price Approach. *The Conference on Pedestrian and Evacuation Dynamics 2014 (PED 2014)* (pp. 783-791). Delft: Transportation Research Procedia.

Golden, B. L., Levy, L., & Vohra, R. (1987). The orienteering problem. *Naval Research Logistics, 34*(3), 307–318.

Golden, B. L., Wang, Q., & Liu, L. (1988). A multifaceted heuristic for the orienteering problem. *Naval Research Logistics, 35*(3), 359-366.

Graves, G., McBride, R., Gershkoff, I., & Anderson, D. (1993). Flight crew scheduling. *Management Science, 39*(6), 657-682.

Gribkovskaia, I., Laporte, G., & Vlček, M. (2007). General solutions to the single vehicle routing problem with pickups and deliveries. *European Journal of Operational Research, 180*(2), 568-584.

Gurobi Optimization, Inc. (2014). Gurobi Optimizer Reference Manual.

Hamacher, H. W., & Tjandra, S. A. (2001). Mathematical Modelling of Evacuation Problems: A State of the Art. In M. Schreckenberg, & S. D. Sharma, *Pedestrian and Evacuation Dynamics* (pp. 227-266). Berlin Heidelberg: Springer.

Hansen, P., & Mladenović, N. (2001). Variable neighborhood search: Principles and applications. *European Journal of Operational Research, 34*(2), 449-467.

Hardy, S., Dudin, J. A., Holford, T. R., & Gill, T. M. (2005). Transitions between States of Disability and Independence among Older Persons. *American Journal of Epidemiology, 161*(6), 575-584.

He, S., Zhang, L., Song, R., Wen, Y., & Wu, D. (2009). Optimal Transit Routing Problem for Emergency Evacuations. *Transportation Research Board 88th Annual Meeting* (p. 13). Washington DC: Transportation Research Board.

Hernández-Pérez, H., & Salazar-González, J. J. (2007). The one-commodity pickup-and-delivery traveling salesman problem: Inequalities and algorithms. *Networks, 50*(4), 258-272.

Hess, D. B. (2009). Access to public transit and its influence on ridership for older adults in two US citie. *Journal of Transport and Land Use, 2*(1), 3-27.

Hess, D. B., & Gotham, J. C. (2007). Multi-modal mass evacuation in upstate New York: A review of disaster plans. *Journal of Homeland Security and Emergency Management, 4*(3).

Hess, D. B., Conley, B. W., & Farrell, C. M. (2013). Improving Transportation Resource Coordination for Multimodal Evacuation Planning. *Transportation Research Record: Journal of the Transportation Research Board, 2376*(1), 11-19.

Higgins, L. L., Hickman, M. D., & Weatherby, C. A. (1999). *Role of Public Transportation Operations in Emergency Management.* Texas: Texas Transportation Institute, College Station.

Hjorthol, R. (2013). Transport resources, mobility and unmet transport needs in old age. *Ageing and Society, 33*(7), 1190-1211.

Holand, J. N. (1975). *Adaptation in natural and artificial systems.* Ann Arbor: The University of Michigan Press.

Hurkens, C. A., & Woeginger, G. J. (2004). On the nearest neighbor rule for the traveling salesman problem. *Operations Research Letters, 32*(1), 1-4.

Ichoua, S., Gendreau, M., & Potvin, J. Y. (2006). Exploiting knowledge about future demands for real-time vehicle dispatching. *Transportation Science*, 40(2), 211-225.

Irnich, S., & Desaulniers, G. (2005). Shortest path problems with resource constraints. In G. Desaulniers, J. Desrosiers, & M. Solomon, *Column Generation* (pp. 33-65). Springer.

Irwin, N. (1969). Public transit and the quality of urban living. *High Speed Ground Transportation Journal, 3*(1), 92-155.

Jepsen, M., Petersen, B., Spoorendonk, S., & Pisinger, D. (2008). Subset-row inequalities applied to the vehicle-routing problem with time windows. *Operations Research, 56*(2), 497-511.

Johnson, J. E. (2002). Why rural elders drive against advice. *Journal of Community Health Nursing, 19*(4), 237-244.

Kallehauge, B., Larsen, J., Madsen, B. G., & Solomon, M. M. (2005). Vehicle Routing Problem with Time Windows. In G. Desaulniers, J. Desrosiers, & M. M. Solomon, *Column Generation* (pp. 67-98). US: Springer Science and Business Media, Inc.

Kaufmann, C. (2014). Application of Scheduling Theory to the Bus Evacuation Problem. *Operations Research Proceedings* (pp. 231-236). Switzerland: Springer International Publishing.

Kim, S. (2011). Assessing mobility in an aging society: Personal and built environment factors associated with older people's subjective transportation deficiency in the US. *Transportation research part F: traffic psychology and behaviour, 14*(5), 422-429.

Kim, S., & Ulfarsson, G. F. (2004). Travel model choice of the elderly: effects of personal, household, neighbourhood, and trip characteristics. *Transportation Research Record, 1894*, 117-126.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science, 220*, 671-680.

Laporte, G. (1992). The vehicle routing problem: An overview of the exact and approximate algorithms. *European Journal of Operation Research, 53*(3), 345-358.

Laporte, G. (2009). Fifty Years of Vehicle Routing. *Transportation Science, 43*(4), 408-416.

Laporte, G., & Martello, S. (1990). The selective traveling salesman problem. *Discrete Applied Mathematics, 26*(2-3), 193-207.

Laporte, G., & Osman, I. H. (1995). Routing problems: A bibliography. *Annals of Operations Research, 61*(1), 227-262.

Larsen, J. (2001). *Parallelization of the Vehicle Routing Problem with Time Windows.* PhD thesis (IMM-PHD-2001-62): Department of Mathematical Modeling, Technical University of Denmark.

Lavoie, S., Minoux, M., & Odier, E. (1988). A new approach for crew pairing problems by column generation with an application to air transportation. *European Journal of Operation Research, 35*(1), 45-58.

Lawler, E. L., & Wood, D. E. (1966). Branch-and-bound methods: A survey. *Operations research, 14*(4), 699-719.

Lee, E. K., & Mitchell, J. E. (2001). Integer Programming: Branch-and-Bound Methods. In *Encyclopedia of Optimization, Volume II* (pp. 509-519). Kluwer Academic Publishers.

Li, H., Raeside, R., Chen, T., & McQuaid, R. W. (2012). Population ageing, gender and the transportation system. *Research in transportation economics, 34*(1), 39-47.

