Agent-based Modelling of Paratransit as an Intelligent Complex Adaptive System to Improve Efficiency

by

Innocent Ndibatya

Dissertation presented in fulfilment of the requirements for the degree of Doctor of Philosophy in Engineering at Stellenbosch University

Promoter : Prof. M.J. (Thinus) Booysen

March 2021

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Plagiaatverklaring / *Plagiarism Declaration*

1. Plagiaat is die oorneem en gebruik van die idees, materiaal en ander intellektuele eiendom van ander persone asof dit jou eie werk is.
   Plagiarism is the use of ideas, material and other intellectual property of another's work and to present it as my own.

2. Ek erken dat die pleeg van plagiaat 'n strafbare oortreding is aangesien dit 'n vorm van diefstal is.
   I agree that plagiarism is a punishable offence because it constitutes theft.

3. Ek verstaan ook dat direkte vertalings plagiaat is.
   I also understand that direct translations are plagiarism.

4. Dienooreenkomstig is alle aanhalings en bydraes vanuit enige bron (ingesluit die internet) volledig verwys (erken). Ek erken dat die woordelikse aanhaal van teks sonder aanhalingstekens (selfs al word die bron volledig erken) plagiaat is.
   Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism

5. Ek verklaar dat die werk in hierdie verhandeling vervat, behalwe waar anders aangedui, my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik ingehandig het vir om enige kwalifikasie te verwerf.
   I declare that the work contained in this dissertation, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for obtaining any qualification.

Abstract

Agent-based Modelling of Paratransit as an Intelligent Complex Adaptive System to Improve Efficiency

I. Ndibatya

Department of Electrical and Electronic Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Dissertation: PhD (Eng)

March 2021

Urban residents in Sub-Saharan Africa (SSA) face mobility challenges that limit their access to jobs, services, markets, and socioeconomic opportunities. In most SSA cities, public transport is predominantly provided by the inefficient paratransit system – a flexible mode of passenger transport consisting of privately-owned, low-capacity unscheduled minibuses and motorcycle taxis. There is growing interest among city authorities and urban transport researchers in addressing the inefficiency problem associated with paratransit. Several approaches, such as complete overhaul to bus rapid transit (BRT), and phased banning of paratransit from the cities have previously been proposed and concomitant implementation projects started. However, most of such projects have either failed to take off, or they have stalled. This is likely because of the huge capital investment required, the unique social and cultural dynamics associated with "third world" countries, and urban sprawl due to poor city planning. This study departs from the common perspective held by several researchers and city authorities who view paratransit as "chaotic", thus, the justification for its total overhaul and banning. Instead, this study aims to leverage the beneficial aspects of existing paratransit – such as flexibility, demand-responsiveness and near-ubiquitous coverage – with the elusive objective of achieving a more efficient paratransit state as a result.

Through theoretical modelling, field study and experimental approaches, this study aimed to improve the efficiency of minibus taxis paratransit systems. The theoretical modelling work involved modelling paratransit systems as complex adaptive systems (CAS) and developing an agent-based model (ABM) for minibus taxi operations in an organically-evolved paratransit setting. The field study involved an in-depth investigation of minibus taxi operations in Kampala's paratransit system, and collection and analysis of minibus taxi movement data that was used to validate the agent-based model. The experimental approaches involved three separate simulation experiments, simulating the minibus taxi transportation dynamics with varying levels of agents' intelligence and situational awareness. Machine learning methods, such as random forests and convolutional neural networks were used to train agents in the subsequent simulation experiment to improve their intelligence during decision making. At each stage, several efficiency metrics' values such as passenger waiting time and minibus taxi occupancy were collected. The results from the experiments showed that there was an improvement in the overall efficiency of the minibus taxi paratransit system. For instance, the average passenger waiting time reduced from 1.2 hours to 30 minutes, indicating a 55% improvement. Whereas the average minibus taxi occupancy increased from 42% to 51%, indicating a 21% improvement. Accordingly, we concluded that improving the micro-level agents' intelligence and situational awareness, results in an overall increase in the efficiency of the paratransit system.

To the transportation researchers, we recommend further work on using ABM to include other modes of paratransit transport such as the three-wheeled rickshaws and motorcycle taxis (boda bodas). To the city authorities, we recommend the integration of smart mobility and ICT applications into the paratransit ecosystem to support journey planning, booking, scheduling, and fare collection.

Key words:
Modelling paratransit; Complex adaptive systems; Agent-based modelling

Uittreksel

Paratransit modellering as 'n intelligente, aanpasbare stelsel om doeltreffendheid te verbeter

I. Ndibatya

Departement Elektroniese en Elektriese Ingenieurswese,
Universiteit van Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Verhandeling: PhD (Ing)

March 2021

Stedelike inwoners in Afrika Suid van die Sahara (SSA) word gekonfronteer met mobiliteitsuitdagings wat hul toegang tot werk, dienste, markte en sosio-ekonomiese geleenthede beperk. In die meeste SSA-stede word openbare vervoer oorheersend aangebied deur die ondoeltreffende paratransit-stelsel -'n buigsame manier van passasiersvervoer wat bestaan uit private lae-volume busse en huur-motorfietse. Daar is toenemende belangstelling onder stadsowerhede en navorsers van stedelike vervoer om die ondoeltreffendheidsprobleem so eie aan paratransit. Verskeie benaderings, soos byvoorbeeld die volledige opknapping van busvervoer (BRT) en 'n gefaseerde verbod op paratransit in stede, is voorheen voorgestel en verwante implementeringsprojekte is van stapel gestuur. Die meeste van hierdie projekte het egter nie daarin geslaag om te begin nie, of hulle is gestaak. Dit is waarskynlik as gevolg van die groot kapitaalinvestering wat benodig word, die unieke sosiale en kulturele dinamika wat verband hou met lande van die "derde wˆereld" en stedelike uitbreiding as gevolg van swak stadsbeplanning.

Hierdie studie wyk af van die algemene perspektief wat deur verskeie navorsers en stadsowerhede gehou word, wat paratransit as "chaoties" beskou, en dus van die regverdiging vir die totale opknapping en verbod daarvan. In plaas daarvan beoog hierdie studie om die voordelige aspekte van paratransit te versterk - soos buigsaamheid, aanvraagresponsiwiteit en byna alomteenwoordige dekking - met die hoop om 'n doeltreffender paratransit-staat as gevolg daarvan te bewerkstellig.

Deur middel van teoretiese modellering, veldstudies? en eksperimentele benaderings, het hierdie studie ten doel gehad om die doeltreffendheid van minibustaxi-paratransitstelsels te verbeter. Die teoretiese modelleringswerk behels die modellering van paratransitstelsels as komplekse aanpasbare stelsels (CAS) en die ontwikkeling van 'n agent-gebaseerde model (ABM) vir minibustaxibedrywighede in 'n organies-ontwikkelde paratransit-omgewing. Die veldstudie behels 'n diepgaande ondersoek na minibustaxibedrywighede in Kampala

se paratransit-stelsel, en versameling en ontleding van minibustaxibewegingsdata wat gebruik is om die agent-gebaseerde model te bekragtig. Die eksperimentele benaderings het drie afsonderlike simulasie-eksperimente behels, wat die minibustaxi-vervoerdinamika met verskillende vlakke van agente se intelligensie en situasiebewustheid simuleer. Masjienleermetodes soos ewekansige woude en evolusionˆere neurale netwerke is gebruik om agente in die daaropvolgende simulasie-eksperiment op te lei om hul intelligensie tydens besluitneming te verbeter. In elke stadium is verskeie waardes vir doeltreffendheid soos die passasierswagtyd en die besetting van minibustaxi's versamel. Die resultate van die eksperimente het getoon dat die algehele doeltreffendheid van die minibus-paratransitstelsel verbeter het. Byvoorbeeld, die gemiddelde passasierwagtyd verminder van 1,2 uur tot 30 minute, wat dui op 'n 55% verbetering. Terwyl die gemiddelde besetting van minibustaxi van 42% tot 51% gestyg het, wat dui op 'n verbetering van 21%.

Gevolglik het ons tot die gevolgtrekking gekom dat die verbetering van die intelligensie en situasiebewustheid van die mikrovlakagente tot 'n algehele toename in die doeltreffendheid van die paratransitstelsel lei.

Vir die vervoernavorser beveel ons verdere ABM-werk aan om ander maniere van paratransit-vervoer soos die driewiel-riksja's en motorfiets-taxi's (boda bodas) in te sluit. Aan die stadsowerhede beveel ons die integrasie van slim mobiliteit- en IKT-toepassings aan in die paratransit-ekosisteem om reisbeplanning, bespreking, skedulering en tariefinvordering te ondersteun.

Key words:

Paratransit van modellering; Komplekse aanpasbare stelsels; Agent-gebaseerde modellering

Acknowledgements

First and foremost, I give thanks and praise to the Almighty God. He has blessed me beyond measure!

I want to thank my supervisor, Professor M.J. (Thinus) Booysen, whose expertise, guidance, encouragement, and generosity was invaluable. Your insightful and timely feedback improved my work amazingly.

I would like to thank the MTN Mobile Intelligence Lab, and the MIH Media Lab at Stellenbosch University for hosting me and providing a conducive environment in addition to financially supporting the various components of my PhD research.

My sincere gratitude goes to Prof. Ddumba Ssentamu, Prof. Gilbert Maiga, Prof. Sarah Ssali, Prof. Ireeta Winston Tumps, Dr. John Kitayimbwa, and Dr. Evelyn Kigozi Kahiigi from Makerere University. Thank you for the support and encouragement during this period.

I would like to acknowledge the contributions made to my academic journey by the following: Prof. Patrick Mangheni (Undergraduate University professor), Mr. Alex Kigwaana (High School teacher), Mr. Patrick James Odong (Secondary School teacher), and Mrs. Hellen Ganyana (Primary School teacher). You provided me with the foundation, the tools, and the guidance that I needed to pursue a successful academic career. You are the "giants" on whose shoulders I stood to explore the academic horizons. May the Almighty God bless you abundantly.

In a special way, I would like to thank my family: Mrs. Jacqui Forde, Mr. Alex Kikoola, Ms. Betty Forde Nansubuga, Ms. Belinder Achieng Okello, Mr. Samuel Kanaye, Dr. Kankaka Edward Nelson, and Frode Mikisa. I am forever indebted to you for your kindness, encouragement, generosity, and thoughtfulness.

To my friends, Dr. Matia Mukama, Mr. Julius Uwimana, Mr. Linoln Talwaana, Mr. John Wabwiire, Mr. Paul Bakaki, Ismael Kaggwa, Mr. Reagan Besimba, Mr. Nasser Kimbugwe, Ms. Flavia Namagembe, and Ms. Justine Mukalele, thank you for being loyal, dependable, and supportive throughout this journey. Proverbs 18:24 "There are 'friends' who ruin each other, but real friends stick closer than brothers."

To Ms. Celene Hunter and Dr. Di Kilpert, thank you for supporting the various stages of this research. I wish to thank Prof. Mark Zuidgeest, Prof. Steven Jones and Dr. Herman Kamper for the insightful feedback.

Dedication

To the loving memories of

Omumbejja. Elizabeth Luwedde Nakimera

and

Rev. Wesley Forde

May their loving souls rest in eternal peace!

Contents

List of Figures

List of Tables

Nomenclature

Variables

v_c	Commercial speed $[km/h]$
d_{l1}	First leg distance $[km]$
t_h	Hold-back time $[h]$
d_l	Intermediate legs distance $[km]$
d_{ln}	Last leg distance $[km]$
l_{count}	Legs count $[-]$
$\mathcal{O}$	Occupancy $[\%]$
v_o	Operating speed $[km/h]$
d_T	Trip distance $[km]$
t_w	Waiting time $[h]$
l	Lévy step length $[km]$
ℓ	Trajectory spatial distance $[-]$
τ	Trajectory tortuosity $[-]$

List of Acronyms

ABM agent-based modelling. 61

ABM agent-based model. xi, 9, 17, 61–65, 69, 79, 80, 87, 98, 100

ACM arrival choice model. 68

BCM boarding choice model. 67

BRT bus rapid transit. 2, 3

CAS complex adaptive system. 7

CER controlled experiment. ix, 87, 99, 117

CNN convolutional neural network. 10, 100, 117

GKMA Greater Kampala Metropolitan Area. 2

GPS Global Positioning System. 30

ISM initial stop model. 67

ITS intelligent transport systems. 118

JICA Japan International Cooperation Agency. 70

KOTSA Kampala Operational Taxi Stages Association. 29

LRT light rail transit. 2, 3

LW Lévy walk. 8

MBT minibus taxi. 9, 61

NDCG Normalised Discounted Gain. 23

PI profitability index. 8

PTM passenger touting model. 68

QGIS quantum geographical information system. 58

RCM route choice model. 68

SSA Sub-Saharan Africa. 2–5, 11, 27, 28, 110

TOR1 test experiment one. 98, 112, 117

TOR2 test experiment two. 98, 112, 117

UTODA Uganda Taxi Operators and Drivers Association. 29

Chapter 1

Introduction

1.1 Motivation

Public transport systems in many developing cities of the Global South are messy, complex, and inefficient. Yet they adapt and serve, though inefficiently, the mobility needs of the often over-populated, poorly-planned cities. Congestion, informality, and inefficiency are the major defining features of numerous public transport systems in the Global South (Pojani and Stead, 2017, 2018; Venter et al., 2018). These transport systems comprise many competing actors that render the systems operationally, economically, and politically complex (Goodfellow, 2017).

Perhaps the essential feature of public transport in the Global South is the widespread presence of "informal transport" or "paratransit", which refers to "a flexible mode of passenger transportation that does not follow fixed schedules," (Behrens et al., 2015a). The paratransit system consists of shared-ride, demand-responsive privately-owned vehicles like minibus taxis in Manila, Lagos, Johannesburg, Nairobi and Kampala; as well as single-passenger vehicles like Kampala's motor-cycle taxis (also referred to as "boda bodas"), and Nairobi's tri-cycle taxis (also referred to as "tuk-tuks") (Mutiso and Behrens, 2011; Booysen et al., 2013; Diaz Olvera et al., 2019). Travel by paratransit contributes approximately 70%, 90%, 91%, and 98% of the road-based public trips in Johannesburg, Lagos, Kampala, and Dar es Salaam, respectively (Behrens et al., 2015a; Evans et al., 2018).

1.1.1 Background

The organic emergence of paratransit systems in the Global South, and their subsequent evolution into fully-fledged quasi-demand-responsive transportation systems is attributed to: the collapse of state-owned transportation enterprises in the 1990s; the World Bank structural adjustment policies (SAPs) of the 1990s; and weak policies for regulating and enforcing paratransit licensing (Kumar, 2011). The SAPs restricted financing to state-owned entities leading to their eventual collapse, thus creating a transportation gap that was filled (initially as complementary) by the loosely regulated, privately-run paratransit services in response to demand. Paratransit has organically-evolved from being a mere complimentary transport service to a way of life for the urban poor in several African cities and the Global South in general (Ajay Mahaputra Kumar et al., 2008). Paratransit is firmly entrenched, primarily due to poor planning by the authorities that has resulted in:

urban sprawl; ‘random’ distribution of public amenities (such as schools, hospitals, and shopping centres); and poor road infrastructure that limits general access to the transport system. Furthermore, paratransit thrives on the irregular daily mobility characteristics of people in developing cities (Kumar, 2011). This is because most of the population in these cities are not formally employed. As a result, their transport schedules and destinations are highly irregular. Thus, in addition to being the only main transport service commuters have access to, paratransit is considered flexible and affordable for the urban poor, and it has a near-ubiquitous coverage to the remote unreachable and highly fragmented settlements in cities.

Despite the massive popularity of paratransit among the urban commuters in Africa, it thrives at high health, environmental and economic cost to the users. Paratransit in Africa substantially contributes to environmental pollution, deaths by crashes, commuter stress, and huge economic losses to the population. In Kampala, for example, a commuter loses 25.4 minutes for every hour travelled using paratransit, and a total of 1.457 million man-hours are lost by commuters every day in the Greater Kampala Metropolitan Area (GKMA) (JICA, 2010; ITP, 2010). The situation is not any better in other developing cities of the Global South (Behrens et al., 2015a). Therefore, paratransit is often described by scholars as inefficient, unsafe, and sometimes dangerous to the commuters (Woolf and Joubert, 2013; Ndibatya and Booysen, 2020a). Consequently, several city authorities have interested themselves in regulating paratransit (Jennings and Behrens, 2017), while others have proposed a total paratransit ban in favour of modern services such as bus rapid transit (BRT) (Plano et al., 2020), a policy that is vigorously resisted by paratransit operators, and often leads to strikes, that seldomly transgress into riots, such as the taxi riots in Cape Town (Bähre, 2014), Kampala (Spooner et al., 2020), and Malawi (Tambulasi and Kayuni, 2008).