Li, J.-Q., Mirchandani, P., & Borenstein, D. (2009a). A Lagrangian heuristic for the real-time vehicle rescheduling problem. *Transportation Research Part E: Logistics and Transportation Review, 45*(3), 419-433.

Li, J.-Q., Mirchandani, P., & Borenstein, D. (2009b). Real-time vehicle rerouting problems with time windows. *European Journal of Operational Research, 194*(3), 711-727.

Litman, T. (2006). Lessons from Katrina and Rita: What major disasters can teach transportation planners. *Journal of Transportation Engineering, 132*(1), 11-18.

Lopes, R. B., Ferreira, C., Santos, B. S., & Barreto, S. (2013). A taxonomical analysis, current methods and objectives on location-routing problems. *International Transactions in Operational Research, 20*(6), 795-822.

Lübbecke, M. E., & Desrosiers, J. (2005). Selected topics in column generation. *Operations Research, 53*(6), 1007-1023.

Madsen, O. B., Ravn, H. F., & Rygaard, J. M. (1995). A heuristic algorithm for the a dial-a-ride problem with time windows, multiple capacities, and multiple objectives. *Annals of Operations Research, 60*(1), 193-208.

Mageean, J., & Nelson, J. D. (2003). The evaluation of demand responsive transport services in Europe. *Journal of Transport Geography, 11*(4), 255-270.

Malucelli, F., Nonato, M., Crainic, T. G., & Guertin, F. (2001). Adaptive memory programming for a class of demand-responsive transit systems. In S. Voss, & J. R. Daduna (Eds.), *Computer-Aided Scheduling of Public Transport* (Vol. 505, pp. 253-273). Berlin, Germany: Springer.

Mamasis, K., Minis, I., & Dikas, G. (2013). Managing vehicle breakdown incidents during urban distribution of a common product. *Journal of the Operational Research Society, 64*(6), 925-937.

Marottoli, R., Mendes de Leon, C., Glass, T., Williams, C., Cooney Jr., L., Berkman, L., et al. (1997). Driving cessation and increased depressive symptoms: prospective evidence from the New Haven EPESE. *Journal of the American Geriatrics Society, 45*(2), 202-206.

Melachrinoudis, E., Ilhan, A. B., & Min, H. (2007). A dial-a-ride problem for client transportation in a health-care organization. *Computers & Operations Research, 34*(3), 742-759.

Mercado, R., & Páez, A. (2009). Determinants of distance traveled with a focus on the elderly: a multilevel analysis in the Hamilton CMA, Canada. *Journal of Transport Geography, 17*(1), 65-76.

Metz, D. (2000). Mobility of older people and their quality of life. *Transport Policy, 7*(2), 149-152.

Metz, D. (2003). Transport policy for an ageing population. *Transport Reviews: A Transnational Transdisciplinary Journal, 23*(4), 375-386.

Mezzetti, M. G. (1998). Triage: military and civilian experience. *Current Anaesthesia and Critical Care, 9*(2), 48-51.

Miller, C., Tucker, A., & Zemlin, R. (1960). Integer Programming Formulation of Traveling Salesman Problems. *Journal of the ACM, 7*(4), 326-329.

Minis, I., & Tatarakis, A. (2011). Stochastic single vehicle routing problem with delivery and pick up and a predefined customer sequence. *European Journal of Operation Research, 213*, 37-51.

Mitrovic-Minic, S. (1998). *Pickup and Delivery Problem with Time Windows: A Survey.* Burnaby, Canada: Simon Fraser University.

Mitrovic-Minic, S., & Laporte, G. (2006). The pickup and delivery problem with time windows and transshipment. *Infor-Information Systems and Operational Research, 44*(3), 217-228.

Mladenović, N., & Hansen, P. (1997). Variable neighborhood search. *Computers and Operations Research, 24*(11), 1097-1100.

Mollaoğlu, M., Tuncay, F., & Fertelli, T. (2010). Mobility disability and life satisfaction in elderly people. *Archives of Gerontology and Geriatrics, 51*(3), e115-e119.

Mu, Q., Fu, Q., Lysgaard, J., & Eglese, R. (2011). Disruption management of the vehicle routing problem with vehicle breakdown. *Journal of the Operation Research Society, 2*(5), 225-234.

Murray-Tuite, P., & Wolshon, B. (2013). Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies, 27*, 25-45.

Musselwhite, C., & Haddad, H. (2010). Mobility, Accessibility and Quality of Later Life. *Quality in Ageing and Older Adults, 11*(1), 25-37.

Muter, İ., Birbil, S. I., & Sahin, G. (2010). Combination of metaheuristic and exact algorithms for solving set covering-type optimization problems. *INFORMS Journal on Computing, 22*(4), 603-619.

Nagy, G., & Salhi, S. (2007). Location-routing: Issues, models and methods. *European Journal of Operational Research, 177*(2), 649-672.

National Consortium on the Coordination of Human S. (2006). *Strategies in Emergency Preparedness for Transportation-Dependent Populations.* Washington, D.C.: U.S. Department of Transportation.

National Transit Database. (2013). *2012 National Transit Summaries and Trends.* U.S. Department of Transportation, Federal Transit Administration.

Nelson, J. D., Wright, S. M., Ambrosino, G., & Naniopoulos, A. (2010). Recent developments in Flexible Transport Services. *Research in Transportation Economics, 29*(1), 243-248.

Newbold, K. B., Scott, D. M., Spinney, J. E., Kanaroglou, P., & Páez, A. (2005). Travel behavior within Canada's older population: a cohort analysis. *Journal of Transport Geography, 13*(4), 340-351.

Nguyen-Hoang, P., & Yeung, R. (2010). What is paratransit worth? *Transportation Research Part A: Policy and Practice, 44*(10), 841-853.

Ninikas, G., & Minis, I. (2014). Reoptimization strategies for a dynamic vehicle routing problem with mixed backhauls. *Networks*, 64(3), 214-231.

Nourbakhsh, S. M., & Ouyang, Y. (2012). A structured flexible transit system for low demand areas. *Transportation Research Part B: Methodological*, 46(1), 204-216.

Pajonk, E. (1982). *Research Project" Telebus" for the Handicapped in Berlin (West)*. Gordon & Breach Science, NewYork: In Mobility and Transport for Elderly and Handicapped Persons.

Papadimitriou, C. H. (1977). The Euclidean travelling salesman problem is NP-complete. *Theoretical Computer Science, 4*(3), 237-244.

Paquette, J., Cordeau, J.-F., & Laporte, G. (2009). Quality of service in dial-a-ride operations. *Computers & Industrial Engineering, 56*(4), 1721-1734.

Parragh, S., & Schmid, V. (2013). Hybrid column generation and large neighborhood search for the dial-a-ride problem. *Computers & Operations Research, 40*(1), 490-497.

Parragh, S., Doerner, K., & Hartl, R. (2008a). A survey on pickup and delivery problems. Part I: Transportation between customers and depot. *Journal für Betriebswirtschaft, 58*(1), 21-51.

Parragh, S., Doerner, K., & Hartl, R. (2008b). A survey on pickup and delivery problems. Part II: Transportation between pickup and delivery locations. *Journal für Betriebswirtschaft, 58*, 81-117.

Petch, R., & Salhi, S. (2003). A multi-phase constructive heuristic for the vehicle routing problem with multiple trips. *Discrete Applied Mathematics, 133*(1-3), 69-92.

Pisinger, D., & Ropke, S. (2010). Large neighborhood search. In M. Gendreau, & J.-Y. Potvin (Eds.), *Handbook of metaheuristics* (2nd ed., pp. 399-419). Springer.

Prescott-Gagnon, E., Desaulniers, G., & Rousseau, L. M. (2009). A branch-and-price-based large neighborhood search algorithm for the vehicle routing problem with time windows. *Networks, 54*(4), 190-204.

Prins, C., Prodhon, C., & Wolfler-Calvo, R. (2006). Solving the Capacitated Location-Routing Problem by a GRASP complemented by a Learning Process and a Path Relinking. *4OR - A Quarterly Journal of Operations Research, 4*(3), 221-238.

Prodhon, C., & Prins, C. (2014). A survey of recent research on location-routing problems. *European Journal of Operational Research, 238*(1), 1-17.

Pugliese, L. D., & Guerriero, F. (2013). A survey of resource constrained shortest path problems: Exact solution approaches. *Networks, 62*(3), 183-200.

Quadrifoglio, L., & Dessouky, M. M. (2008). Sensitivity Analyses over the Service Area for Mobility Allowance Shuttle Transit (MAST) Services. *Computer-aided Systems in Public Transport Lecture Notes in Economics and Mathematical Systems, 600*, 419-432.

Quadrifoglio, L., Dessouky, M. M., & Ordonez, F. (2008). Mobility allowance shuttle transit (MAST) services: MIP formulation and strengthening with logic constraints. *European Journal of Operational Research, 185*, 481-494.

Quadrifoglio, L., Dessouky, M. M., & Palmer, K. (2007). An insertion heuristic for scheduling mobility allowance shuttle transit (MAST) services. *Journal of Scheduling, 10*, 25-40.

Quadrifoglio, L., Hall, R. W., & Dessouky, M. M. (2006). Performance and design of mobility allowance shuttle transit (MAST) services: Bounds on the maximum longitudinal velocity. *Transportation Science, 40*(3), 351-363.

Raidl, G. R., Puchinger, J., & Blum, C. (2010). Metaheuristic hybrids. In M. Gendreau, & J. Y. Potvin, *Handbook of metaheuristics. 2nd ed.* (pp. 469-496). Berlin: Springer.

Raitanen, T., Törmäkangas, T., Mollenkopf, H., & Marcellini, F. (2003). Why do older drivers reduce driving? Findings from three European countries. *Transportation research part F: traffic psychology and behaviour, 6*(2), 81-95.

Ralphs, T. K., & Galati, M. V. (2005). Decomposition in integer linear programming. In J. K. Karlof, *Integer Programming: Theory and Practice (vol. 3 of Operations Research Series)*. CRC Press.

Ravulaparthy, S., Yoon, S., & Goulias, K. G. (2013). Linking Elderly Transport Mobility and Subjective Well-Being: A Multivariate Latent Modeling Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 28-36.

Rekiek, B., Delchambre, A., & Saleh, A. (2006). Handicapped person transportation: An application of the grouping genetic algorithm. *Engineering Applications of Artificial Intelligence, 19*(5), 511-520.

Renne, J. L., Sanchez, T. W., & Litman, T. (2008). *National Study on Carless and Special Needs Evacuation Planning: A Literature Review*. New Orleans: Federal Transit Administration.

Renne, J. L., Sanchez, T. W., & Peterson, R. C. (2009a). *National Study on Carless and Special Needs Evacuation Planning: Case Studies*. New Orleans: Federal Transit Administration.

Renne, J. L., Sanchez, T. W., Jenkins, P., & Peterson, R. (2009b). Challenge of evacuating the carless in five major US cities. *Transportation Research Record: Journal of the Transportation Research Board, 2119*(1), 36-44.

Repoussis, P. P., Tarantilis, C. D., & Ioannou, G. (2007). The open vehicle routing problem with time windows. *Journal of the Operational Research Society, 58*(3), 355-367.

Righini, G., & Salani, M. (2008). New dynamic programming algorithms for the elementary shortest path problem with resource constraints. *Networks, 51*(3), 155-170.

Ropke, S., & Cordeau, J. F. (2009). Branch-and-Cut-and-Price for the Pickup and Delivery Problem with Time Windows. *Transportation Science, 43*(3), 267-286.

Ropke, S., & Pisinger, D. (2006). An Adaptive Large Neighborhood Search Heuristic for the Pickup and Delivery Problem with Time Windows. *Transportation Science, 40*(4), 455-472.

Ropke, S., Cordeau, J.-F., & Laporte, G. (2007). Models and Branch-and-Cut Algorithms for Pickup and Delivery Problems with Time Windows. *Networks, 49*(4), 258-272.

Rosenbloom, S. (2001). Sustainability and automobility among the elderly: An international assessment. *Transportation, 28*(4), 375-408.

Rosenbloom, S. (2004). The mobility of the elderly: there is good news and bad. In *Transportation in an Aging Society: A Decade of Experience, Technical Papers and Reports from a Conference* (pp. 3-21). Maryland: Transportation Research Board.

Salhi, S., & Petch, R. (2007). A GA based heuristic for the vehicle routing problem with multiple trips. *Journal of Mathematical Modelling and Algorithms, 6*(4), 591-613.

Sandia National Laboratories. (2004). *Identification and Analysis of Factors Affecting Emergency Evacuations: Main Report.* Washington, DC.

Savelsbergh, M. W., & Sol, M. (1995). The General Pickup and Delivery Problem. *Transportation Science, 29*(1), 17-29.