Cities in the Global South are experiencing unprecedented growth. It is estimated that by 2050 ninety per cent of urban growth will be concentrated in Africa and Asia (Jahan, 2016), and eighty per cent of the population in these continents will reside in cities (Awumbila, 2017). At the same time, public authorities have announced intentions to reform, phase out, or ban paratransit in favour of more modern, high-capacity forms of public transit, such as, light rail transit (LRT) and BRT (Ommeh et al., 2015). Throughout the last two decades, there has been a growing agitation among authorities in Africa advocating for paratransit-to-BRT transition. These crusades are in part instigated by the international donors, who make such transition plans prerequisites for city development loans (Gauthier and Weinstock, 2010). Consequently, since the early 2000s, the number of BRT projects have increased in the African cities of Lagos, Dar es Salaam, Nairobi, Kampala, Cape Town and others (Nkurunziza et al., 2012; Vermeiren et al., 2015; Wood, 2015). However, except for a few success stories from South Africa, Addis Ababa, and Rwanda, most BRT projects, especially in Sub-Saharan Africa (SSA) are struggling. This is mainly because of the high BRT infrastructure costs, corruption by the implementing authorities, slow local acceptance, inadequate infrastructure maintenance, and poor integration with existing paratransit (Wood, 2015). In many SSA cities, for example, BRT rollout was planned on a corridor-by-corridor basis following the standard BRT design (Gauthier and Weinstock, 2010; JICA, 2008, 2014; MoWT and Transport, 2009). However, ten years later, many of the BRT projects have not taken off (e.g. Kampala KCCA (2016)), and those that have taken off have not realised the full benefits that come with a BRT (e.g. in Dar es Salaam) leaving paratransit still dominating public transportation, its downsides notwithstanding.

Experts agree that many urban dwellers in the Global South face travel conditions that limit their ability to live healthy and productive lives. For many, travelling to work, to school, to medical facilities, or for social activities requires long or unsafe walks, long waits between poorly connected services in inconvenient locations, or expensive trips in unsafe and uncomfortable vehicles. Venter et al. (2019) categorises urban commuters in the Global South into four categories, i.e. the *mobile under-served*, *stranded under-served*, *well-located commuters*, and *well-located urbanites*. The under-served commuters (who are the majority in SSA) face severe transport constraints, with many commuting by foot or bicycles for longer distances, and others spend above average amounts of time and money on commuting (Venter et al., 2019; Ndibatya and Booysen, 2020a). They (the experts) also rightly believe that the existence of efficient public transportation is the key to solve mobility problems. However, most of the solutions authorities in the Global South are pursuing are not feasible for the context of their cities because they are costly, unsustainable, and they do not guarantee an efficient public transport system for the future vulnerable urban poor. Most of these solutions are copied (as-they-are) from the Global North cities in Europe and North America, and 'pasted' in the African underdeveloped cities whose social contexts are completely different. Often, heading the transport transformation agenda is the complete system overhaul (often phased) in favour of fancy modern Metro, LRT, and BRT systems that require a long time (4-10 years) to implement at a very high cost. For instance, the proposed BRT corridors planned for Kampala are estimated to cost $14 million per kilometre and will take eight years to complete (KCCA, 2016). A close analysis of the SSA cities master plans reveals the absence of a clear plan for paratransit in these cities, save for a few excerpts acknowledging how paratransit is chaotic, unsafe, and unreliable. We found no policy document bearing a comprehensive plan to organise and transform paratransit (Ndibatya and Booysen, 2020b). Whereas transport problems are universal, culture and circumstances vary per country. It is thus, imperative that authorities and researchers study the unique characteristics of the organically-evolved paratransit in the Global South and suggest economically feasible, sustainable, and socially acceptable solutions that incorporate the current technological trends and scientific research.

1.1.2 The proposed paradigm shift

Three key shifts in thinking can fundamentally reform paratransit in the Global South:

1. A paradigm shift from perceiving paratransit in developing cities as an alienating agent of public transport, to considering it as an integrated complimentary service with flexibility, adaptability, and near-ubiquitous coverage as its main assets.

2. Deconstruct the 'colonial mentality' of copying and pasting BRT design and implementation of the Global North style, to redesigning BRT corridors that fit into the existing road infrastructure, encouraging shared lanes in addition to low-cost semi-permanent bus stop platforms and paratransit integration nodes, or exchange centres.

3. Explore investment in smart mobility information and communications technology (ICT) systems in all aspects of the transportation service ecosystem to meet the

> rapidly increasing mobility needs of the developing cities. Smart mobility seamlessly integrates all modes of transport (first-to-last mile of commute) via wireless communications. It applies real-time data analytics and machine learning to make transportation safer and more efficient. There is evidence suggesting that smart mobility has the potential to enable safe, sustainable, and efficient movement of people, thus saving lives, reducing pollution and saving commuters billions of dollars annually.

Cognisant of the technological disruption wave that has swept over cities in the last decade, developing cities in the Global South must not only find more ways to integrate existing paratransit systems, they must also strive to understand the behaviour of passengers who rely on these services every day. This is possible with macro-level integration of smart mobility ICT systems in transportation and paratransit in general.

This dissertation takes into consideration the first and third fundamental shifts in thinking (mentioned earlier) to study the complex nature of organically-evolved paratransit systems with a focus on minibus taxis (which is the significant component of paratransit) in Sub-Saharan Africa, specifically in Uganda's capital, Kampala. First, we quantitatively investigated the minibus taxi operations and movement characteristics in Kampala, intending to understand the main actors in the system; scientifically describing their movement patterns, and estimating (at macro-level) the efficiency of the minibus taxi system as part of the broader paratransit system. Second, we modelled the minibus taxi paratransit system as a complex adaptive system (CAS) to scale up and simulate micro-level behaviour of individual actors (referred to as agents) in the minibus taxi systems. Lastly, we incorporated elements of smart mobility into the minibus taxi simulation and used machine learning methods to optimise selected metrics, hence, improving individual agents' situational awareness. Then we compared the macro-level efficiency gain because of improved agents' situational awareness, with the results where interacting agents had low situational awareness.

Results from this dissertation brings us a step closer to sustainably solving the efficiency problem affecting the mobility of the urban poor in developing cities of Africa who depend on paratransit systems for their daily mobility needs.

1.2 Paratransit, complex adaptive systems (CAS) and agent-based models (ABM)

In this section, we highlight the relationship between paratransit, complex adaptive systems, and agent-based models as used in this dissertation.

Paratransit emerged because of private initiatives that developed spontaneously in Global South cities as a stop-gap solution to the inadequacy (or lack) of institutional transport. It organically evolved into a transport structure composed of many fragmented and connected, formal and informal actors that operate autonomously (or independently) without a proper form of centralised coordination. Thus, paratransit can be described as a *complex system.*

Despite the complex nature of paratransit, it *adapts* in a pragmatic way to the local context in many Global South cities where the institutional framework is inadequate and where urban sprawl, topography, poor road infrastructure, and lack of funds are obstacles to the development of large-bus services. The complexity and adaptive properties of paratransit prompted us to model it as a *complex adaptive system (CAS)*.

A CAS is a dynamic system that represents individual agents and their collective behaviour. An agent-based model (ABM) is a computational simulation model of a many-agent system that captures the behaviours of the system's autonomous agents and their interactions with each other (Blume, 2015). In other words, an ABM is a computational instantiation of a complex adaptive system (CAS). In this research, we designed and implemented intelligent agents with optimised situational awareness, thus qualifying our approach as the intelligent complex adaptive systems (ICAS) approach.

1.3 Research focus

This research focusses on: analysing the operations of minibus taxis in Kampala's organically-evolved, quasi-demand-responsive paratransit system; modelling the minibus taxi system as a complex adaptive system; and simulating and optimising the minibus taxi system to achieve efficiency at a macro-level. This gives rise to the following research questions, the corresponding objectives, and the associated original contributions.

1.3.1 Research questions

Five research questions were formulated and investigated in this study:

RQ 1: How do minibus taxis operate in organically-evolved, quasi-demand-responsive paratransit systems?

RQ 2: Are their operations efficient?

RQ 3: How do individual-level operations and autonomous interactions between minibus taxis and passengers shape the higher-level (macro-level) system behaviour in an organically-evolved, quasi-demand-responsive paratransit system?

RQ 4: What metrics can be used to measure efficiency in such a system?

RQ 5: What is the macro-level effect on system efficiency, of intelligent routing of autonomous and situationally aware minibus taxi agents with self-selected origins and destinations in an organically-evolved, quasi-demand-responsive paratransit system?

1.3.2 Hypothesis

This research will test the following hypotheses:

1. The transportation dynamics of organically-evolved paratransit systems in Sub-Saharan Africa are shaped by local interactions of autonomous agents at the micro-level of the system giving rise to a stable (often inefficient) state at macro-level through demand and supply.

2. Improving the intelligence and situation awareness of autonomous agents in organically-evolved paratransit systems leads to agents adapting more optimal travel behaviour resulting in improved macro-level paratransit system efficiency.

1.3.3 Research Objectives

Objective 1: Analyse and describe the operations of minibus taxis in an organically-evolved, quasi-demand-responsive paratransit system, and assess the system's efficiency.

- **1.1** Describe the operations of minibus taxis in Kampala's organically-evolved, quasi-demand-responsive paratransit system.
- **1.2** Estimate the system efficiency from the passengers' and drivers' perspectives.
- **1.3** Characterise the movement patterns of minibus taxis in Kampala's paratransit system.

Objective 2: Develop and validate an agent-based model and simulator that describes minibus taxis and passengers interaction dynamics in a quasi-demand-responsive paratransit system and establish user-centric efficiency metrics.

- **2.1** Design and describe an agent-based model of minibus taxis and passengers in Kampala's organically-evolved, quasi-demand-responsive paratransit system.
- **2.2** Implement and validate the designed agent-based model in a simulator. This includes the study of the micro-level semi-autonomous interactions between Kampala's minibus taxis and passengers and analysis of the emergent behaviour at the macro level of the system.
- **2.3** Establish user-centric metrics for evaluating the efficiency of minibus taxi transportation in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

Objective 3: Use the agent-based model simulation to assess potential improvements to the efficiency of Kampala's minibus taxi transportation systems and make appropriate system improvement recommendations.

- **3.1** Optimise selected efficiency metrics of Kampala's simulated minibus taxi transport system and evaluate the associated gain in system efficiency at a macro level.
- **3.2** Recommend a plan for improving the efficiency of minibus taxis operations in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

1.3.4 Original contributions

This dissertation has introduced a new method of studying organically-evolved paratransit systems as complex adaptive systems and provides insight into the mobility dynamics of agents (minibus taxis and passengers) in the paratransit system and how they autonomously interact at micro-level, synchronising their behaviour in a *self-organising* process, leading to *emergence* of pseudo-order at the macro level of the system. The specific contributions are summarised below.

1. It scientifically describes the characteristics associated with minibus taxi operations in an organically evolved paratransit system. They include: *quasi-demand-responsiveness*; gradual *route evolution*; and they exhibit a *Lévy walk process* movement pattern when searching for passengers.

2. Developed an agent-based model for minibus taxi transport and simulated the transportation dynamics in a quasi-demand-responsive setting.

3. It contributes to public transport policy paratransit improvement and streamlining by proposing a radical shift from the current capital intensive and unsustainable approach that emphasises transport system overhaul devoid of paratransit to a more economically, culturally feasible and sustainable approach.

1.4 Dissertation synopsis

This research pursues a three-stage process to achieve the main objectives and address the research questions. Each Stage is presented substantially in two dissertation chapters and is associated directly with one or two research objectives and research questions (refer to Figure 1.1). This dissertation reports on the research process completed. The results from Stage one were reported in two journal articles ((Ndibatya and Booysen, 2020a) and the other under review in the Journal of Transport Geography), and one conference paper (Ndibatya et al., 2016). Part of Stage three results were published in one conference paper (Ndibatya and Booysen, 2020b). The order of the research stages (I-III) corresponds to the logical sequence of the research process as summarised in Figure 1.1. Below is the stage-by-stage summary of the research as presented in this dissertation.

1.4.1 Stage I (Ch3 and Ch4): Analysis and description of minibus taxi operations in Kampala's paratransit system

The first stage aims at analysing and describing the operations of minibus taxis in Kampala's organically-evolved paratransit system and assessing the system efficiency. At this stage of the research, we carried out two distinct but related empirical studies that we described substantially in Chapters 3 and 4. The first study (refer to Chapter 3) uses a three-pronged approach. (i) We studied the operations of minibus taxis in Kampala's paratransit system, from basics like regulation, management, and routes, to the industry's unique and peculiar practices. (ii) We studied and evaluated the economics of running a minibus taxi business, and we estimated the drivers' profitability index (PI). (iii) We assessed the efficiency of the minibus taxi transportation system from the passengers' and drivers' perspectives. Subsequently, in Chapter 4, we studied the movement characteristics of minibus taxis in Kampala's paratransit system using floating car data. We were interested in discovering whether minibus taxi movement patterns were consistent with Lévy walk behaviour; whether the routes the taxis used changed topology or shape over time, in other words, evolved; and whether their movements could suggest anything about their level of determination when searching for passengers.

Results from Stage one show that minibus taxis in Kampala are semi-organised and 'quasi-demand-responsive'. Their stops and routes are not published for the general public, and the routes gradually evolve in response to demand. When searching for passengers, minibus taxis adopt (either subconsciously, or through experience) a scale invariant supper-diffusive structure synonymous with *Lévy walk* pattern, consisting of many short 'steps' interspersed with long and rare steps. The long steps fit into a power law probability distribution defined by $f(x) \sim l^{-\alpha}$ where l is the step length, and α (referred to as the Lévy exponent) is in the range $1 < \alpha < 3$.

We made two main contributions during this stage. First, we described in geospatial terms the network of routes and stops used by minibus taxis in Kampala, and quantitatively estimated the waiting time, *hold back* time, and the average cost of travel per kilometre when using minibus taxis in Kampala. Second, we have discovered and proved (with empirical evidence) a scientific method (the Lévy walk) that explains the behaviour of minibus taxis when searching for passengers in an organically-evolved paratransit system. The details about methods and results are presented in detail in Chapters 3 (p. 27) and 4 (p. 27).

Correspondingly, Stage I covers research Objectives 1 and 2, and it addresses the associated research questions RQ1 and RQ2, respectively.

1.4.2 Stage II (Ch5 and Ch6): Designing, simulating, and validating a minibus taxi agent-based model

This stage aims at developing, simulating, and validating an agent-based model of minibus taxi transportation dynamics in an organically-evolved, quasi-demand-responsive paratransit system setting. Chapter 5 (p. 61) presents the design and description of the agent-based model (ABM). The ABM describes travel by minibus taxi in Kampala as a collection of autonomous decision-making entities called agents. Chapter 6 (p. 79) presents a controlled agent-based simulation experiment which is the process of model execution that takes the ABM (designed earlier in Chapter 5) through discrete state changes over time. During the simulation, each agent (e.g., passenger or driver in control of a minibus taxi) individually assesses its situation and makes travel decisions based on a set of rules. The agents repetitively interact (with self, with other agents and with their surroundings) in a common environment, executing various actions (such as boarding a taxi and searching for passengers). Agents in the ABM are capable of learning and adapting their behaviour to achieve desired goals, thereby allowing new and sometimes unanticipated behaviour to emerge.

Specifically, the agents in the controlled simulation experiment were designed to represent (as closely as possible) the behaviour dynamics in Kampala's minibus taxi system. Thus, the minibus taxi agents were designed to have: limited situational awareness, quasi-demand-responsiveness, random passenger search, Lévy walk behaviour and the occasional abandonment of routes that were considered 'unprofitable' (with persistently low passenger occupancy). The passenger agents were designed to have: limited situational awareness, considerable persistence (able to move from stop to stop waiting for a taxi), and little memory – they depended on episodic memory to make some decisions such as where to wait for a taxi.

Analysis and validation of the simulation results indicated distributions statistically close to the distributions obtained from the field study in Kampala. We therefore concluded that the agent-based simulation closely represented the minibus taxi transport dynamics in Kampala's organically-evolved, quasi-demand-responsive paratransit system. We further identified four primary metrics for evaluating the efficiency of a paratransit system. These included: the passenger waiting time t_w; minibus taxi hold-back time t_h; passengers' first leg distance d_{l1}; and minibus taxi occupancy $\mathcal{O}$.

Correspondingly, Stage II addresses objectives 2.1, 2.2 and 2.3, and answers the associated research questions RQ3 and RQ4.

1.4.3 Stage III (Ch7 and Ch9): Optimising minibus taxi operations' metrics to improve efficiency

This stage aims at testing hypothesis two by optimising selected metrics and improving the intelligence and situational awareness of the simulated minibus taxi and passenger agents. In Chapter 7 (p. 97), we set up two test agent-based simulation experiments of minibus taxi transportation dynamics. In the first test experiment (TOR1), we improved the agents' decision making based on supervised learning to improve the agents' situation awareness. The passengers and minibus taxi agents in TOR1 observe the status of the world; they form the current and future beliefs based on a supervised learning algorithm (Random forest). The supervised learning algorithm generates situational awareness scores that are used to evaluate the alternative with high utility. The second test experiment (TOR2) further improves the agents' decision making and situation awareness based on a deep learning method, namely, a convolutional neural network (CNN). The CNN was trained to optimally rank (or order) a set of choices an agent has to choose from, such that the option with the highest utility is on top. Thus, the agent chooses the one with higher utility.

Results from optimising and analysing selected metrics from the two test experiments (TOR1 and TOR2) indicate substantial improvement in minibus taxi transport system efficiency at macro-level. Thus, hypothesis two is correct.

Correspondingly, stage III addresses Objectives 3.1, and Objective 3.2, and answers the associated research question RQ5.

1.5 Structure of the dissertation

The contents of this dissertation are presented in nine chapters, broadly grouped into two parts. The first part consists of Chapters 1 and 2. Chapter 1 presents the motivation behind this research, the justification for categorising organically-evolved paratransit systems and complex adaptive systems (CAS), and the relationship between CAS and agent-based modelling are presented in Section 1.2. Section 1.3 focuses the research on the research questions, hypotheses, objectives, and it further expounds on the original contributions made to the wider research community and the public transport industry in developing cities of Sub-Saharan Africa. Chapter 2 introduces the theoretical background to the research and presents a review of the literature related to this research.

The second part consists of Chapters 3 to 9. In these chapters, three distinct but chronologically related research processes that were undertaken are presented (grouped into Stages I, II, and III), and their scientific context to the broader objectives explained. Summaries of key findings and contributions at each stage of the research process are presented at the end of each Chapter. Subsequently, in Chapter 8, the most relevant results and individual contributions from the previous chapters are put into a broader context, and the substance of the findings are discussed in relation to the research questions and objectives in Section 1.3. Finally, the dissertation ends with conclusions and an outlook to future research in Chapter 9. Figure 1.1 summarises the chronological relationship between the research stages, dissertation chapters, research questions and objectives.

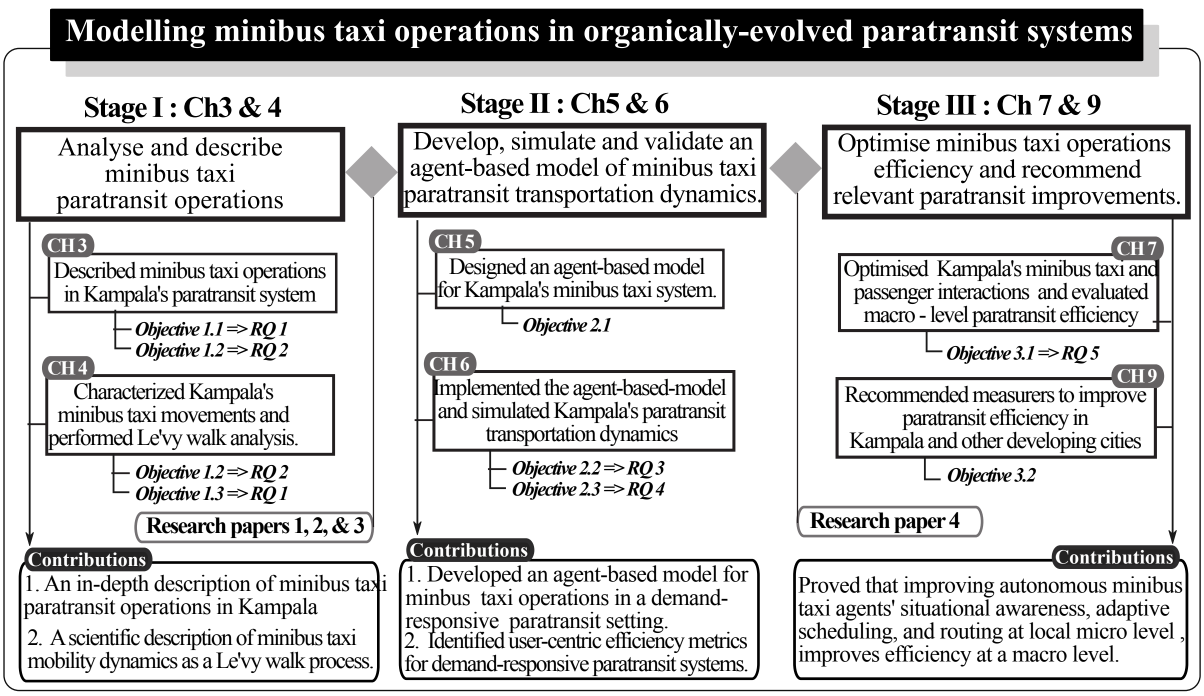


Figure 1.1: Research process of the dissertation.

1.6 Research Papers

⇒ **Research Paper 1: (Conference)**

Ndibatya, I., Coetzee, J., and Booysen, M. J. (2016). *Mapping the informal public transport network in Kampala with smartphones: Making sense of an organically evolved chaotic system in an emerging city in sub-Saharan Africa.* In Proc. 35th Southern African Transport Conference, Pretoria, pages 4–7.

⇒ **Research Paper 2: (Journal)**

Ndibatya, I. and Booysen, M. J. (2020a). *Minibus taxis in Kampala's paratransit system: Operations, economics and efficiency.* Journal of Transport Geography, 88.

⇒ **Research Paper 3: (Journal)**

Ndibatya, I. and Booysen, M. J. (Submitted, Oct, 2020). *Characterizing the movement patterns of minibus taxis in Kampala's paratransit system*, Journal of Transport Geography.

⇒ **Research Paper 4: (Conference)**

Ndibatya, I. and Booysen, M. J. (2020b). *Transforming Paratransit in Africa's congested Cities: An ICT- enabled Integrated Demand Responsive Transport (iDRT) approach.* In Miriam, C. and Paul, C., editors, IST-Africa 2020 Conference Proceedings, pages 1–10, Kampala, Uganda. IST-Africa Institute and IIMC.

Chapter 2

Literature review

This chapter presents the literature related to the scientific methods used in this dissertation. It starts with the background to complex systems research. It then discusses how organically-evolved paratransit systems in developing cities of Africa in particular, and the Global South, in general, fit the description of a complex adaptive system (CAS). It further explores the approaches to complex adaptive systems modelling, where the concept of agent-based modelling is discussed in detail. It concludes with a discussion of the different approaches to designing and training autonomous agents to improve their intelligence and situational awareness.

2.1 Background to complex systems

The foundations for complex systems research dates to the works of several mathematicians such as Poincare (1881, 1908), Prigogine (1961), and Lorenz (1963) in understanding the unpredictable behaviour of non-linear dynamical systems. Prior to this, between 1500–1700, there had been a fundamental shift in perception of the world, from a worldview governed by Christian theology and ethics, to that of a machine-like world governed by natural forces and mathematical equations (Capra and Luisi, 2014). This period is also known as the *Newtonian era*. In his book "*On the Shoulders of Giants: The Great Works of Physics and Astronomy*", Hawking (2002) compiled the major scientific discoveries that epitomised the Newtonian era. Most notable of all include Galileo Galilei's "Two New Sciences"; Johannes Kepler's "Mystery of the Cosmos"; Albert Einstein's "Principle of Relativity"; and Sir Isaac Newton's *Principia* (Newton, 1687, 1713, 1726). Post the Newtonian era, complexity science has since developed two schools of thought: the *reductionist* and the *emergentists* schools of thought.

Reductionists believe that systems can be entirely explained in terms of their components, and that the overall functioning of the system is a sum of the individual components. Therefore, to explain new emergent properties, reductionists reduce system elements into lower-level components, and they develop interaction rules between lower-level components and higher-level components (Raisio and Lundström, 2014). The *reductionism philosophy* was held by the scientific community since the publication of Newton's *Principia*, and various theories were developed around it, e.g., the chaos theory. As demonstrated by Lorenz (1963), chaos is some kind of order without periodicity. Lorenz's talk entitled "*Predictability: - Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?*" which he delivered to the 139^{th} meeting of the "*American Association for* *the Advancement of Science*" fundamentally paved the way for the application of chaos theory in other fields such as natural, social, and engineering sciences using the principle known as the *butterfly effect*. The *butterfly effect* principle simply indicates that a small change in the initial conditions of a modelled system can produce a significant change in the final state of the system (Riley, 2012). Chaos is deterministic and linear, with mathematical meaning (Hamstra, 2017), and is sensitive to its initial conditions (Holte, 1993). However, existing models of chaos describe dynamics of one (or a few) variables (also known as low-dimensional chaos). Thus, it is not suitable for systems with many variables.

The emergentists, on the other hand, view systems as being non-linear, thus, future states are unpredictable (Raisio and Lundström, 2014). As systems transition from simple to complex, the predictive mechanisms become less reliable and cannot be completely explained in terms of their individual components. In the emergentists school of thought, the world is viewed as an *organic* entity, composed of interacting components (Schneider and Somers, 2006). The interactions tend to lead to new system states, contributing to the system's unpredictability. Often, what *emerges* is more than the sum of the processes from which it emerges. For example, ants are capable of building very complex colony structures using simple local interaction rules (Odell, 2002). Based on the *emergence* principle, scientists have been able to study a whole spectrum of complex phenomena in different fields ranging from biological sciences, physics, chemistry, and engineering; to social, economic, and cognitive sciences (Bennet and Bennet, 2004; Shiell et al., 2008).

2.1.1 Complexity theory and complex systems

Complexity theory concepts emerged in the late 19^{th} century. They were applied in many multidisciplinary works. These works include Prigogine's (Prigogine, 1961) work on dissipative structures in non-equilibrium thermodynamics; Lorenz's (Lorenz, 1963) work on weather systems, non-linear pathways (i.e., the butterfly effect) and chaos theory; as well as evolutionary thinking informed by Lamarck's views on learning and adaptation. Larmark argued that complex dynamic systems often acquire characteristics that sometimes become inherited traits (Marion, 1999). Complexity theory allows us to understand better how order emerges in complex, non-linear systems as diverse as cells, human beings, galaxies, ecologies, forest ecosystems, markets, and social systems that are only partially understood by traditional scientific methods (Schneider and Somers, 2006). Complexity theory spans across vast disciplines in the physical, biological, and social sciences, and greatly influences how we think about and act within the world (Schneider and Somers, 2006).

The complexity theory body of knowledge is composed of four different theories that are often combined to model and analyse complex systems (Hooker, 2011). The theories are as follows:

1. Systems theory: This deals with the concepts of self-organisation and adaptability (Luhmann and Gilgen, 2012). In systems theory, a system is considered a cohesive collection of interrelated and interdependent parts which can be natural or human-made. Every system is bounded by space and time, influenced by its environment, defined by its structure and purpose, and expressed through its functioning. A system may be more than the sum of its parts if it expresses synergy or emergent behaviour.

2. Chaos theory: This is the study of non-linear systems, or things that may look completely random but still have an underlying cause that may not seem obvious on the surface (Hayles, 1991). From chaos theory, we gain an understanding of feedback loops and non-linear systems.

3. Network theory: This uses network graphs as a representation of either symmetric or asymmetric relations between discrete objects (Cohen and Havlin, 2010). Network theory relies less on models and more on real-world data (Saleh et al., 2017).

4. Adaptive systems theory: This deals with the study of interacting or interdependent entities, real or abstract, forming an integrated whole that together can respond to environmental changes or changes in the interacting parts, in a way that either maintains stable internal state or evolves and adapts to new states (Martín H. et al., 2008). Adaptive systems theory studies the organisation of things that do not have centralised control and are governed by simple rules that emerge through interaction.

In general, complexity theory views systems as a process that is self-organising. It applies mathematical modelling of linear and predictable states when dealing with chaos or chaotic systems, whereas it employs complex adaptive systems (CAS) to deal with unpredictable non-linear systems.

2.2 Complex adaptive systems

Despite the ubiquity of complex adaptive systems (CAS), there is no standard definition. For purposes of this dissertation, we shall adopt a definition suggested by Abbott and Hadžikadić (2017). A CAS is "*a system composed of a large number of independent simple components that locally interact in an independent and non-linear fashion, exhibit self-organisation through interactions that are neither completely random nor completely regular and are not influenced by some central or global mechanism and yield emergent behaviour at large scales that are not predictable from observation of the behaviour of the components*". CAS can further be described from the atomic definitions of its constituent words, i.e., complex, adaptive and system. The word "complex" implies diversity, through a great number, and wide variety of interdependent, yet autonomous parts. "Adaptive" refers to the system's ability to alter, change, and learn from past experiences. The "system" portion refers to a set of connected, interdependent parts – a network (Zimmerman et al., 2001).

While there are a significant number of CAS existing at different scales, complexity theory reveals that there are common, interrelated principles which can be observed across all CAS (Zimmerman et al., 2001). These principles include: Path dependence, non-linearity, emergence, and adaptiveness (Lindberg and Schneider, 2013; Ladyman et al., 2013). *Path-dependent* systems are sensitive to their initial conditions. Thus, the same condition might affect seemingly similar systems differently, based on their histories. If small changes in the system can lead to big effects and, within the same system, big changes can have minimal effects, then the effects are difficult to predict; thus, the system is *non-linear*. According to Lindberg and Schneider (2013), *emergence* refers to the system's interactions that could result in new system states that are different from the initial states. The emergence also makes CAS irreducible due to its emergent properties. A phase transition thus occurs, changing the initial lower-level states. One of the unique characteristics of CAS is the ability to *adapt* and operate between chaos and order (Turner and Baker, 2019). By operating between chaos and order, CAS avoids the status quo while at the same time avoiding complete chaos. This balance is self-organising and allows CAS to learn and evolve into new emergent states (Turner and Baker, 2019).

The smallest components of CAS are often referred to as agents (Abbott and Hadžikadić, 2017). Agents can respond to stimuli. A few simple rules govern this stimulus/response behaviour of an agent. In CAS, there are homogeneous and heterogeneous local interactions of groups of agents in a variety of configurations. In systems with few agents, these interactions can be predicted, as there are usually a limited set of interactions that each agent can perform. These random local interactions generally yield outcomes that approximate to the sum of each individual interaction; in some cases, however, as the combinations of agents increase in varying proportions, acting in different ways, complex and potentially novel behaviours emerge from these combinations of agents that often yield significantly greater outcomes than expected (Ladyman et al., 2013). When a specific collection of agents combines to produce these emergent behaviours through the process of aggregation, they are referred to as aggregate agents (Holland, 1995). These aggregate agents group together with other aggregate agents to form larger complex adaptive systems with richer sets of emergent behaviours and interactions.

Two approaches are normally used to study complex adaptive systems. The first is the creation of simplified *mathematical models* that try to abstract the most important qualitative elements into a solvable framework from which we can gain scientific insight. In this approach, numerical methods are used. The second is *Agent-based modelling and simulation*, where more comprehensive and realistic computer models representing the interacting parts of the complex system are developed down to the lowest level of detail, and the interactions simulated to measure the emergent behaviour (Newman, 2011). The flexibility, ability to capture emergent phenomena, and provision for a natural description of the system, have made agent-based modelling popular among researchers seeking to solve problems involving real-world complex systems (Bonabeau, 2002).

2.3 Paratransit as complex adaptive systems

Available literature indicates that paratransit systems in the developing cities of the Global South *organically emerged* in the late 1990s to fill the void left by the gradual collapse of the state-run transport enterprises (Kumar, 2011). Since the emergence of paratransit, it has gradually *evolved* into a fully-fledged quasi-demand-responsive transport system with little or *no centralised control* (Ndibatya and Booysen, 2020a). The paratransit system is composed of many interdependent components that interact non-linearly, often operating between "chaotic" and semi-orderly states (Godard, 2013; Behrens et al., 2015b; Diaz Olvera et al., 2019). Goodfellow (2012) documented some of the *many components* that play vital roles in the paratransit system of Kampala city in Uganda, whereas Booysen et al. (2013) identified the distributed and atomic ownership of the vehicle as one of the factors that create many competing centres of decentralised control. However, despite the chaotic and disorganised nature of paratransit, it *adapts* and serves (though inefficiently) the mobility needs of the urban population in the developing cities of the Global South (Pablo, 2015). This is evident in the way it has successfully responded to the *urban sprawl* problem that has characterised most poorly planned developing cities in the last decade (Woolf and Joubert, 2013; Veng and Tetsuo, 2016; du Preez et al., 2019). In response to the urban sprawl, existing and new paratransit routes organically evolve to serve the transport needs of the emerging dispersed settlements (Ndibatya and Booysen, 2020a,b). We can therefore state that paratransit exhibits *self-organising* properties.

Based on Abbott and Hadžikadić (2017)'s definition of a complex adaptive systems, we can classify paratransit systems in developing cities of the global South as CAS. This is because they exhibit properties such as: Having a large number of interdependent components, non-linear interaction of components, absence of centralised control, adaptation and emergence of new states, and self-organisation. Thus, when modelling paratransit, researchers ought to explore the emergentists' school of thought.

2.4 Agent-based modelling

Agent-based modelling is the computational modelling of systems as collections of autonomous interacting entities. In other words, an agent-based model (ABM) is a computational instantiation of a complex adaptive system (CAS) (Blume, 2015). Initially, ABMs were applied in biological and social sciences and they focused on theory and hypothesis development (Hammod, 2015). In biological sciences, ABM applications included: Holland (1992) and Ohtsuki et al. (2006)'s work on evolutionary biology; and DeAngelis and Mooij (2005)'s work on ecology. ABM use in social sciences included: Epstein (2002)'s work on conflict; Schelling (1971)'s work on segregation; and Bendor et al. (2003)'s work on electoral dynamics.