Sayyady, F., & Eksioglu, S. (2010). Optimizing the use of public transit system during no-notice evacuation of urban areas. *Computers & Industrial Engineering, 59*(4), 488-495.

Schönfelder, S., & Axhausen, K. W. (2003). Activity spaces: measures of social 12 exclusion? *Transport Policy, 10*(4), 273-286.

Schwanen, T., & Páez, A. (2010). The mobility of older people–an introduction. *Journal of Transport Geography, 18*(5), 591-595.

Schwanen, T., & Ziegler, F. (2011). Wellbeing, independence and mobility: an introduction. *Ageing and Society, 31*(5), 719-733.

Schwartz, M. A., & Litman, T. A. (2008). Evacuation Station: The Use of Public Transportation in Emergency Management Planning. *ITE Journal on the Web*, 69-73.

Schwarzlose, I. A., Mjelde, J. W., Dudensing, R. M., Jin, Y., Cherrington, L. K., & Chen, J. (2014). Willingness to pay for public transportation options for improving the quality of life of the rural elderly. *Transportation Research Part A: Policy and Practice, 61*, 1-14.

Shaw, P. (1998). Using constraint programming and local search methods to solve vehicle routing problems. *In: Proceedings CP-98.* Fourth international conference on principles and practice of constraint programming.

Sigurd, M., Pisinger, D., & Sig, M. (2004). Scheduling transportation of live animals to avoid the spread of diseases. *Transportation Science, 38*(2), 197-209.

Song, R., He, S., & Zhang, L. (2009). Optimum Transit Operations during the Emergency Evacuations. *Journal of Transportation Systems Engineering and Information Technology, 9*(6), 154-160.

Soumis, F. (1997). Decomposition and column generation. In M. D. Amico, F. Maffioli, & S. Martello, *Annotated Bibliographies in Combinatorial Optimization* (pp. 115-126). Chichester: John Wiley and Sons.

Spinney, J. E., Scott, D. M., & Newbold, B. K. (2009). Transport mobility benefits and quality of life: A time-use perspective of elderly Canadians. *Transport Policy, 16*(1), 1-11.

Stjernborg, V., Emilsson, U., & Ståhl, A. (2014). Changes in outdoor mobility when becoming alone in the household in old age. *Journal of Transport & Health, 1*, 9-16.

Su, F. (2007). *Understanding and satisfying older people's travel demand.* (Ph.D diss) London: Imperial College.

Su, F., & Bell, M. G. (2009). Transport for older people: Characteristics and solutions. *Research in transportation economics, 25*(1), 46-55.

Su, F., Schmocker, J. D., & Bell, M. G. (2008). *Mode choice of older people before and after shopping: study with London data.* (No. 08-0986): In Transportation Research Board 87th Annual Meeting.

Taillard, E. D. (1999). A heuristic column generation method for the heterogeneous fleet VRP. *RAIRO-Operations Research, 33*(1), 1-14.

Taillard, E. D., Laporte, G., & Gendreau, M. (1995). Vehicle routing with multiple use of vehicles. *The Journal of the Operational Research Society, 47*(8), 1065-1070.

Tang, H., & Miller-Hooks, E. (2005). A tabu search heuristic for the team orienteering problem. *Computer and Operations Research, 32*(6), 1379-1407.

Tarantilis, C. D., Zachariadis, E. E., & Kiranoudis, C. T. (2008). A hybrid guided local search for the vehicle-routing problem with intermediate replenishment facilities. *INFORMS Journal on Computing, 20*(1), 154-168.

Tatarakis, A., & Minis, I. (2009). Stohastic single vehicle routing with a predefined customer sequence and multi depot returns. *European Journal of Operation Research*(197), 557-571.

Tesfahuney, M. (1998). *Imag(in)ing the Others – Migration, Racism and the Discursive Construction of Immigrants.* (Ph.D diss.), Uppsala University.

Todorov, A., & Kirchner, C. (2000). Bias in proxies' reports of disability: data from the National Health Interview Survey on disability. *American Journal of Public Health, 90*(8), 1248-1253.

Toth, P., & Vigo, D. (1997). Heuristic algorithms for the handicapped persons transportation problem. *Transportation Science, 31*(1), 60-71.

Toth, P., & Vigo, D. (2002). *The vehicle routing problem.* Philadelphia: SIAM.

Transportation Research Board. (2008, TRB). *The Role of Transit in Emergency Evacuation.* Washington, DC: Transportation Research Board.

Tsiligirides, T. (1984). Heuristic methods applied to orienteering. *Journal of the Operational Research Society, 35*(9), 797-809.

Tsirimpas, P., Tatarakis, A., Minis, I., & Kyriakidis, E. (2007). Single vehicle routing with a predifined customer sequence and multiple depot returns. *European Journal of Operation Research*(187), 483-495.

U.S. Department of Homeland Security. (2006, DHS). *Report to Congress on Catastrophic Hurricane Evacuation Plan Evaluation.* Washington, DC.

U.S. Government Accountability Office. (2007, GAO). *Transportation-Disadvantaged Populations: Actions Needed to Clarify Responsibilities and Increase Preparedness for Evacuations.* Washington, D.C.: (GAO-07-44).

United Nations, Department of Economic and Social. (2013). *World Population Ageing 2013.* United Nations publication, ST/ESA/SER.A/348.

Urry, J. (2007). *Mobility.* Cambridge: Polity Press.

Vanderbeck, F. (2000). On Dantzig-Wolfe decomposition in integer programming and ways to perform branching in a branch-and-price algorithm. *Operations Research, 48*(1), 111-128.

Vansteenwegen, P., Souffriau, W., & Oudheusden, V. (2011). The orienteering problem: A survey. *European Journal of Operational Research, 209*(1), 1-10.

Wolsey, L. A. (1998). *Integer programming* (Vol. 42). New York: Wiley.

Wolshon, B. (2002). Planning for the evacuation of New Orleans. *Journal of the Institute of Transportation Engineers, 72*(2), 44-49.

Wolshon, B., Urbina, E., Levitan, M., & Wilmot, C. (2005b). Review of policies and practices for hurricane evacuation. II: Traffic operations, management, and control. *Natural Hazards Review, 6*(3), 143-161.

Wolshon, B., Urbina, E., Wilmot, C., & Levitan, M. (2005a). Review of policies and practices for hurricane evacuation. I: Transportation planning, preparedness, and response. *Natural hazards review, 6*(3), 129-142.