With the exponential growth in memory, computing power, and datasets, ABM application have expanded to other fields such as: transportation (Ciari et al., 2014), land use (Berger et al., 2007), economics and finance (Dawid and Neugart, 2011), marketing (Rand and Rust, 2011), and education (Maroulis et al., 2014). Recently, ABMs are being used to inform policy or decisions in many fields. Examples of policy-related ABMs applications include: Agriculture and land-use policy (Brady et al., 2012; Berger and Troost, 2014); natural resource management (Heckbert et al., 2010); smart electricity grids (Ringler et al., 2016); health service supply chain design (Rouzafzoon and Helo, 2016); and control of communicable diseases outbreak (Lee et al., 2010; Eubank et al., 2004; Toroczkai and Eubank, 2006).

In their work entitled "*How to halt a smallpox epidemic*", Toroczkai and Eubank (2006) developed an ABM of the spread of smallpox through a city. The model was calibrated on pre-existing data, and then the effects of several different vaccination regimens were simulated. They concluded that if detection is sufficiently fast and targeted response is effective, then mass vaccination would not be necessary. Likewise, Weidlich and Veit (2008) made a critical review of ABM in electric power systems, specifically the wholesale electricity markets. They concluded that the characteristics of these markets are challenging to analyse with conventional economic optimisation and equilibrium modelling. The market characteristics deviate from the known conditions of the "perfect competition" assumption, and the participants in the markets engage in "strategic behaviour"; i.e., in making decisions they also consider the potential decisions, and reactions, of other participants (Sanstad, 2015).

2.4.1 Transportation modelling

Transportation engineers and planners have relied on transportation forecasting models for decades to study vast complex transport-related dynamics that range from congestion and air quality to social equity. In the past, two major approaches to travel demand modelling were developed and used (Ortúzar and Willumsen, 2011). These were the *trip-based* and *activity-based* approaches. The types of models developed using these approaches are also referred to as *equilibrium assignment models* (Florian and Hearn, 2008). The trip-based approach takes individual trips as the elementary subjects and considers aggregate travel choices in four steps: trip generation, trip distribution, modal split, and route assignment.

During *trip generation* (trip production & attraction), the study area is divided into zones that are similar in their traffic characteristics. For each zone, the number of trips produced and attracted are calculated. For instance, in the morning, residential areas are trip producers and workplaces are trip attractors Brotcorne et al. (2002). *Trip distribution* involves matching the trips produced in one zone to the trips attracted in another zone. The matching follows a gravity model – trips produced at an origin and attracted to a destination are directly proportional to the total trip productions at the origin and the total attractions at the destination (see Equation 2.1) Brotcorne et al. (2002). For instance, a traveller is more likely to go to work at a place closer to them than a place further away. The result of this step is an origin-destination (OD) matrix. *Modal Split* involves assigning each OD pair an estimated mode split of the available transport services (Ortúzar and Willumsen, 2011). For example, in the high-income areas, the majority are assumed to travel by car. In *route/network assignment*, the route to each assignment is assigned to the underlying road network. This uses volume-delay functions to do the assignment. The network loading function finds an equilibrium assignment so that all the links are evenly loaded.

$$T_{ij} = \frac{A_j F_{ij} K_{ij}}{\sum_{z=1}^{n} A_z F_{ij} K_{iz}} x P_i \tag{2.1}$$

where T_{ij} = trips produced at i and attracted at j; P_i = total trip production at i; A_j = total trip attraction at j; F_{ij} = a calibration term for interchange ij, (friction factor) or travel time factor ($F_{ij} = \frac{C}{t_{ij}^n}$); C= calibration factor for the friction factor; K_{ij} = a socioeconomic adjustment factor for interchange ij; i = origin zone; n = number of zones.

Though the equilibrium assignment models have been used for decades by transport planners to support transport decisions, they are not a good representative of the reality in developing cities, and they are not very intuitive to understand and validate. Therefore, there is a growing interest in using agent-based modelling to support decision making. While agent-based models are not commonly used in travel demand forecasting as such, many activity-based models are agent-based models of a sort – at least in part – though the behaviours of the agents are typically very complex. In agent-based modelling, we model agents. Each agent makes autonomous decisions, interacts with one another, and interacts with their environment. ABMs are much more intuitive to understand and validate. They also produce more detailed results that can be used in many more different comprehensive analyses.

2.4.2 Agent-based models of transportation

In transport applications, the important advantage of ABM is the ability to include structurally rich, dynamic, and heterogeneous representations of social or environmental exposures and influences (Hammod, 2015). For example, ABM can incorporate a representation of geographic information system (GIS) data, or detailed network structures (Zhang and Levinson, 2004).

Several agent-based modelling frameworks have been developed. They include general frameworks such as Repast (North et al., 2013), Mason, Netlogo (Robertson, 2005), MESA (Masad and Kazil, 2015) and other frameworks specific to modelling transportation problems. Table 2.1 shows a sample of ABM frameworks developed for transport. Figure 2.1 shows the MESA ABM framework architecture.

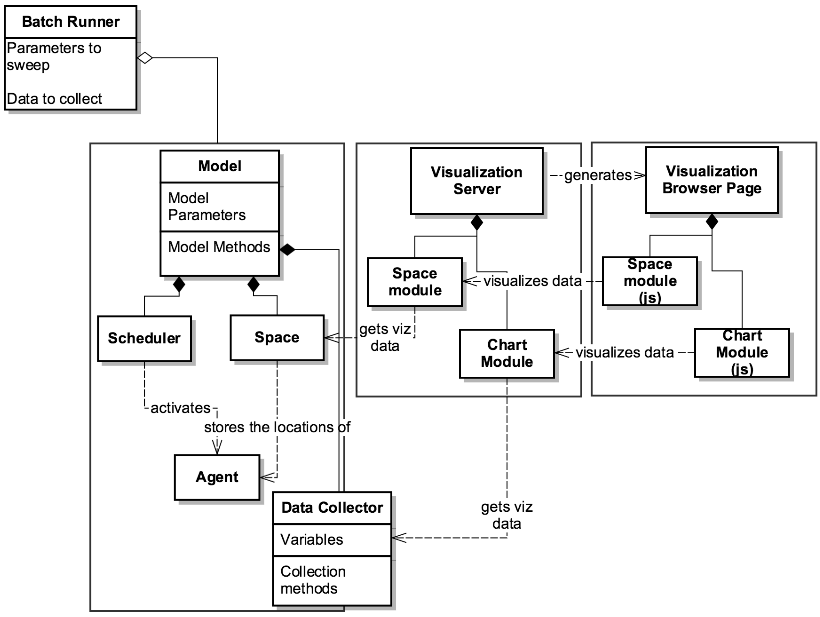


Figure 2.1: MESA architecture for agent-based modelling.

Table 2.1: Summary of ABM frameworks for modelling transportation.

Source	Model/Tool	Applications	Characteristics	Benefits	Limitation
Jakob et al. (2012)	AgentPolis	Strategies to prevent fare evasion	On-demand vehicle allocation and routing mechanism	Wide range of metrics	Trip-oriented matrices
Bell et al. (2012)	UrbanSIM	Zone-level mobility forecast	Integrates the 4-step model and trip-based OD matrix	Rich public transport ecosystem	Derived from aggregate demand
Ciari et al. (2014)	MATSim	Car sharing and autonomous on-demand taxi service	Mode choice based on utility maximisation	Incorporates land-use and agents preference	Difficulties in calibration
Azevêdo et al. (2016); Oh et al. (2018)	SimMobility	Connected and Autonomous Vehicles to replace public transport	Three components to simulate short, medium, and long-term time horizon	Seamless integration	No interaction with public transport services
Laine and Bruun (2017)	Brutus	Planning for MaaS	Multi-modal transport demand forecasting	Trip-chains	No information on deployment

2.5 Intelligent agents

An intelligent agent (IA) refers to an autonomous entity which acts rationally towards achieving goals in an environment perceived using sensors (Franklin and Graesser, 1997; Jokinen and Wilcock, 2019; Mills and Stufflebeam, 2017). Burgin and Dodig-Crnkovic (2009) identified four main characteristics an intelligent agent should possess. These characteristics are: (1) Usefulness: i.e., it should be of service to others and may assist others or be part of a larger process. (2) Context-awareness (or situational awareness (SA)). Naderpour et al. (2012) defines situational awareness as "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future*". (3) Ability to learn: an IA should be capable of learning from previous experiences. (4) Adaptability: an IA should be able to adapt to apply what it has learned.

Russell and Norvig (2003) classified intelligent agents into five types based on the degree of perceived intelligence and capability. The intelligent agents were classified into simple reflex, model-based, goal-based, utility-based and learning-based agents as explained below:

i) Simple reflex agents: These are the basic form of agents that function only in the current state. Their intelligence capability is very low, and they often ignore history. Their responses are based on pre-defined rules. They perform well only when the environment is fully observable.

ii) Model-based agents: These agents choose actions in the same way as reflex agents, but they have a more comprehensive view of the environment in addition to the capacity to store past internal states. Model-based agents update the internal state at each step. To perform any action, it relies on both internal state and current perception. However, it is almost impossible to find the exact state when dealing with a partially observable environment.

iii) Goal-based agents: These agents expand upon the information model-based agents store by including goal information, or information about desirable situations. To attain its goal, it uses the search and planning algorithm. One drawback of goal-based agents is that they do not always select the most optimal path to reach the final goal. This shortfall can be overcome by using the utility agent described below.

iv) Utility agents: These agents are similar to goal-based agents but provide an extra utility measurement which rates each possible scenario on its desired result and chooses the action that maximises the outcome. Rating criteria examples could be the probability of success or the resources required.

v) Learning agents: These agents can gradually improve and become more knowledgeable about an environment through an additional learning element. The learning element uses feedback to determine how performance elements should be changed to improve gradually.

2.5.1 Rationality among intelligent agents

In game theory, decision theory, artificial intelligence, and economics, a rational agent is an agent that: has clear preferences; models uncertainty via expected values of variables or functions of variables; and always chooses to act with the optimal expected outcome for itself from among all possible actions (Martín H. et al., 2008). Russell and Norvig (2003) view the agent's goal-directed behaviour to be the essence of intelligence. In this rational-action paradigm, an IA possesses an internal "model" of its environment. This model encapsulates all the agent's beliefs about the world. The agent also has an "objective function" that encapsulates all the IA's goals. Such an agent is designed to create and execute whatever plan will, upon completion, maximise the expected value of the objective function (Russell and Norvig, 2003). Simon (1957) introduced the term "*bounded rationality*" to tailor the concept of rationality with cognitively limited agents. Bounded rationality has since come to refer to a wide range of behaviour which departs from the assumptions of "*perfect rationality*".

Perfect rationality assumes a hypothetical agent having complete information about the options available, perfect foresight of the consequences from selecting any of those options, and the computational capability to solve an optimisation problem that maximises the agent's utility, which is often complex (Wheeler, 2020). However, "bounded rationality", introduces a risk factor associated with decision making such that the agents' decision model accounts for the individual preferences, the existence of partial information and the weaknesses in the individual computational capability of agents (Wheeler, 2020).

2.5.2 Situational awareness among intelligent agents

Situational awareness (SA) was defined by Jiang (2020) as the progression by an intelligent agent through three levels: (1) the *perception* of environments with respect to time or space within the situation, (2) the *comprehension* of their meaning, and (3) the *projection* of their future status. Within a multi-agent system or team perspective, SA can be defined as "*the degree to which every team member possesses the situation awareness required for his or her responsibilities*" (Endsley, 1995). SA has been recognised as a critical foundation for successful decision-making in agent-based systems. As observed by Burgin and Dodig-Crnkovic (2009), SA is one of the major characteristics that makes the agents intelligent. Thus, agents in ABMs need to adopt methods that enhance their levels of situational awareness.

2.5.3 Learning among intelligent agents

Intelligent agents (IA) that have the ability to learn often use machine learning algorithms. Machine learning (ML) investigates how computers can learn to recognise complex patterns and make intelligent decisions based on data (Pedregosa et al., 2011). There are three basic machine learning paradigms: supervised, unsupervised and reinforcement learning (Talabis et al., 2015). *Supervised learning* provides powerful tools to classify and process labelled data. *Unsupervised learning* is synonymous with clustering. The learning process is considered unsupervised since the input examples are not class labelled. Clustering can also be used to discover classes within the data. *Reinforcement learning* is a machine learning approach that lets users/agents play an active role in the learning process (Schmidhuber, 2015). A reinforcement learning agent can have a "reward function" that allows the programmers to shape the IA's desired behaviour as illustrated in Figure 2.2a. The reinforcement learning scenario in Figure 2.2a shows that an agent takes actions in an environment, which is interpreted into a reward and a representation of the state, which are fed back into the agent.

In supervised learning, there are four main types of classification tasks: binary, multi-class, multi-label, and imbalanced classification (Schmidhuber, 2015). Widely used classification tools during supervised learning include, logistic regression, linear discriminant analysis, K-nearest neighbours, trees, random forests, support vector machines, and neural networks (Talabis et al., 2015; Schmidhuber, 2015).

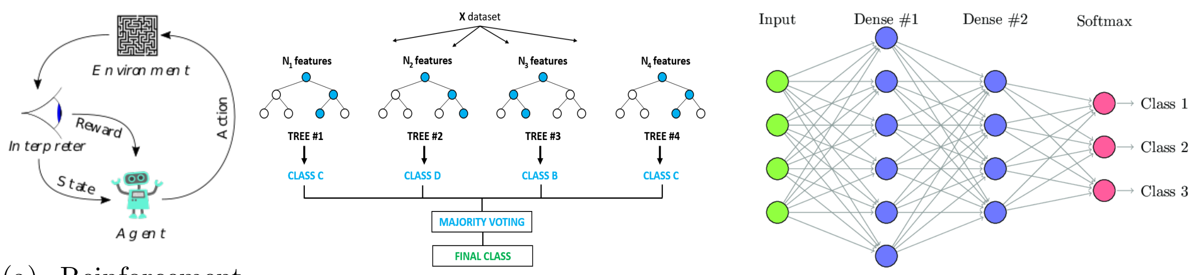
Learning among intelligent agents can take on many facets depending on the agents objective. For the purposes of this dissertation, we shall restrict ourselves to two learning objectives: learning how to classify, and learning how to rank, using random forests, and convolutional neural network (CNN), respectively. Classification problems usually involve predicting classes of items based on predefined labels, which are often discrete values. For example, classifying photos (Kussul et al., 2017), emails, and customer risks in banking applications. Tools such as random forests (Breiman, 2001) and neural networks (Pelletier et al., 2019) are used for classification problems. Recently, classification problems have evolved to ranking problems which usually involves both classifying and ordering (Dong et al., 2009).

2.5.3.1 Random forests

Random forests are tree-based machine learning algorithms popularly used to solve supervised learning problems flexibly. A random forest consists of a collection of tree structured classifiers $\{h(\mathbf{x}, \Theta_k), k = 1, ...\}$ where the $\{\Theta_k\}$ are independent, identically distributed random vectors and each tree casts a unit vote for the most popular class at input $\mathbf{x}$ (Breiman, 2001). A random forest algorithm goes through four steps as illustrated in Figure 2.2b. First, the algorithm selects random samples from the dataset provided. Second, it creates a decision tree for each sample selected; then it gets a prediction result from each decision tree created. Third, voting is performed for every predicted result. Finally, the algorithm selects the most voted prediction result as the final prediction.

2.5.3.2 Convolutional neural network

Deep Convolution Neural Networks (CNN) have been used successfully for many machine learning applications such as object recognition, machine translation, remote sensing data classification, as well as missing data reconstruction (LeCun et al., 2015; Schmidhuber,



(a) Reinforcement learning

(b) Random forest

(c) Convolutionary neural network

Figure 2.2: (a) The Reinforcement Learning (RL) scenario, (b) the process of classification using Random Forests; (c) Example of a dense (fully connected) neural network (Pelletier et al., 2019)

2015). Deep learning networks are based on a concatenation of different layers of neurons – simple connected processors each producing a sequence of real-valued activations– where each layer takes the outputs of the previous layer as inputs (Pelletier et al., 2019). Figure 2.2c shows a neural network composed of two *dense* (fully connected) layers, and a "SoftMax" activation layer. The green, blue and red neurons represent the inputs, hidden layers, and outputs, respectively. The SoftMax layer is a special case of a dense layer that maps the output of the previous layer to a vector of class probabilities. The activation for neuron i (class i of the total number of classes C) is an extension of multiclass sigmoid function that can be written as:

$$A_i^{[L]} = \frac{e^{Z_i^{[L]}}}{\sum_{j=1}^{C} e^{Z_j^{[L]}}} \tag{2.2}$$

Where $Z_i^{[L]}$ is the result of the linear combination of neuron i of the SoftMax layer, i.e., $Z_i^{[L]} = W_i^{[L]} A^{L-1} + b_i^{[L]}$, C. For a given training instance, the C activations sum to one and can be interpreted as a probability distribution over the class (Pelletier et al., 2019).

2.5.3.3 Learning to rank

Learning to rank or machine-learned ranking (MLR) is the application of supervised machine learning to construct ranking models. Ray and Triantaphyllou (1999) defined ranking as an ordering on a set of alternatives. Dong et al. (2009) classified learning to rank models into three classes. The first is *pointwise ranking*. Here the model inputs one item at a time and assigns it a probability of being relevant but ignores the relationship between the different items in the list. The second is *pairwise ranking*. Here models perform pairwise comparisons of each item in the list and learn the probability of one item being preferred to another item. This does not capture the entire list; it only compares pairs (Cao et al., 2007). The third is *listwise ranking*. Here the function inputs one item at a time and optimises the ordering of the list producing an optimal permutation on the items. In some high-level APIs – such as "TensorFlow-ranking" – multi-item scoring was introduced. Here the model inputs all the items at a time and produces the optimal ordering (Svore et al., 2011). This allows the scoring function to use the context of other items to make better ranking decisions (Pasumarthi et al., 2019). Ranking results are evaluated using a range of metrics for all the ranking methods. The metrics include: Normalised Discounted Gain (NDCG), Average Relevance Position (ARP), and Mean Average Precision (MAP) (Dong et al., 2009).

The popular NDCG method of evaluating ranking results evolved from the Discount Cumulative Gain (DCG) method developed by Freund and Schapire (1996); Freund et al. (2003) and is described below. First, a Cumulative Gain (CG) is computed (Equation 2.3), followed by a DCG (Equation 2.4), and finally a NDCG (Equation 2.5).
Each item with different relevance in the ranking list is assigned with a gain such as 0, 1, 2, or 3. And then the list built by different gains would be $\{3, 2, 3, 0, 0, 1, 2, 2, 3, 0, ...\}$. The item's ranking cannot be observed from the list, yet it can be solved by calculating the Cumulated Gain (CG).

$$CG[i] = \begin{cases} G[1] & \text{, if } i = 1 \\ CG[i-1] + G[i] & \text{, otherwise.} \end{cases} \tag{2.3}$$

The list built by CG is $\{3, 5, 8, 8, 8, 9, 11, 13, 16, 16, ...\}$. The larger the CG is, the greater the position of the item in the list. However, the gain would be too large to be calculated if there is a large number of items, which can be solved through computing DCG.

$$DCG[i] = \begin{cases} G[1] & \text{, if } i = 1 \\ DCG[i-1] + G[i]/logi & \text{, otherwise.} \end{cases} \tag{2.4}$$

And then the list built by DCG is $\{3, 5, 6.89, 6.89, 6.89, 7.28, 7.99, 8.66, 9.61, 9.61, ...\}$. However, people hope that there is a normalized DCG to compare the ranking results. Finally, they proposed NDCG to estimate ranking results. The definition of NDCG is as below (Dong et al., 2009):

$$\begin{aligned} NDCG = \{v_1/i_1, v_2/i_2, ..., v_n/i_n\} \\ v : non - ideal \quad DCG\ list \\ i : ideal \qquad DCG\ list \end{aligned} \tag{2.5}$$

Learning to rank has previously been applied in: search engines to rank documents in response to user queries (Joachims, 2002); recommendation systems to rank items for a given user (Elsas et al., 2008); dialogue systems to rank responses for user requests; and in question & answer systems to rank answers in response to user questions. (Dong et al., 2009; Liu, 2011; Valizadegan et al., 2009)

2.6 Conclusion

Despite the several studies performed on complex adaptive systems modelling and, in particular, agent-based modelling of transport systems, only a few have targeted public transport in the developing cities of the Global South. Currently, there is no agent-based model of paratransit systems in Sub-Saharan Africa and, more particularly, in Uganda. However, available literature indicates that transport planning projects undertaken in East Africa for the last decade were mainly based on the reductionist school of thought. They depended on the equilibrium assignment models for transport planning and transport-related decision making (JICA, 2014, 2008; ITP, 2010).

Notably, paratransit dominates the road-based public trips in several developing cities of the Global South. For instance, paratransit contributes 70%, 90%, 91%, and 98% of the road-based public trips in Johannesburg, Lagos, Kampala, and Dar es Salaam, respectively (Behrens et al., 2015a; Evans et al., 2018). Additionally, paratransit exhibits several properties of complex adaptive systems as discussed in Section 2.3. Therefore, the traditional equilibrium assignment models that are based on the reductionist school of thought may not provide the complete understanding of the complex dynamics associated with paratransit systems. Therefore, in this dissertation, we depart from the traditional practice and explore the emergentists school of thought to study the complex dynamics in paratransit systems using complex adaptive system approaches such as agent-based modelling.

Studies such as Hammod (2015), Azevêdo et al. (2016) and Ciari et al. (2014) demonstrate the feasibility for the development and deployment of agent-based models in transportation application. The studies illustrate the possibilities to model social interaction, environmental interaction, and seamless integration dynamics in addition to incorporating geographical and detailed network structures into the transportation agent-based models (Zhang and Levinson, 2004). Furthermore, flexible implementation architectures such as MESA (Masad and Kazil, 2015) present the opportunity to customise micro-level agents' structure and behaviour. Thus, overcoming the lack of flexibility encountered in earlier transport ABM frameworks such as MATSim (Ciari et al., 2014) and AgentPolis (Jakob et al., 2012).

The advances in machine learning and pattern recognition research have unlocked new methods of applying agent-based modelling to transport problems. Using evolutionary methods such as random forest (Breiman, 2001), convolution neural networks (Pelletier et al., 2019) and reinforcement learning, intelligent and situationally aware agents can be designed and trained (using data) to make more optimal micro-level decisions. Hence, new emerging complex dynamics of transport systems at macro-level – such as self-organisation and evolution – can be studied. Such system properties could not be fully studied using the traditional equilibrium assignment models.

Part I:

Research Stage I

Chapter 3

Minibus taxis in Kampala's paratransit system: Operations, economics, and efficiency

Chapter 3 objectives
This chapter aims at achieving the Research Objective 1.1 and RO 1.2 of the dissertation to answer research questions RQ1 and RQ2, respectively.

• ⇒**Research Objective 1.1**

> Describe the operations of minibus taxis in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

• ⇒**Research Objective 1.2**

> Estimate the system efficiency from the passengers' and driver's perspectives.

To answer research questions RQ1, RQ2, and subsequently achieve Objective 1.1 and Objective 1.2 of the dissertation, this chapter uses a three-pronged approach. First, we studied the operations of minibus taxis in Kampala's paratransit system, from basics like regulation, management, and routes, to the industry's unique and peculiar practices. Second, we studied and evaluated the economics of minibus taxis in terms of passenger fares, drivers' daily cash flow, profits, and we estimated the drivers' profitability index (PI) of the minibus taxi business. Third, we assessed the efficiency of the paratransit system from the passengers' and drivers' perspectives, and how it affects the overall minibus taxi operations and the subsequent operational economics. To estimate the system efficiency, we introduced the concept of *hold-back time* (t_h), which means the accumulated time a taxi waits at stops along a route (during a trip), waiting for and loading passengers. The hold-back time can be likened to the *dwell time* in scheduled public transport. However, hold-back time is unique to minibus taxis in organically-evolved, quasi-demand-responsive paratransit systems like those in Kampala, the Sub-Saharan Africa (SSA) and many developing cities of the Global South, because of their unscheduled nature. The hold-back time in such systems vary depending on the *occupancy* status of the taxi and the anticipated demand along the route.

3.1 Background to paratransit in Kampala

Across developing cities of Sub-Saharan Africa, transport authorities are struggling to fulfil the mobility needs of rapidly growing populations, especially the urban poor. The transport systems that are supposed to connect commuters to jobs, services and markets have limited capacity and are loosely regulated and inefficient (Behrens et al., 2015a; Daniel E. Agbiboa, 2018). Since the 1990s these typically quasi-demand-responsive services – referred to in the academic literature as "paratransit" or "informal transport" – have filled the gap left by the collapse of the colonial-era state-owned transport companies (Cervero and Golub, 2007; Ajay Mahaputra Kumar et al., 2008; Kumar, 2011). The introduction of the World Bank's structural adjustment policies in the 1990s, coupled with weak policies for regulating and enforcing paratransit licensing, created a barrier-free entry into the informal transport industry in response to demand. The paratransit industry, consisting mostly of fourteen- to twenty-seater minibus taxis, therefore expanded to dominate public transport in many cities of SSA (Booysen et al., 2013). In Uganda, two bus companies – the Uganda Transport Company (UTC) and the People's Transport Company (PTC) – provided public transport services in the 1970s and 1980s. However, after a litany of problems comprehensively documented by Kumar (2011), the companies collapsed in the early 1990s, paving the way for the organic evolution of privately-run paratransit with atomised ownership and dispersed financial capital.

Uganda's informal minibus taxi industry is a vital and lucrative component of the urban economy. It directly employs more than 100,000 people in Kampala (Uganda's capital city), and it provides many indirect jobs through the motor vehicle repair and servicing industry (Pablo, 2015; Goodfellow, 2017). In 2015, Kampala had 1.5 million residents and 16,000 minibus taxis that transported 82.6% of the commuters across its five divisions (i.e., Central, Kawempe, Makindye, Nakawa and Lubaga) (Vermeiren et al., 2015; KCCA, 2016; Aggrey, 2017). The remaining 17.4% of commuter travel was shared among private cars, busses, and motorcycle taxis (boda-bodas). Though many minibus taxis are not officially registered, Kampala's minibus taxi fleet is estimated to be growing at a rate of 5.4% annually (Pablo, 2015; Aggrey, 2017; Jean et al., 2018).

Although the minibus taxi ownership structure is fundamentally opaque, reports point to wealthy politicians owning fleets of minibus taxis, several private citizens owning one or two, and groups of individuals pooling funds to co-own one (Stewart-Wilson et al., 2017). The capital outlay of about $15,000 for a single second-hand minibus from Japan substantially contributes to the atomised ownership of minibus taxis in Kampala (Dorothy, 2018). The $15,000 start-up capital is out of reach for many Ugandans since Uganda's GDP per capita is only $642 (Dorothy, 2018). The minibus taxi business has a potential annual cash flow of $10,000 per minibus taxi and a profit of $35,000 over five years accrued to the owner (Aggrey, 2017; Dorothy, 2018). However, the operations of minibus taxis are opaque to the commuters and new drivers joining the industry; the economics of fares paid by the commuters and the daily cash flow to drivers and owners are not well documented; and the efficiency of paratransit in terms of passenger travel time has not been well studied.

3.2 Paratransit regulation and management in Kampala

Regulation in public transport universally covers three dimensions: quality regulation to ensure vehicle safety; quantity regulation to set targets for and limit the number of vehicles operating in the system; and fare regulation. The principal objectives of transport regulation are to ensure that: (1) services are operated in accordance with government policy; (2) demand for public transport is satisfied as far as possible; (3) standards of quality and safety are maintained; and (4) fares are controlled to affordable levels (Richard, 2005).

Efforts were made to regulate Kampala's paratransit system through statutory laws and regulations enacted by the Ugandan Parliament and loosely enforced by the Kampala Capital City Authority (KCCA) – the local governing body of Kampala City. The laws included: the Traffic and Road Safety Act of 1998; the Kampala Capital City Act of 2010; and the subsequent amendments in 2012 (Uganda Parliament, 1998, 2011, 2012). The laws were strongly resisted by paratransit operators, backed by influential politically connected owners (Goodfellow, 2010). These laws provided for the formation of a transport licensing board and taxi owners' and drivers' associations to manage minibus taxi affairs and enable collective bargaining between taxi owners, drivers, and the city authorities. The regulations and the means of enforcing them were not clear and thus left considerable room for discretion. For example, they left room for self-regulation through self-organised associations like the Uganda Taxi Operators and Drivers Association (UTODA) and the Kampala Operational Taxi Stages Association (KOTSA). Self-regulation of the quality of service, number of vehicles and fares charged by drivers generally worsened the quality of service provided by paratransit in Kampala. It also led to excessive competition between drivers seeking to maximise their profits, and left commuters with no choice but to travel in often overloaded and unroadworthy vehicles that seldom adhered to traffic rules. Goodfellow (2017) argues that the laxity in regulation enforcement and the subsequent chaos in the paratransit industry are not coincidental: the situation serves the economic and political purposes of the political elites who use the minibus taxi industry for political mobilisation.

3.2.1 Minibus taxi management

Management controversy, exploitation, inefficiency and political interference are rife in Kampala's minibus taxi industry (Goodfellow, 2010, 2012, 2017; Phillips and Mesharch, 2018). In the early 1990s the first attempt to manage the informal taxi industry through the monopolistic Uganda Taxi Operators and Drivers Association (UTODA) resulted in a deadlock that Goodfellow (2017) refers to as the "double capture". In the double capture, political elites infiltrated the taxi industry and subsequently, UTODA wielded enormous influence over the city authorities. UTODA engaged in behind-the-scenes multi-institutional informal bargaining, played one arm of the state against the other, and made it very difficult to implement transport policy in Kampala (Goodfellow, 2010). The results of this laissez-faire manner of running the transport industry were exploitation of drivers and an inefficient transport service. Today the bad practices of UTODA still haunt the minibus taxi industry: the regulations are not fully implemented; the drivers and other sector employees do not fully benefit from the industry; and commuters have to suffer an inefficient service riddled with delays, high and inconsistent fares, and lacking standardised routes and schedules information (Ndibatya et al., 2016; Goodfellow, 2017).

3.3 Methods and data

This section presents our economic and efficiency metrics, data collection process, and data processing methods. The efficiency of a public transport system can be analysed using both subjective measures, such as user opinions and satisfaction surveys, and objective measures, such as numeric values for attributes like load factor (percentage occupancy), travel time and total passenger-kilometres covered (Sampaio et al., 2008; Eboli and Mazzulla, 2011, 2012; Gorkem et al., 2014; Marcius et al., 2015). To study the operations of Kampala's paratransit system, both qualitative and quantitative data was collected between January and March 2016. The collected data was associated with pre-selected paratransit travel attributes: taxi ranks, stops, routes and route-related attributes such as fare, hold-back time, commuter waiting time and minibus taxi occupancy. The conditions in Kampala's minibus taxi system have not changed substantially since 2016. Therefore, the data that was collected and the results discussed in the subsequent sections are still relevant.

3.3.1 Metrics

We used *economic metrics* – minibus taxi fares (β), taxi occupancy ($\mathcal{O}$) and drivers' profitability index (PI) – and *efficiency metrics* – passenger waiting time (t_w), taxi hold-back time (t_h), operating speed (v_o) and commercial speed (v_c) – to study the operational economics and to estimate the efficiency of the minibus taxi transportation system in Kampala's paratransit. We define the metrics used for this study in Table 3.1.

The hold-back time (see Table 3.1) has a knock-on effect on the trip duration for passengers passively completing their journey in the taxi. This concept is similar to a bus waiting at a stop mid-journey due to being too early according to its schedule. But, in the minibus taxi case, the taxi stops for an unspecified time because the driver considers the vehicle occupancy to be too low to make the trip profitable.

3.3.2 Data collection methods

We adopted a participatory observation data collection method in which fieldworkers acted as passengers. They used Global Positioning System (GPS)-enabled devices to record minibus taxi movements on pre-selected routes. Global Positioning System data is widely used by researchers to study movements because it provides precise spatiotemporal characteristics of travel (Usyukov, 2017; Jonker and Venter, 2019). While on board and en route to the selected destinations, the fieldworkers also engaged in open-ended questioning of drivers to obtain qualitative information such as minibus taxi industry practices, driver revenue and expenses.

We used two quality assurance and control strategies described by Whitney et al. (1998) to preserve the data integrity, to detect and prevent errors in the collected data (such as sampling errors), and to ensure the scientific validity of the results. First, the error prevention strategies included talking to key people in the paratransit industry to get a general understanding of minibus taxi operations before designing the data collection instruments and recruiting data collection assistants (volunteers) who were familiar with the paratransit system. We trained the volunteers in the use of GPS devices and the data collection mobile application before starting the data collection process. Second, the error detection strategies used during and after this process involved a daily early morning meeting with the volunteers to update and ensure data consistency, exit interviews with each volunteer to record further critical observations, and redoing (or validating) routes that had erroneous data.

3.3.2.1 Sampling methods and data

We used stratified sampling to select 155 routes from the 307 routes distributed across the five divisions of Kampala as follows: Central-99, Kawempe-10, Makindye-12, Nakawa-10 and Lubaga-24 (refer to the "*Routes count*" column in Table 3.4 and Figure 3.5g for the detailed distributions of sampled and studied routes). Four data collection volunteers participated in the data collection exercise. The volunteers used a standard data collection instrument to collect data about the routes, i.e., route name, route fare, stops, hold-back time, waiting time and occupancy. Additionally, the volunteers randomly engaged in informal chats with 54 minibus taxi drivers and collected extra data about the taxi industry practices, drivers' revenue and expenditure. We also talked to the leadership of KOTSA and KCCA to get data about minibus taxi ranks, routes, the general operations, and regulations of minibus taxis in Kampala.