World Health Organization. (2011). *World report on disability 2011.* Geneva: WHO Press.

World Health Organization. (2014). *World health statistics 2014.* Geneva: WHO Press.

Yang, W., Mathur, K., & Ballou, R. (2000). Stochastic Vehicle Routing Problem with Restocking. *Transportation Science*(34), 99-112.

Zhang, X., & Chang, G. L. (2014). A Transit-Based Evacuation Model for Metropolitan Areas. *Journal of Public Transportation, 17*(3), 129-147.

Zheng, H. (2014). Optimization of bus routing strategies for evacuation. *Journal of Advanced Transportation, 48*(7), 734-749.

Ziegler, F., & Schwanen, T. (2011). 'I like to go out to be energised by different people': an exploratory analysis of mobility and wellbeing in later life. *Ageing and Society, 31*(5), 758-781.

Appendices

Appendix A: Generation of test data for scheduled paratransit transport systems

Appendix A.1 Introduction

In this Appendix we describe the creation of random test data for the case of Scheduled Paratransit Transport Systems (SPTS). SPTS focus on flexible public transport services that address the needs of the elderly, as well as of people with a disability. In this flexible transport services a bus operating in a public transport route may diverge from its nominal path to pick-up passengers with limited mobility and drop them off at their destination. The generator described in this Appendix aims to create appropriate test instances that simulate

practical problems within three distinct environments: (a) Urban, (b) Suburban, and (c) Rural. The data generator uses special techniques to produce test instances that incorporate the characteristics of each of the three aforementioned environments.

Appendix A.2 Random test data generator for the SPTS

The data generator for SPTS deals with four main aspects: (a) the locations of the regular bus stops, (b) the locations of the pick-up and delivery requests, (c) the latest time of arrival at each regular bus top, and (d) the earliest and latest drop-off times of the paratransit requests. The processes that address aspects (a), (c) and (d) are the same for all environments, while the procedure for aspect (b) differs depending on the type of environment. The generator constructs problem instances given the following parameters (inputs):

(a) m the number of regular bus stops
(b) k the number of the scheduled bus trips
(c) n the number of paratransit requests
(d) ρ the average distance between two sequential regular bus stops
(e) u the average vehicle speed
(f) f the time period between two consecutive bus trips (headway)
(g) d the coefficient of the allowed delay w.r.t. the nominal bus trip at each bus stop
(h) ξ the width of the drop-off time windows for the paratransit requests.

Regular Bus stop locations: The location of a regular bus stop is defined w.r.t. the location of the previous regular bus stop. The coordinates from this latter origin are the distance r and the angle θ, where r is randomly generated in the interval $[\frac{4}{5}\rho, \frac{6}{5}\rho]$ and θ is randomly generated in the interval $[-60^o, +60^o]$. Figure A.1 shows the region for the location of bus stop $i+1$ with respect to location of bus stop i. The location of the first bus stop is considered to be known.

Pick-up and drop-off locations of paratransit requests in the Urban environment: The allowable regions in which the pick-up and drop-off locations lie are ellipses. Each ellipsis has as foci two sequential regular bus stops and its longest diameter (major axis) is equal to ω. The related construction procedure for the generating a pick-up location is the following:

(a) A regular bus stop h is selected uniformly from $\{1, \dots, m-1\}$, where m is the last regular bus stop of the sequence

(b) The ellipse is created using the locations of regular bus stops h and $h+1$ as foci points, and

(c) The pick-up location is selected randomly from the area of the ellipse using a uniform distribution.

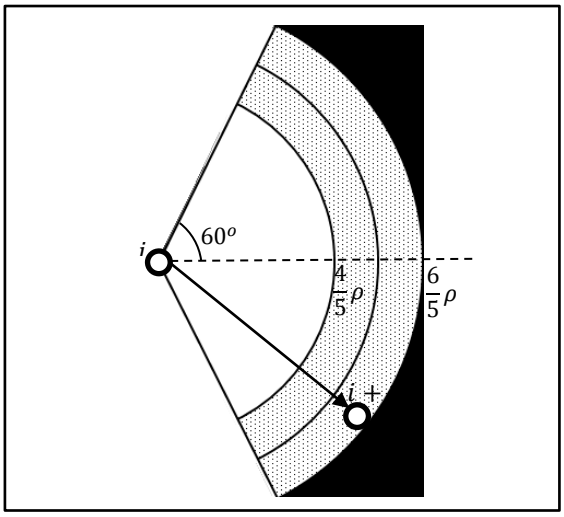


Figure A.1: Area for the locating bus stop $i+1$ w.r.t. to bus stop i

The drop-off locations are created by the same procedure, but in this case the reference bus stop g is selected randomly from the set $\{h+1, \dots, m\}$. Figure A.2 illustrates the geometry of the above procedure.

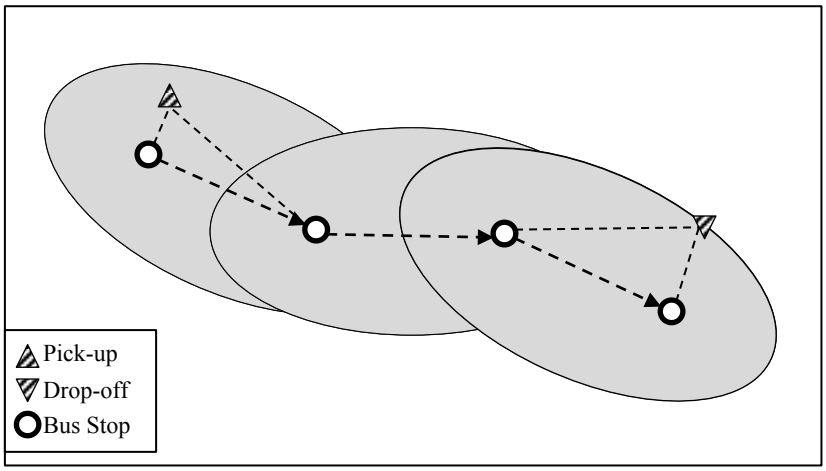


Figure A.2: Creation of a pick-up and a drop-off location in urban environment

Pick-up and drop-off locations of paratransit requests in a Suburban environment: A procedure similar to the one above is followed. However, in this case the selection of the reference bus stops h and g is different; h is selected from $\{1, \dots, m-2\}$ considering that the probability of selecting points $\left\{1, \dots, \left\lceil \frac{m}{4} \right\rceil\right\} \cup \left\{\left\lceil \frac{2m}{4} \right\rceil, \dots, \left\lceil \frac{3m}{4} \right\rceil\right\}$ is two times the probability of

selecting the remaining points in the set. g is selected randomly with respect to the reference bus stop h as follows:

- If the reference bus stop h is in the first half of the route ($h < \left\lceil \frac{m}{2} \right\rceil$), then g is selected from the set $\{h+1, \dots, \left\lceil \frac{m}{2} \right\rceil\}$, considering that the probability of selecting points $\{\left\lceil \frac{m}{4} \right\rceil, \dots, \left\lceil \frac{2m}{4} \right\rceil\}$ is two times the probability of selecting the remaining points in the set
- If the reference bus stops h is in the second half of the route ($h \geq \left\lceil \frac{m}{2} \right\rceil$), then g is selected from the set $\{h+1, \dots, m-1\}$, considering that the probability of selecting points $\{\left\lceil \frac{3m}{4} \right\rceil, \dots, m-1\}$ is two times the probability of selecting the remaining points in the set.

This attempts to model the behavior of passengers that travel from a sub-urban area to the center of the city (the first half of the route) or vice versa (the second half of the route). Figure A.3 illustrates the geometry of the regular bus stops, the pick-up and drop-off locations, and the ellipse areas for the suburban environment.

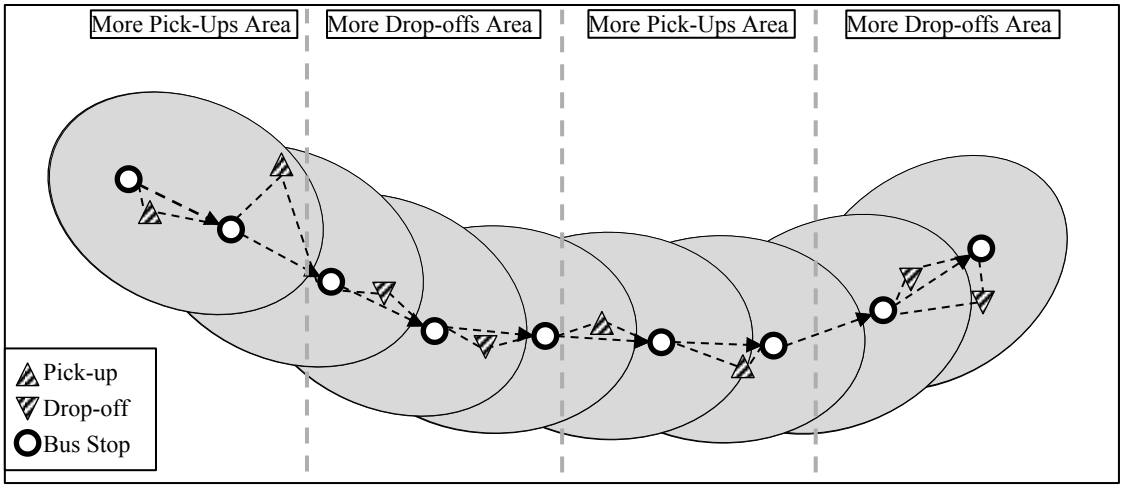


Figure A.3: Creation of a pick-up and a drop-off location for a suburban environment

Pick-up and drop-off locations of paratransit requests in a Rural environment: The selection process for the reference bus stops h and g is similar to the one used for the Urban environment. The difference lies in the allowable regions, which, in this case, are circles of radius $\bar{\omega}$ centered at the locations of the reference bus-stops, respectively. This construction procedure attempts to model the fact that residences in rural areas tend to be distributed around the center of the communities. Figure A.4 illustrates this process.

Earliest and latest times of arrival at a regular bus stop: The earliest time of arrival to a regular bus stop is the minimum time to travel from the route origin (first regular bus stop) to the one under consideration. The latest time of arrival is determined by adding to this earliest

time the time required to travel an additional distance $d \times \rho$, where ρ is the average distance between two sequential regular bus stops. Hence, d determines the service level for the regular bus line passengers.

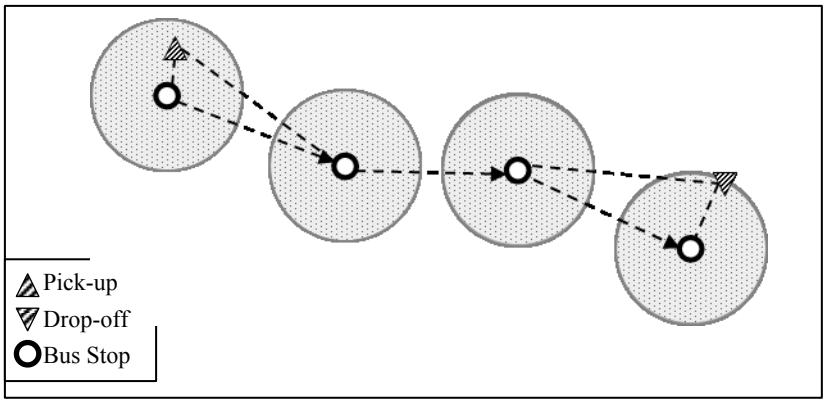


Figure A.4: Creation of a pick-up and a drop-off location in rural environment

Earliest and latest drop-off times of the paratransit requests: In the formulation of Section 3.1, only the earliest and latest drop-off times are defined, since the pick-up time is a decision variable related to the ride time. Thus, in the test cases of Section 3.4 the latest drop-off time is defined as follows:

- a random bus trip is selected, e.g. τ
- the minimum time to travel from the first regular bus stop to the bus stop that is closer to the drop-off location through the respective pick-up location is calculated, e.g. ζ
- the time from the latter bus stop to the drop-off location itself, e.g. φ
- then, the actual latest drop-off time will be equal to $\tau \cdot f + \zeta + 2 \cdot \varphi$, where f is the headway between consecutive trips.

The earliest drop-off time is defined as the difference of the latest drop-off time minus a service parameter ξ. This parameter defines the maximum allowable time to drop-off a paratransit customer before her/his requested time.

Appendix B: Tuning the probabilistic return to shelters (PRS) heuristic

As described in subsection 5.2.2, a vehicle may pay preemptive visits to shelters even if it is not full, depending on its location with respect to the shelters and the state of the system (shelter loads, vehicle loads, etc.). The decision to visit a shelter is made in a probabilistic way by the Probabilistic Return to Shelters (PRS) heuristic, which is a constructive heuristic that produces feasible routes.