Table 3.1: Description of economic and efficiency metrics

Metric	Description
Minibus taxi fare (β)	The amount of money paid by a minibus taxi passenger for a complete one-way trip on selected a route. Note: Full trip fare was paid even for trips abandoned midway.
Occupancy ($\mathcal{O}$)	The number of passengers on board a minibus taxi as a percentage of the total taxi capacity. Note: only fourteen-seater minibus taxis were studied.
Profitability Index (PI)	The profitability index represents the relationship between the minibus taxi driver's revenue and expenses for day under study. Given one minibus taxi for one day, the driver's profitability index is computed using the equation $PI = \frac{\sum_{i=1}^{n} E_i}{R}$, where E_i is the expense incurred on a single item and R is the total daily revenue. R is computed as a function of (i) trip-based passenger fare (β), (ii) the average occupancy per trip ($\mathcal{O}$), (iii) the total number of trips per day (γ).
Waiting time (t_w)	The total time a passenger waits for a taxi at a stop or taxi rank.
Hold-back time (t_h)	The accumulated time a taxi stays at stops along a route waiting for (or in anticipation of) passengers. It includes the time spent at the stop of origin loading the minibus taxi before embarking on the trip.
Operating speed (v_o)	The average speed at which a minibus taxi could travel from origin to destination without stopping en route. When computing the operating speed, the hold-back time was excluded.
Commercial speed (v_c)	The overall average speed of the minibus taxis during the trip, including the time spent at the stops (hold-back time). $v_c = \frac{d}{T}$, where, d is the total trip distance, and T is the total time taken to complete a trip.

3.3.2.2 Participatory observation

We recruited and trained the volunteers and deployed them to collect geospatial data using a customised mobile application (GoMetro Pro) that was pre-installed on their smartphones running on an Android operating system (GoMetro, 2016). They travelled around the five divisions of Kampala to locate the gazetted and ungazetted taxi ranks and document the GPS coordinates and major routes originating from the taxi ranks. At some taxi ranks the routes and fares were written on small placards and carried from one minibus taxi to another in order of departure sequence (see the yellow placard in Figure 3.1b); at others, the volunteers asked drivers to provide information about the routes. The volunteers rode in taxis from the taxi ranks to the destinations on pre-selected routes and recorded aspects of interest to our research such as stop location and hold-back time. Shortly before setting off, they recorded data about the taxi fare, waiting time and the number of passengers in the taxi. During transit, they recorded data about the location of stops and the hold-back time at each stop. The GoMetro Pro mobile application automatically recorded the minibus taxi's GPS locations every 30 seconds, and these were later processed into the route's GPS profile. The volunteers also observed the drivers' behaviour during the journey, such as the hand signals used for communication between drivers and passengers on the road.

3.3.3 Data processing and analysis

We used a language-neutral Protobuf protocol developed by Google to serialise the collected data for transfer to custom-developed Python libraries to be cleaned, transformed and loaded into other Python scripts for further analysis (Blyth et al., 2019). During the transformation process, we developed and implemented several algorithms in Python programming language, such as a map-matching algorithm to remove multi-path errors in the data. However, in this chapter we report only the results from these algorithms.

Preliminary data analysis was done using the Quantum Geographic Information System application to map the geospatial layout of minibus taxi routes, taxi ranks and stops, relative to the administrative regions (divisions) and sub-regions (parishes) of Kampala, as shown in Figure 3.2. We developed other customised Python libraries to analyse further the relationships between routes, route lengths, fares per route, passenger waiting time and minibus taxi hold-back time. In Section 3.4, we present the results from the field study and data analysis for minibus taxi operations, economics, and efficiency.

3.4 Results

We present the results in three subsections below, i.e., minibus taxi operations, minibus taxi economics, and minibus taxi efficiency.

3.4.1 Minibus taxi operations

Kampala's minibus taxi operations are quasi-demand responsive, being mostly based on, or in response to, passenger demand. The routes, stops and schedules are not static but evolve according to passenger demand and drivers' prior knowledge or anticipation of demand.

3.4.1.1 Minibus taxi ranks and stops

Kampala has seven officially gazetted minibus taxi ranks (locally known as "taxi parks"), and 307 routes (locally known as "stages") originate from the taxi ranks to destinations across Kampala city and the neighbouring districts. The taxi ranks include: Old, New, Kisenyi, Namayiba, Usafi, Namirembe and Natete Taxi Park. Of the 307 routes, KCCA manages 122 and KOTSA manages 185 routes. Figure 3.1b shows a typical minibus taxi rank in Kampala (the Old Taxi Park). Table 3.2 summarises the minibus taxi ranks, stops, and routes originating from the five divisions of Kampala (Central, Kawempe, Makindye, Nakawa and Lubaga – see Figure 3.2c). The routes in Table 3.2 indicate the routes whose origins we documented during the study. They are according to the frequency of minibus taxi departures, i.e., high (every 20 minutes), medium (every 40 minutes) and low (every 60 minutes). The taxi ranks and stops in Table 3.2 are according to the frequency of passenger pickups, drop-offs, or departures in the case of taxi ranks. If the pickups/drop-offs were frequent at a stop (every 5 to 30 minutes), we regarded it as a "major stop", and one with fewer (every 30 to 60 minutes) as a "minor or informal stop".

All routes are supposed to originate from the taxi ranks. However, from interacting with minibus taxi drivers, we established that – in response to demand and for passenger convenience – several illegal origins are scattered around the city at ungazetted informal stops, such as: Clock Tower, Mini Price, Namirembe Road, City Square, Nasser Road, Mutaasa Kafeero, Mega Standard, and many others. Each route attached to a taxi rank is managed by a committee comprising a chairman, vice-chairman, secretary, treasurer, and welfare officer, all selected from drivers of taxis attached to the route. The committee resolves disputes, registers new drivers joining the route and manages the welfare needs of the members, such as loans and contributions in case of the death of any of its members. We did not find any formally documented details about the routes and the stops. However, the drivers we interviewed described the routes according to the significant towns they go through. Each route from the taxi rank operates on a first-in, first-out (FIFO) queuing system based on a ticket number booked in advance. The driver with the lowest ticket number loads passengers first. Two people operate each minibus taxi: a driver and a conductor. The conductor is responsible for touting passengers, negotiating, and collecting the taxi fares.

We did not find documentation indicating the presence of formally gazetted minibus taxis stops along the roadsides in Kampala. However, we observed that there were several stops along the major roads. We learned – from minibus taxi drivers and other officials – that the roadside stops are organically established (informally by drivers) according to passenger demand, often because of increased economic, leisure and other passenger travel attracting activities. When the demand at the location diminishes, the stop ceases to exist. Thus, we categorised the stops as either formal or informal depending on the frequency of pickups and drop-offs at the stop (see Table 3.2). We observed with keen interest the presence of waiting areas for motorcycle taxis (known as "boda bodas") at most minibus taxi stops (major and minor), as can be seen in Figure 3.1a. We could not verify independently whether the boda bodas attracted the minibus taxi stops or the stops attracted the boda bodas. Every major minibus taxi stop has two to four self-appointed wardens who often tout for passengers on behalf of the minibus taxi drivers. They are given a commission for each pickup. The commission is negotiated according to the number of passengers picked up from the stop: it ranges from \$0.11 to \$0.3 per minibus taxi.

Three basic characteristics of a minibus taxi paratransit system can clearly be seen in Figure 3.1a. The first is the informality of the stop: there is no signage to say it is a stop, but drivers stop there because of their previous knowledge of passenger demand at that location. The second is the presence of a pseudo-modal exchange centre between two modes of transport: motorcycle taxis and minibus taxis. We saw boda bodas dropping off and picking up passengers from the location. The third is the dominant presence of minibus taxis on the roads during peak hours (07:00 - 9:00 and 17:00 - 19:30). Only three ordinary vehicles can be seen in Figure 3.1a, as opposed to at least 12 minibus taxis on the same road segment. However, we observed that most of the minibus taxis were only half full. This gives an idea of the low load factor (taxi occupancy) which is an indication of the minibus taxi system inefficiency. We later confirmed (from our results in Table 3.4 and Figure 3.4a(v) that the average occupancy for minibus taxis in our sample was 69%.

3.4.1.2 Mid-trip change or abandonment of network routes

The minibus taxi route network consists of the major and minor nodes (gazetted and ungazetted taxi ranks) that are the most common origins and destinations for Kampala commuters. The nodes are interlinked with a series of stops (intermediate nodes), some known and gazetted, others ad-hoc and organically evolving into major stops if passenger demand grows. Figure 3.2a shows the distribution of taxi ranks and major and minor stops and 3.2b and 3.2c show the aggregated number of routes originating from parishes

Table 3.2: Number of minibus taxi ranks, stops and routes (categorised according to the frequency of taxi departures) originating from the five divisions of Kampala.

		Taxi ranks		Stops		No. of routes & categories			
Code	Division	Major	Minor	Major	Informal	High	Medium	Low	Totals
101	Central	3	6	53	194	55	17	50	122
102	Kawempe	0	2	20	91	6	4	2	12
103	Makindye	0	4	41	102	4	5	6	15
104	Nakawa	0	1	19	241	3	3	6	12
105	Lubaga	0	2	19	16	14	12	4	30
	Totals	3	15	152	644	82	41	68	
		18		796		191			

(a) (b)

Figure 3.1: Two types of stops showing the taxis, boda bodas, route signage and touts. (a) An informal stop at Port Bell Road; b) A formal minibus taxi rank (Old Taxi Park).

and divisions, respectively (the circles represent parish and division centres). Minibus taxi movements through the route network follow a near-straight course in some areas and a winding course in others, as shown in Figure 3.3a. Several factors are responsible for such behaviour: passenger demand (drivers wander off the main route looking for passengers); the presence of traffic police on the main route; the state of the roads; requests by passengers in the taxi; weather conditions; and sometimes the premature termination of the trip in pursuit of a new route considered to be more profitable. However, we observed that most of the diversions were made in search of passengers. Figure 3.3a shows the winding characteristics of taxi movements on selected routes. We observed that in zones Z_1, Z_2 and Z_3 drivers diverted off the main route in search of passengers. In zone Z_4 the driver abandoned the trip before getting to the destination originally communicated to the passengers and abruptly ordered all of them to disembark. He then diverted to a new route, went south a little way searching for passengers and then headed to zone Z_3, where he resumed the search.

3.4.1.3 Industry practices

Minibus taxi drivers have developed a host of strategies to improve profits from trips and avoid getting caught breaking the traffic rules. They believe these strategies to be moderately effective if correctly interpreted and applied.

i) ***Hand and headlamp signalling***

Minibus taxi drivers in a paratransit system frequently signal each other using hand signs and headlamp flash signals for the taxis coming from the opposite direction. Some of Kampala's hand signs are similar to those documented in Johannesburg by Susan Eve Woolf (2014), apart from some local differences in meaning. Taxi hand signalling is a useful language developed out of a desperate need for transport amongst multi-cultural and multi-lingual city travellers (Susan Eve Woolf, 2014). Each hand sign conveys a message, such as a warning, a clue to passenger demand

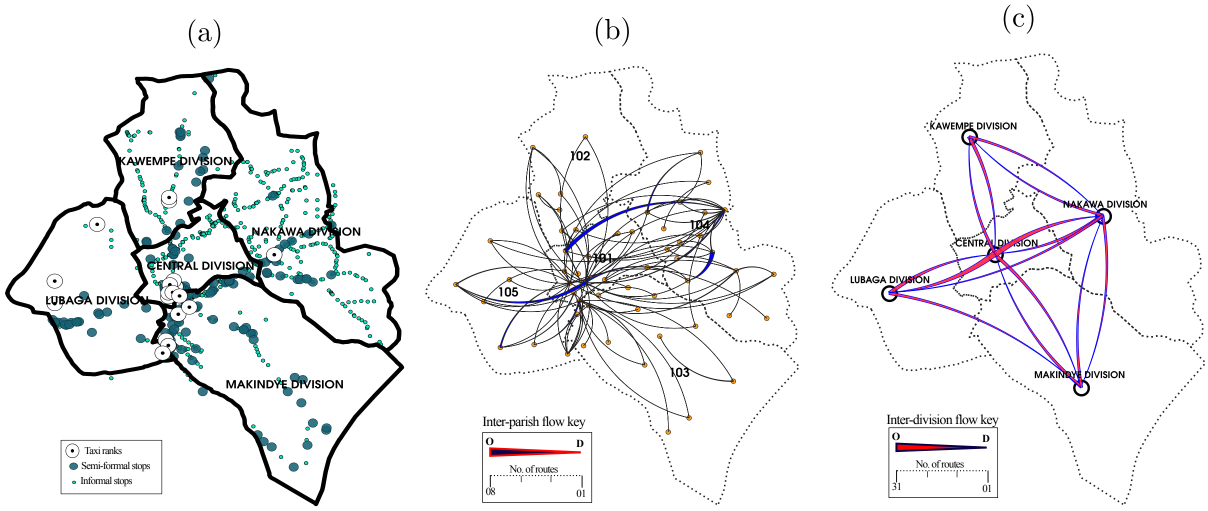


Figure 3.2: (a) Spatial distribution of taxi ranks and stops; (b) inter-parish routes flows; (c) inter-division routes flows. Note: The key shows the origin and destination routes flow count.

or an indication of the taxi's destination. Figure 3.3b shows five of these. Signal (i) is usually for driver-to-driver communication and is associated with a hilly terrain ahead (usually within about three kilometres). It means that after the indicated number of hills there is a high probability of passenger demand. If preceded by a double headlight flash signal, it warns of the presence of traffic officers or mobile speed trap camera ahead, alerting the other driver to reduce speed or drop off any excess passengers. Signal (ii) is for communication both driver-to-driver (indicating the presence of traffic police or a speed camera usually about 4 km ahead) and driver-to-passenger (indicating to waiting passengers that the taxi is headed to destinations far out of Kampala). Signal (iii) is used for communication both driver-to-driver (indicating the presence of traffic police or a speed camera about 1 km ahead) and driver-to-passenger (indicating that the taxi is about to reach its final destination, usually less than 2 km away). Signal (iv) is a driver-to-driver signal indicating the presence of many passengers waiting, usually within a distance of 1 km. Signal (v) is a driver-to-passenger signal indicating that the taxi is about to take a detour off the main route, usually to the informal settlements.

ii) ***Passenger touting strategies***

Passenger touting is mainly done by taxis circulating within the town rather than heading for distant destinations, especially those that originate from the informal roadside stops within the city. Taxis usually start the trip with few or no passengers in anticipation of collecting some en route, a strategy referred to locally as "*okuvuga ekkubo*" (random passenger search). To ensure a profitable trip, various complementary sub-strategies are used. One such is called "*okubala gap*" (strategic demand estimation), where the driver observes the taxis coming from both directions to either receive a signal of passenger demand ahead or estimate the presence of demand ahead on the basis of the number of competing taxis heading in the same direction. Another is a sub-strategy called "*okukyeebakamu*" (random back off) where, depending on estimated or received negative feedback from the *okubala gap* sub-strategy, the driver interrupts the trip and waits at a strategic stop for a random period to allow for commuter demand replenishment before continuing with the trip.

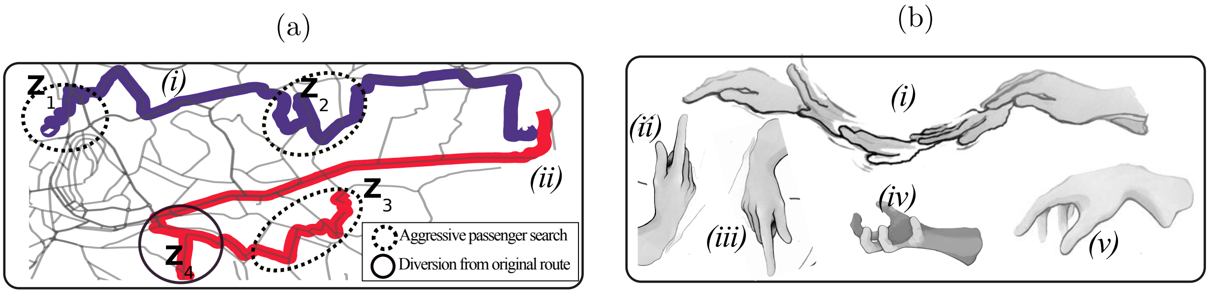


Figure 3.3: (a) Movement characteristics and route abandonment; (b) Gestures used by drivers (Susan Eve Woolf, 2014).

3.4.2 Minibus taxi economics

We looked at the economics of the minibus taxi operations from both the passengers' and the drivers' perspectives. Passengers are concerned with the fares they must pay; drivers are concerned with their overall cash flow and profit margins to keep their taxi business afloat. Figure 3.4b shows a positive correlation between hold-back time and route fare (r=0.299) and between waiting time and route fare (r=0.256), and a negative correlation between waiting time and route length (r=-0.265). Figure 3.5 shows the distributions of each attribute between and within divisions.

3.4.2.1 Passenger fares

We related our minibus taxi fare data to the length of the routes (total distance travelled from origin to destination). Table 3.4 shows the average route lengths, average fare per one-way trip, and approximate fare per kilometre for the inter-divisional routes in Kampala. Figures 3.4a(*i-iii*) show the general distributions of route lengths, route fares and fare per kilometre, and Figures 3.5a and 3.5b show the distributions of route lengths and route fares categorised according to the origins and destinations. The distances range from 1.2 to 11.8 km.

The average cost of travel by minibus taxi at the time of our study (January–March 2016) was \$0.55 (UGX 1,980) and the average fare per kilometre was \$0.12 (UGX 432). Routes within Makindye (103–103) were the least expensive at \$0.22 (UGX 799); routes from Kawempe to Nakawa (102–104) were the most expensive at \$1.08 (UGX 3,884) while routes from Central to Lubaga (101–105) were moderately priced at \$0.46 (UGX 1,656). The taxi fare per unit of distance travelled (fare per km) represents the unit cost of travel by minibus taxi, which ranged from \$0.06 (UGX 216) for trips within Makindye to \$0.28 (UGX 1,008) for those within Kawempe, as shown in Table 3.4. There is a positive correlation (r = 0.406) between route length and route fare as shown in the scatter matrix in Figure 3.4b. Figures 3.4a(*ii-iii*) show the general distribution of route fares and fare per unit distance (fare per km), respectively, while Figure 3.5b shows the division-level distributions of minibus taxi route fairs categorised according to division origin and destination pairs.

3.4.2.2 Drivers' profitability index (PI)

As noted earlier, Kampala taxis are privately owned by sole and multiple proprietors. After acquiring a taxi, the owner usually puts it up for hire by drivers. Drivers and taxi owners often execute an agreement (sometimes unwritten) where the driver rents the taxis and remits to the owner a daily rental fee ranging from \$26 to \$30 in addition to paying for the daily running costs of the taxi such as fuel, washing, overnight security and a commission to the tout. The owner pays for the routine vehicle maintenance and a monthly fee (\$32) to the city authorities for the right to operate in the city.

Preliminary results from analysing data about minibus taxi occupancy, route length and taxi fares (see Table 3.4), indicated that the minibus taxi driver's revenue is generally low. The average fare per kilometre is \$0.12, and the average minibus taxi occupancy is 69% (10 passengers). It means that, for an average 5km trip, a driver earns \$6, which represents a 29% revenue loss due to low occupancy. Figures 3.4a(*ii-v*) show the general distribution of route fare, fare per km and minibus taxi occupancy. There were some cases of overloading as shown in Figure 3.4a(v). Where taxi occupancy exceeded 100%, there is a positive correlation ($r = 0.46$) between route fare and route length.

Based on the preliminary results about route fare, fare per km, and occupancy, we estimated the minibus taxi drivers' profitability index. Table 3.3 summarises the estimated minibus taxi driver's daily expenses and revenue for an average 5 km route. The "Base" column shows the driver's revenue and profitability index on a normal fifteen-hour working day; the "Practice" column shows the driver's revenue when some strategies described in Section 3.4.1 are used; and the "Objective" column shows the revenue a taxi driver could ideally achieve, given an improvement of the system with all 14 passenger seats occupied and 15 trips per day – an objective expressed by the drivers who work for approximately 15 hours a day.

3.4.3 Minibus taxi efficiency

For this study, we focused on passenger waiting time, hold-back time, and travel time as measures of efficiency, and we assumed all other factors to be constant. The waiting time (time spent at a stop or taxi rank waiting for a minibus taxi) and hold-back time were recorded by the data collection assistants per route. The figures shown in Table 3.4 are of all-day averages for all the routes studied during the research period. The travel time, operating speed, and commercial speed presented in this section were computed from the timestamped routes profile GPS data collected using the GoMetro Pro mobile application.

Conventionally, travel time is a function of velocity and the geometry of direct and subsidiary routes (Sampaio et al., 2008). In a quasi-demand responsive paratransit system, however, travel time is greatly influenced by the passengers' waiting time and the drivers' hold-back time. We thus included these in our analysis. Table 3.4 shows the average waiting time, average hold-back time, average hold-back per km, operating speed, and commercial speed for various minibus taxi routes within and between Kampala's divisions. Accordingly, the waiting time in Kampala's paratransit system ranges from 22 minutes (Central to Nakawa) to 59 minutes (Central to Lubaga), with an overall average waiting time of 39 minutes. We found that minibus taxis from Central to Lubaga spent less time holding back and waiting for passengers than those from Kawempe to Nakawa –

Table 3.3: Driver's daily cash flow and estimating the profitability index for a single rented minibus taxi.

Average driver's expenses per day		Average driver's revenue per day			
Item	Cost ($)	Item	Base	Practice	Objective
Taxi rent (E_1)	$27.80	Occupancy per trip ($\mathcal{O}$)	10	11	14
Fuel (E_2)	$16.70	Avg fare per passenger (β)	$0.45	$0.55	$0.45
Washing bay (E_3)	$2.80	Trip duration (in hours) (ϱ)	1.5	1.7	1
Security (E_4)	$11.10	Trips per day (γ)	10	11	15
Touts commissions (E_5)	$1.10	Total revenue per trip ($\lambda = \mathcal{O} \times \beta$)	$4.5	$6.05	$6.3
Total expenses $E = \sum_{i=1}^{5} E_i$	*$59.40*	*Total daily revenue* ($R = \lambda \times \gamma$)	*$45.0*	*$66.55*	*$94.5*
		Profitability index ($PI = \frac{R}{E}$)	0.76	1.12	1.59

Note: (i) Profitability index is computed as a ratio of total revenue to total expenses. (ii) A driver works for 15 hours a day. (iii) The Objective column assumes filling every seat (14) for an objective expressed by the driver of 15 trips per day. (iv) Occupancy $\mathcal{O}$ is the average number of passengers in a taxi per trip.

Table 3.4: Average operational attributes for origin–destination (O-D).

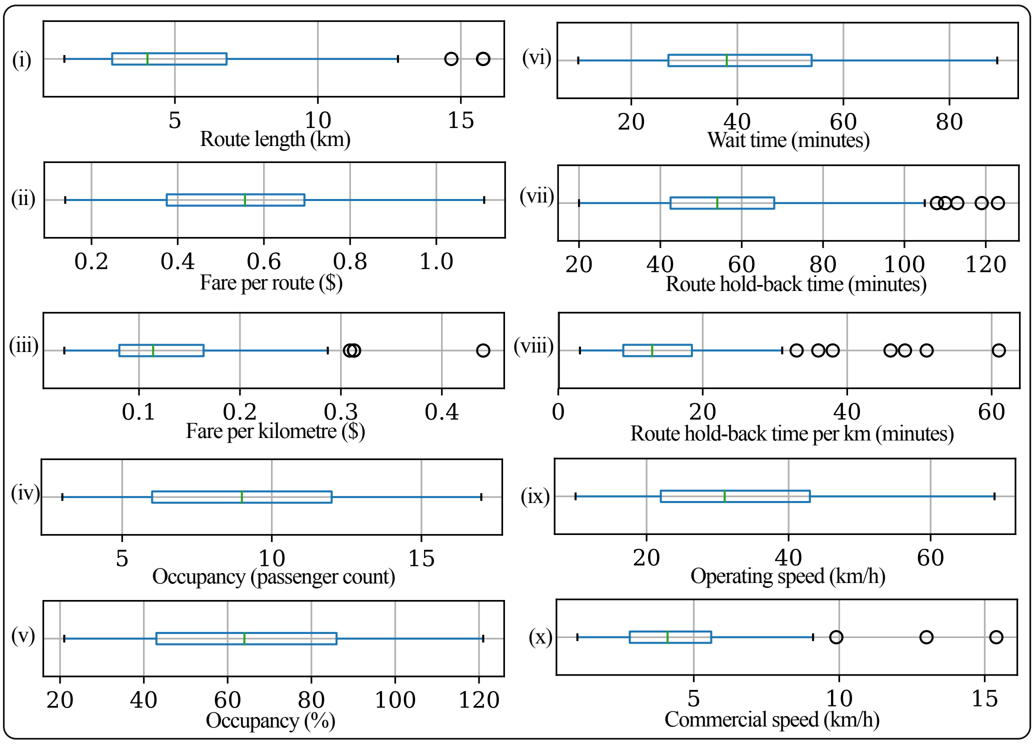
			Economic metrics (averages)			Efficiency metrics (averages)				
O–D (div IDs)	Routes count	Length (km)	Fare ($)	Fare per km ($)	$\mathcal{O}$ (%)	t_w (min)	t_h (min)	t_h per km (min)	v_o (km/h)	v_c (km/h)
101–101	13	2.5	0.37	0.15	52%	31	41	17	32	3.4
101–102	17	3.2	0.68	0.22	60%	44	53	17	27	3.2
101–103	13	5.2	0.76	0.17	71%	51	52	12	32	4.8
101–104	29	6.5	0.58	0.10	80%	22	63	12	35	5.3
101–105	27	3.7	0.47	0.14	59%	59	40	11	31	4.7
102–101	7	6.3	0.40	0.07	90%	25	52	11	36	5.4
102–102	2	1.2	0.35	0.28	32%	36	66	54	31	3.1
102–104	1	11.8	1.08	0.09	57%	34	110	9	51	5.7
103–101	7	4.7	0.69	0.16	67%	51	64	15	35	3.9
103–103	3	5	0.22	0.06	78%	26	45	14	46	6.3
103–104	2	8.7	0.70	0.08	78%	48	84	10	44	5.5
104–101	6	8.7	0.44	0.06	85%	23	62	9	43	6.8
104–103	1	8	0.71	0.09	71%	54	79	10	60	5.5
104–104	3	6.9	0.38	0.07	95%	29	54	12	45	7.5
105–101	24	5.3	0.46	0.10	63%	48	84	20	27	3.3
Mean μ		5.85	0.55	0.12	69%	39	63	16	38	4.96
STD σ		2.9	0.21	0.07	28%	17	22	9	15	2.2

Summary statistics

	Min	Max		Min	Max
Route length	(102–102) 1.20	(102–104) 11.8	*Wait time (t_w)*	(101–104)22	(101–105) 59
Routes fare	(103–103) 0.22	(102–104) 1.08	*Hold-back (t_h)*	(101–105)40	(102–104)110
Route fare/km	(104–101)0.057	(102–102)0.281	*Hold-back/ km*	(102–104) 9	(102–102) 54
Occupancy ($\mathcal{O}$)	(102–102) 32%	(104–104) 95%	*Speed (v_c)*	(102–102)3.1	(104–104)7.5

Note: IDs for the divisions are Central 101, Kawempe 102, Makindye 103, Nakawa 104, Lubaga 105.
t_w - Waiting time,t_h - Hold back time,v_o - Operating speed, v_c - Commercial speed,
$\mathcal{O}$ - Percentage minibus taxi occupancy.

(a)



(b)

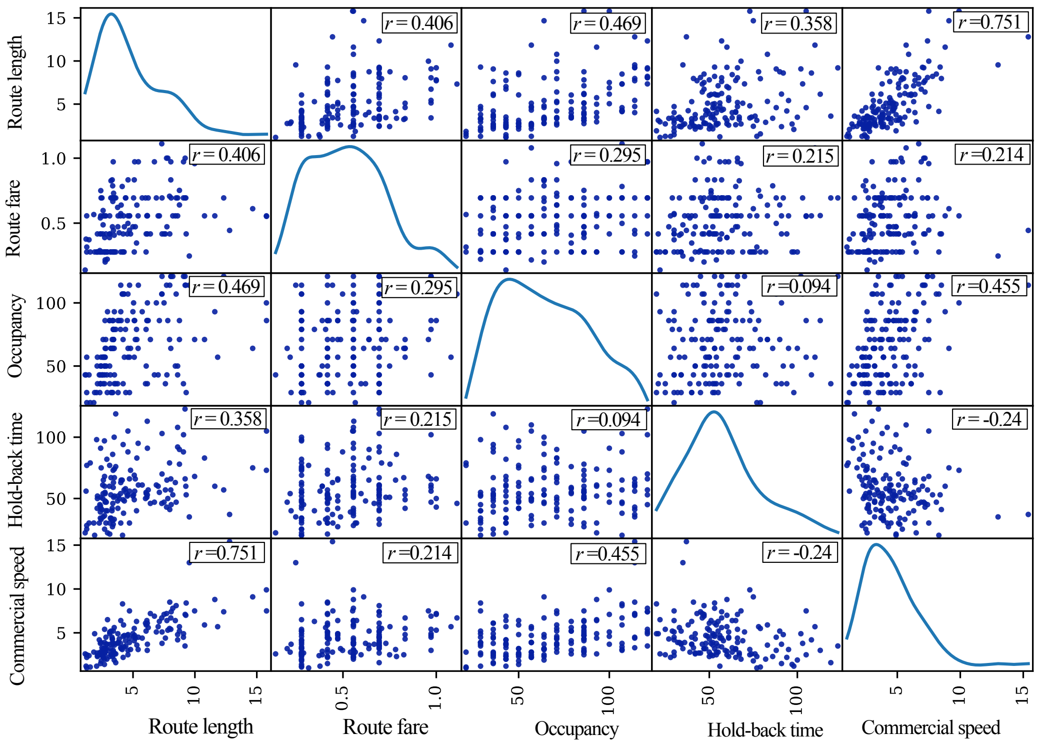


Figure 3.4: Operational attributes of division-level origin–destination routes. (a) Scatter matrix plot and corresponding correlation coefficients for different routes' attributes; (b) General statistical distribution of operational route attributes (n=155 routes).

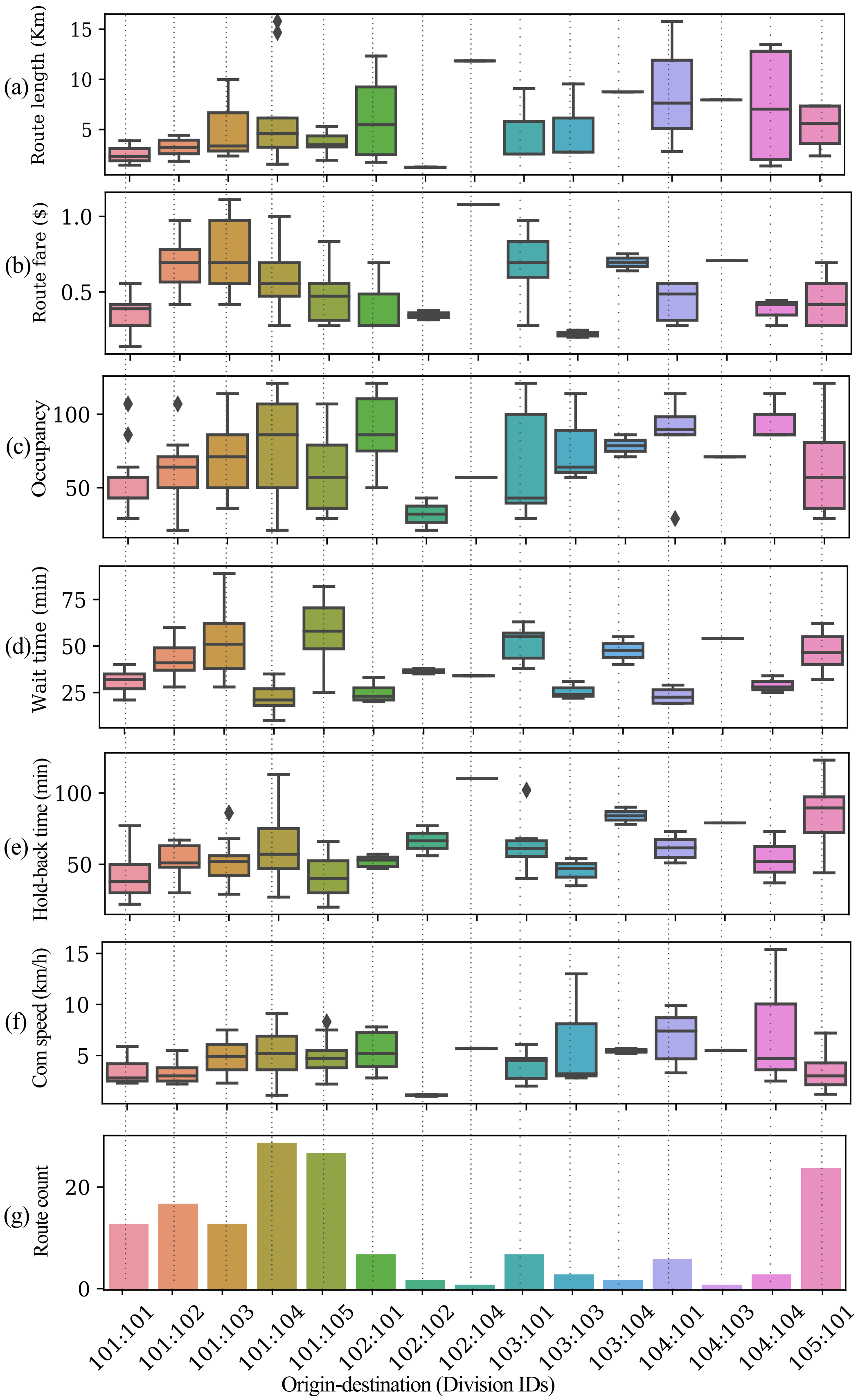


Figure 3.5: (a - f) Detailed independent distributions of route attributes; (g) Barplot showing origin-destination route counts.

40 minutes as compared with 110 minutes. The hold-back time per kilometre travelled ranged from 9 minutes (Kawempe to Nakawa) to 54 minutes (within Kawempe) with a median of 16 minutes.

We found that the general operating speed for minibus taxis in Kampala ranges between 10 km/h and 69 km/h with a mean and standard deviation of 38 km/h and 15 km/h, respectively. In contrast, the commercial speed is low, ranging between 3 km/h and 15.4 km/h with a mean and standard deviation of 4.9 km/h and 2.2 km/h.

3.5 Discussion

We discuss our results according to our three major themes: operations, economics, and efficiency.