In this Appendix firstly we study the tuning of the following key aspects of the proposed framework: (a) probabilities to return to a shelter $P_Q = e^{\frac{q-Q}{T}}$ and $P_D = e^{\frac{-\Delta t/\bar{t}}{aT}}$, (b) the appropriate combination of P_Q and P_D among $P_p = (P_Q \cdot P_D)$ and $P_m = \max\{P_Q, P_D\}$, and (c) the effect of T to the convergence of the framework. Note that parameter T depends on the state of the system as described below. For higher values of T, the probability to visit a shelter, P_p or P_m, is higher.

Subsequently, we examine two versions of the greedy and k-regret insertions: (a) one slower but more efficient, and (b) one faster but less efficient. The slower and faster versions differ in terms of the frequency of application of the PRS heuristic. Specifically, we examine how often we should execute the slow version and how often the fast version.

For the studies of the current Appendix, we used the test problems of Prins *et al.* (2006), and modified them appropriately to meet the BEP characteristics (i.e. by reducing the number of available shelters and the number of vehicles). Table B.1 shows the characteristics of the resulting problems.

Table B.1: Parameters of the test problems

Instance Name	Evacuees Sites	Shelters	Vehicles	Vehicle Capacity	Evacuees	Shelter Capacity
Coord20-5-1	20	5	5	70	315	350
Coord20-5-2	20	5	5	70	310	340
Coord50-5-3	50	5	12	70	756	840
Coord50-5-1	50	5	12	70	775	865
Coord50-5-2	50	5	12	70	761	850

Tuning the temperature and probability criterion of the PRS heuristic

Figure B.1 illustrates how P_Q changes as a function of vehicle occupancy for different values of T. It is observed that for higher values of T a preemptive visit to a shelter with fewer

evacuees on board is probable, while for lower values of T the vehicle will return only when the number of evacuees on board is close to its capacity.

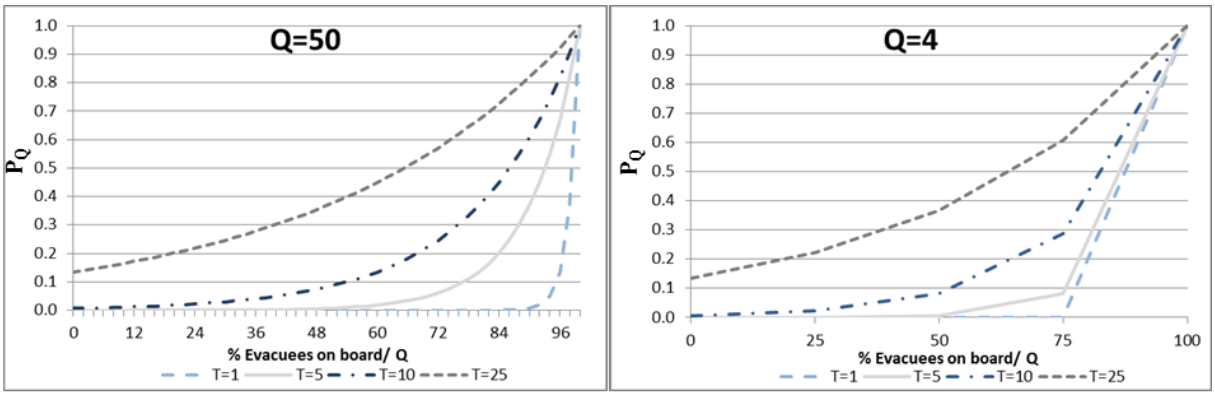


Figure B.1: Return probability P_Q as a function of evacuees on board for different values of the temperature T

Figure B.2 shows the variation of P_D with respect to the ratio $\Delta t/\bar{t}$, i.e. the time difference Δt between travelling to the next pick-up location of the route through the nearest shelter or directly, over the average travelling time $\bar{t}$ for different values of T. Visiting a remote shelter is more probable for higher values of T.

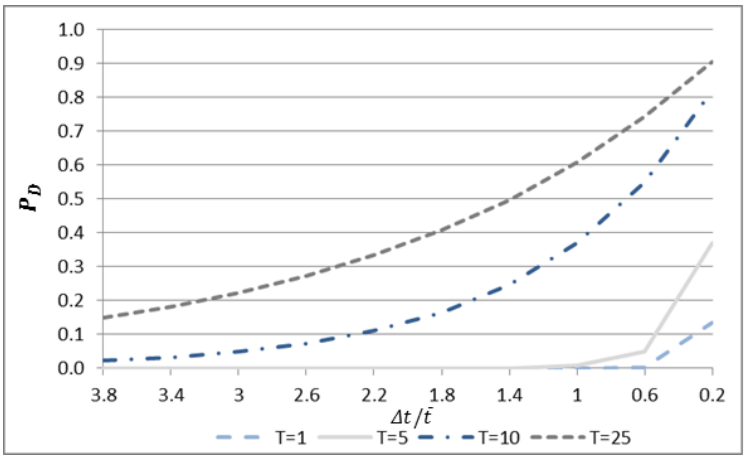


Figure B.2: Return probability P_D as a function of Δt over the average travelling time for different values of temperature T

Concerning the temperature, the issue of interest is the rate of convergence of HCG to the final solution. Note that in cases of emergency, the computational time may prove critical in decision making, and, therefore rapid convergence is a distinct advantage. For this study we skip Phase I of the framework and we focus on Phase II in which the value of T changes in every iteration. An initial solution is produced by the combination of the NN-BEP and one run of LNS.

Figure B.3 shows how the two terms of the objective function of BEP are decreasing from the initial solution within 25 iterations of Phase II when: (a) T starts from value 1 and increases

by 1 in every iteration of the main algorithm, and (b) T starts from the value 25 and decreases by 1 in every iteration of the main algorithm. It seems that starting from lower values of T leads to a better solution in less iterations, i.e. less computational time. Thus, the schedule used throughout the framework involves: (a) for Phase I, $T = 1$; and (b) for Phase II, initial value T=1, and then increased by 1 in every iteration of Phase II.

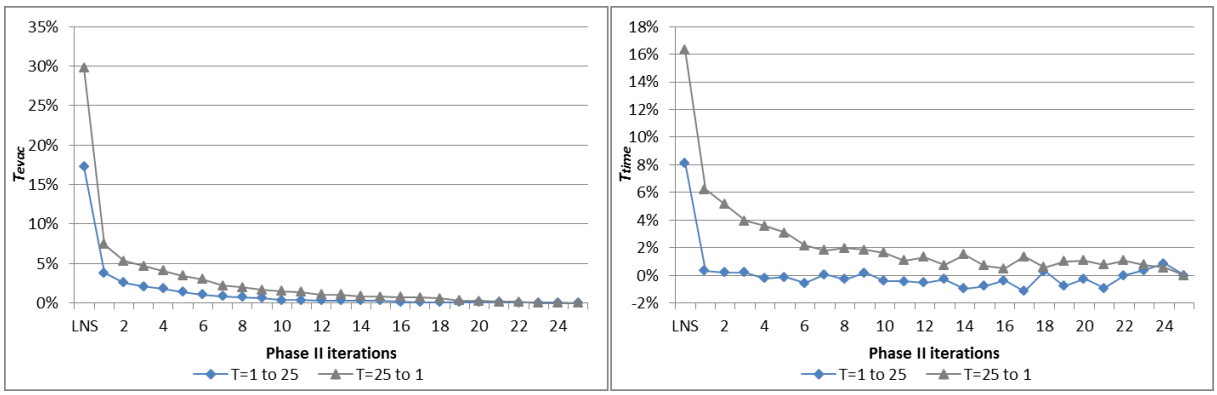


Figure B.3: Convergence of the solution with respect to temperature change

Concerning the probabilistic criterion, Fig. B.4 presents the efficiency of HCG under two options: The first, P_p, uses the product $P_Q \cdot P_D$ as the criterion of step (d) of the PRS heuristic described in Section 3.2. The second, P_m, uses $max(P_Q, P_D)$. The efficiency of HCG is presented regarding the two terms of the objective function of BEP, i.e. the duration of the evacuation T_{evac} and the total routing time T_{time}. Note that since the terms of the objective function are lexicographically ordered (first objective T_{evac}, second T_{time}), the (combined) value of the objective function is in line with the values of T_{evac}.

From the results of Figure B.4 it seems that the P_p probability function achieves better results in all cases, and is overall more effective than the P_m probability function. Thus, the former is used throughout this study.

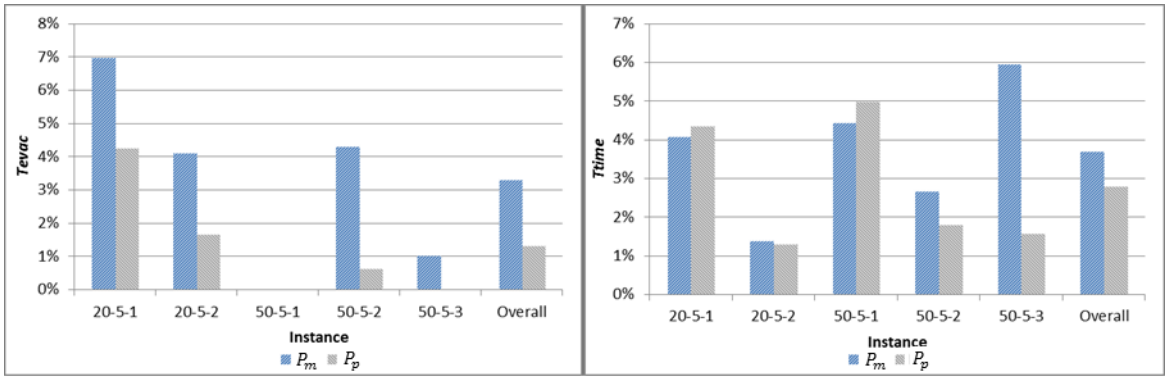


Figure B.4: Average deviation from the best solution for the two probability functions

Tuning the frequency of implementation of the PRS heuristic

Figure B.5 displays results with respect to the frequency of application of the Probabilistic Return to Shelters (PRS) heuristic (described in subsection 5.3.2). Specifically, we examine two versions of the greedy insertion and the *k*-regret insertion. In the first version (FV) we consider that PRS is applied for every possible insertion position, while in the second version (SV) we consider that the insertions are performed initially without PRS (note this may lead to temporary infeasible insertions), and PRS is applied subsequently after the selection of the preferable insertion position. The second version is expected to require less computational time. Of course in FV the best insertion is selected among feasible paths, while in SV the best insertion may involve infeasible solutions. This difference may affect HCG efficiency.

To study this difference we tested the following FV-SV application combinations to solve the instances of Prins *et al.* (2006) described in Table B.1:

i. Apply only FV
ii. Apply FV for 50% of insertions and SV for the rest 50%
iii. Apply FV 10% of the insertions and SV for the rest 90%

Five runs of each combination case were performed for each instance, and Figure B.5 reports the average deviation of each case from the best identified solution. It seems that all settings produce solutions of similar quality and, as expected, the computational time decreases as the usage of SV increases. Based on these results in the HCG framework we applied the (10%-90%) combination, since it requires significantly less computational time while it produces efficient solutions.

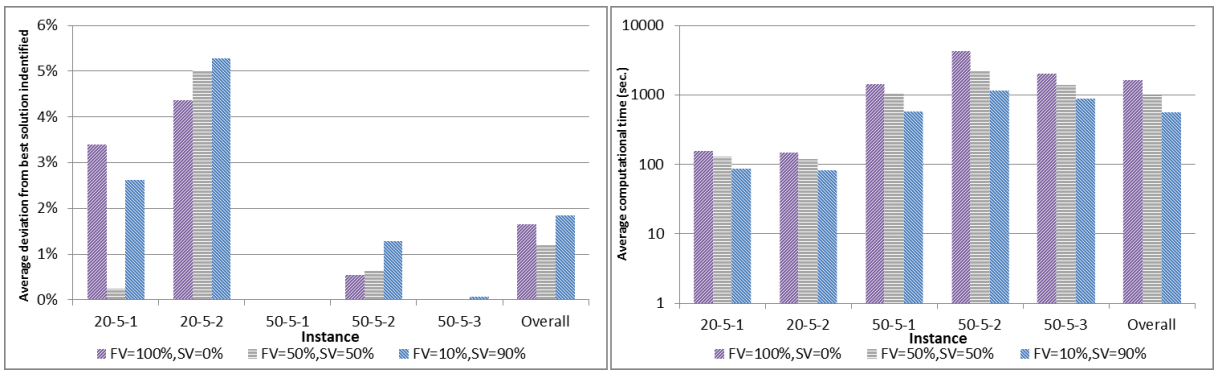


Figure B.5: Impact of the frequency of feasibility check to the efficiency and performance of the total framework