3.5.1 Operations

Minibus taxi operations in Kampala are characterised by informality in all their aspects: regulations, regulation enforcement, management, stops and routes. The minibus taxi routes and stops are not clearly established and labelled. The routes between two stops or ranks are difficult to present in static maps because they are not fixed: they vary according to demand, traffic conditions, competition, and sometimes drivers' preference. Ndibatya and Booysen (2020a) presented a static route map that shows an attempt by KCCA to develop a static public transit route map for Kampala in 2017. The route map consists of 190 stops and 110 minibus taxi routes. When compared with our results in Table 3.2, we found a difference of 606 stops and 81 routes that do not appear on the KCCA route map. The most probable causes of the difference in routes and stops count are either that the routes and stops have evolved because of changing passenger demand at different locations or that they were considered insignificant by the KCCA mapping team.

We observed desperate attempts by drivers to make the trips profitable through rudimentary strategies such as starting trips with no passengers (*okuvuga ekkubo*), strategic observation (*okubala gap*) and random back off (*okukyeebakamu*), as discussed in Section 3.4.1.3(ii). These strategies are not as effective as claimed by drivers. In fact, they may be responsible for the fluctuating minibus taxi occupancy and subsequent low profitability index presented in Table 3.3. The minibus taxi drivers adopt such ineffective strategies due to lack of proper minibus taxi scheduling, booking and demand forecast systems.

Kampala's public transport sector affects the livelihood of millions every day, especially the urban poor. It is dangerous to leave it to private management by manipulative wealthy elites (often referred to as "transport mafia"). Kampala can learn from cities like Accra, Lagos, and Cape Town, where paratransit regulation and reforms are gradually gaining acceptance among taxi operators. Accra and Lagos, for example, have combined regulation with financial support to overcome resistance to reforms. This arrangement includes an ownership reorganisation scheme whereby informal minibus owners form co-operatives and jointly invest in higher capacity buses. Cape Town authorities have plans to provide incentives to paratransit operators on selected routes to complement scheduled trunk services (du Preez et al., 2019). Alternatively, Uganda could implement a taxi recapitalisation policy similar to the one implemented by the South African government (Schalkwyk and Sadie, 2011) but avoid the mistakes that were made by that government by involving the paratransit operators during all stages of policy formulation. It is worth noting that the city authorities in Kampala are making strides towards engaging paratransit operators in transportation planning for the city (ODA, 2020).

The complementary role played by motorcycle taxis (locally known 'boda bodas') in Kampala's paratransit industry is interesting and deserves further research. Almost every major minibus taxi stop has a 'boda boda' stop nearby. Even informal minibus taxi stops (such as the one shown in Figure 3.1a) always act as terminal positions for boda boda trips. The boda boda stops act as inter-mode exchange centres, and they serve the passengers' last kilometre of commute. In a paratransit system, boda bodas are a necessary evil because of their ability to manoeuvre and to penetrate the sprawling townships' deeper locations that are often unreachable using other vehicles. They become a menace only when allowed onto highways, causing traffic interruptions and accidents. If regulated to serve only the first and last kilometre of commute, their role in paratransit would be substantial and a net positive.

3.5.2 Economics

Kampala's paratransit system operates a risky but generally profitable business model (to the taxi owners). It is characterised by restricted access to capital, no subsidies from the government, and exploitation of drivers, especially those who do not own taxis. Entry level capital for individual drivers and owners is mainly through personal savings, and soft loans from friends and family members. Drivers depend on passenger payments to cover all the operational costs. Most of the drivers rent the vehicles from owners (who pay for repairs) at fees that vary according to the vehicle's condition. While the taxi owners' cash flow is almost guaranteed, drivers are exploited, and they often make losses as illustrated by a low profitability index in Table 3.3 ("Base" column). To make profits, drivers work for long hours (15 hours and more per day) (see Table 3.3 "Practice" column). Vehicles are often shared by several drivers, leading to rapid degradation due to overuse, and sometimes overloading (see Figure 3.4a(v)). Paradoxically, the rapid vehicle degradation in the paratransit system generates many informal, indirect, and unstable jobs through the repair industry (Pablo, 2015).

3.5.3 Efficiency

We used two main operational attributes to measure the efficiency of minibus taxi transport in Kampala's paratransit system, i.e., the waiting and hold-back time. As summarised in Table 3.4 and illustrated in Figures 3.4a(vi-x), 3.5d, 3.5e and 3.5f we found that travel by minibus taxi was inefficient and characterised by long passenger waiting times (22 to 59 minutes), long hold-back times (35 to 110 minutes) and low commercial speeds (3.1 to 15.4 km/h). The result is that a large portion of minibus taxi commuters' travel time consists not of actual travel but of sitting in a stationary vehicle waiting for more passengers to fill up the minibus taxi. From the driver's perspective, the high hold-back time leads to fewer trips per day and thus a substantial loss in revenue resulting in a low profitability index as illustrated in Table 3.3.

We identified three factors that could be the root causes of the minibus taxi system inefficiency in Kampala. First, the absence of market entry controls. The collapse of the UTC and PTC in the early 1990s and the subsequent adoption of the World Bank structural adjustment policies left the public transport industry open to the free market forces of demand and supply (Kumar, 2011). With little or no entry controls, the minibus taxi industry emerged and boomed with fragmented ownership of often old vehicle fleets (Cervero and Golub, 2007; Kumar, 2011). To date, there are insufficient entry controls into the minibus taxis business. This often leads to oversupply or undersupply of low-quality vehicles and thus system inefficiency.

Second, inadequate regulation and enforcement thereof. Minibus taxis operations in Kampala are largely self-regulated (Goodfellow, 2010, 2017). The taxi drivers often determine individually the route to take for a particular trip, the fares to charge from the commuters (leading to fare variations shown in Table 3.4), the number of passengers to load, and when to take the vehicle for servicing (Ndibatya et al., 2016; Goodfellow, 2017). Self-regulation often leads to overloading, overpricing, trip abandonment (illustrated in Figure 3.3), driving unserviced vehicles and generally inadequate service provision. There is general laxity in regulation enforcement by KCCA and KOTISA, resulting in general system inefficiency.

Third, the lack of known minibus taxi scheduling and booking mechanisms. This leads to wide variations in taxi occupancy: sometimes taxis are overloaded while others are half loaded. Passengers and drivers depend on personal experience and sometimes on random guesses to determine supply and demand. Hence, the quality of service is poor, driver profits are low, and vehicle quality rapidly degrades, causing traffic jams and pollution. This has a knock-on effect on businesses: they lose efficiency because the waiting times and hold-back times prevent the workforce from getting to work on time.

3.6 Summary

In this chapter, we used economic metrics (i.e., taxi fares, occupancy, drivers' revenue and expenses) to estimate minibus taxi drivers' profitability index, and efficiency metrics (i.e., waiting time, hold-back time, and commercial speed) to estimate the efficiency of the minibus taxi transportation system in Kampala. We found that the driver profitability index is low – ranging between 0.76 and 1.12 – and the waiting and hold-back times are high – ranging between 22 to 59 minutes and 35 to 110 minutes, respectively. This indicates an overall minibus taxi system inefficiency. The absence of market entry controls into the minibus taxi business, coupled with inadequate regulation enforcement, poor minibus taxi scheduling and non-existent booking mechanisms, render Kampala's minibus taxi system inefficient to both the drivers and commuters. Furthermore, we found that the operations of minibus taxis are riddled with informalities, from management, regulations, to informal stops and routes.

Thus, we have described minibus taxis' operations in Kampala's organically-evolved, quasi-demand-responsive paratransit system, and estimated the system efficiency from the passengers' and drivers' perspectives. Hence, we have answered research questions RQ1 and RQ2.

Chapter 4

Characterising paratransit movements

Chapter 4 objectives

This chapter aims to achieve the research Objective 1.3 and further support Objective 1.2 of the dissertation to answer research questions RQ1 and RQ2.

- ⇒**Research objective 1.2**

> Estimate the system efficiency from the passengers' and driver's perspectives.

- ⇒**Research objective 1.3**

> Characterise the movement patterns of minibus taxis in Kampala's paratransit system.

To further support the answers to the research questions RQ1 and RQ2 (see Chapter 3) and achieve Objective 1.3, in this chapter we studied the operations of minibus taxis in terms of their movement patterns. To do this, we used floating car data (timestamped geo-localisation data collected by moving vehicles) to characterise the movement patterns (or trajectories) of minibus taxis in Kampala's organically-evolved, quasi-demand-responsive paratransit system. We were interested in discovering whether minibus taxi movement patterns were consistent with Lévy walk behaviour; whether the routes the taxis used changed topology or shape over time, in other words, evolved; and whether their movements could suggest anything about their level of determination when searching for passengers.

4.1 Introduction

In most African countries, minibus taxis are the backbone of public transportation. They transport more than 70% of the total urban travellers and dominate most social and economic aspects of urban mobility (Behrens et al., 2015a). They form part of the broader organically-evolved paratransit system that operates with little or no regulation in many developing cities of Africa and the Global South (Behrens et al., 2015a). Minibus taxi transport is flexible and semi-adaptive, with stops, schedules, fares, and routes primarily determined by demand (Klopp and Cavoli, 2019). Unlike traditional bus rapid transit

(BRT) systems that use buses on fixed routes and schedules developed a long time in advance, minibus taxi drivers in a paratransit system often plan their routes according to the occupancy status of the taxi and anticipated demand (Gauthier and Weinstock, 2010).

Rapid urban population growth (6% per annum) is reshaping urban settlements and changing economic and social population dynamics in Africa (Awumbila, 2017; Lucas et al., 2019). The population surge in cities, coupled with weak and non-transit-oriented city development policies will increase the problems of urban sprawl, scattered public amenities and unemployment. The mobility characteristics of urban dwellers will consequently change, triggering a change in minibus taxi movement characteristics in response: the static minibus taxi route maps proposed by Klopp et al. (Klopp and Cavoli, 2019) will no longer be useful. By exploring the evolution of minibus taxi routes in Kampala's paratransit system, our study could pave the way for solutions to the future minibus taxi travel problems.

Minibus taxis rarely get enough passengers to fill up before departure unless they start trips from the major taxi ranks (which are typically few, and travellers often shun them). They therefore search for passengers on the way to make the trips profitable. Sometimes, they wait ("hold back") at selected stops in anticipation of passengers turning up; sometimes they go off the main route to search for passengers in sparsely distributed places where they anticipate demand for their services. The taxis go up and down the streets in an apparently chaotic fashion, hooting repeatedly, calling out their destinations, randomly inviting pedestrians to board the taxi, and stopping anywhere to tout for potential passengers in total disregard of traffic and municipal laws. We suspected that analysis of these movements of minibus taxis would demonstrate a Lévy walk pattern during the passenger search process (shown in Figure 4.1a). A "Lévy walk" (a term we use synonymously with "Lévy flight") is a pattern of movements made by a random walker, where many short movements are randomly interspersed with long ones and occasionally very long ones, as illustrated in Figure 4.1b(ii) (James et al., 2011). In Figure 4.1a, a minibus taxi moves from origin (O) to destination (D) but in the process makes many detours to hunt for passengers in off-route locations L, S, S′ and L′.

The Lévy walk theory broadly combines an organisms' need for resources (e.g., food, shelter, or customers) and the need to reduce risks (e.g., from predators or competitors) with the density and renewability of resources to explain the organisms' movement in space. This study focuses on human movement where the minibus taxi driver represents the 'random walker' and the minibus taxi travellers (trips demand) represent the resources being searched for under some predatory risks such as police enforcement, and competition from other minibus taxis on the same route.

4.2 Literature and applications to this study

We divided the literature into three categories: the current status of paratransit in African cities, Lévy walk behaviour in animal and human movements, and spatial similarity analysis of mobility trajectories.

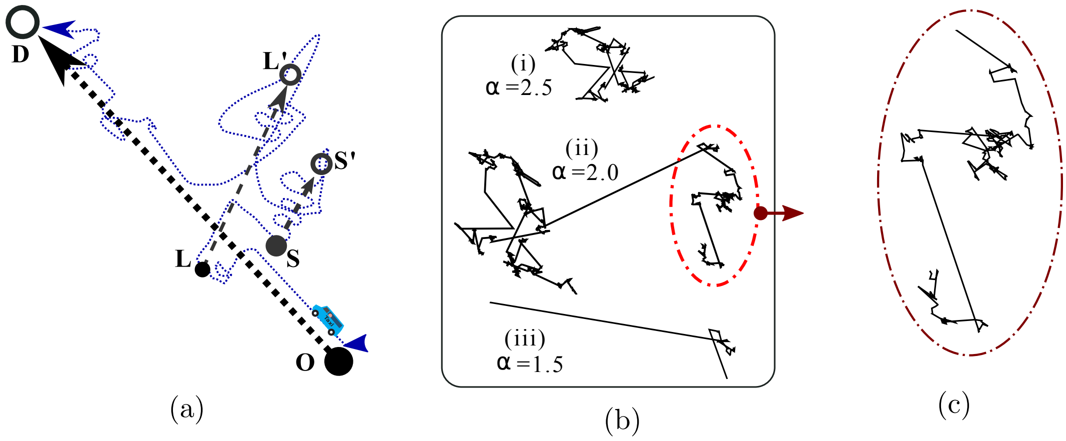


Figure 4.1: The concepts of Lévy walk (LW), and minibus taxi movement behaviour: a) Minibus taxi passenger search behaviour from origin (O) to destination (D); b) Lévy walks with different values of Lévy exponent α; c) Scale-invariant and fractal properties of a Lévy walk.

4.2.1 Current status of paratransit in African cities

Until recently, the term "paratransit" meaning "beyond standard transit" or "alongside of standard transit" was used (mostly in the United States of America) to refer to supplementary public transport services that do not have fixed routes or timetables but instead respond to travel demand and preferences and are often used by the elderly and the disabled. However, transport researchers have also adopted the term in the context of developing cities of Africa and the Global South to describe the informal transport that is synonymous with public transport in these cities (Behrens et al., 2015a). Paratransit in developing African cities is composed of diverse modes, such as minibus taxis (Booysen et al., 2013), tricycle taxis, bicycle taxis (Mutiso and Behrens, 2011) and motorcycle taxis (Diaz Olvera et al., 2019, 2016; S Kisaalita and Sentongo-Kibalama, 2007; Bradbury and Howe, 2002; Ehebrecht et al., 2018). In some African countries, motorcycle taxis dominate the modal share in terms of vehicle composition (e.g., in Lomé, Togo), but minibus taxis dominate the total share of passengers transported per day (Diaz Olvera et al., 2016; Lucas et al., 2019). In Kampala, for example, the Kampala Capital City Authority (KCCA) estimates that motorcycle taxis comprise 42% of vehicles and carry 9% of people, minibus taxis comprise 21% of vehicles and carry 82% of people. Private cars comprise 37% of vehicles and carry 9% of people (Evans et al., 2018).

There are five main actors involved in minibus taxis system: the owner, the driver, the conductor, the authorities, and the users (Booysen et al., 2013; Plano et al., 2020). The owner provides the vehicle, pays for the operating license and is responsible for the maintenance of the vehicle (Dorothy et al., 2016). The driver rents the minibus taxi from the owners at a pre-negotiated daily fee and makes operation-specific decisions such as, when to provide the service, the route for a given trip, and the trip fare depending on the demand and where to stop to pick up passengers (Dorothy et al., 2016). The conductor, if present, is responsible for touting and collecting fares from the passengers (Ndibatya and Booysen, 2020a; Plano et al., 2020). Most of the paratransit users in Africa's cities are not formally employed and thus tend to have variable and highly irregular commuting schedules and destinations (Ndibatya and Booysen, 2020b). This influences the movement patterns of minibus taxis in the paratransit system.

Several mapping projects have used floating car data to describe the routes taken by the informal paratransit minibus taxis in developing cities, such as Accra (Saddier et al., 2016), Nairobi, Maputo (Klopp and Cavoli, 2019), Kampala (Ndibatya et al., 2016), Dar es Salaam and Stellenbosch (Ndibatya et al., 2014). These projects have in some instances produced route maps, such as Digital Matatu for Nairobi (matatu referring to "minibus" in Kenya (Robert Heinze, 2018)) and the Mapas Dos Chapas for Maputo–chapa refers to "minibus" in Mozambique–as well as the standardised data in the general transit feed specification (GTFS) format used by developers to build mobile applications (Klopp and Cavoli, 2019). However, the paratransit mapping projects produced static maps, and the researchers ignored the possibility of changes in the routes that would render the maps irrelevant after less than five years. Thus, the need to explore the concept of route evolution in the minibus taxi system.

4.2.2 Lévy walk behaviour in animal and human movements

Movement by organisms is a biological process of great significance. In the reviewed literature, researchers have studied biologically motivated movement (searching for habitats, avoiding predators, or foraging) of organisms with cognitive abilities ranging from the relatively simple (e.g. bacteria), to the cognitively complex (e.g. humans), that demonstrates their ability to respond to external stimuli and memorise past movement experiences. The mechanisms by which organisms make movement-related decisions have evolved, as has the biological context that determines the "currency of fitness" or "reward" associated with the movement (such as net food intake, predatory risk, profit, or time savings). To optimise the "currency of fitness", models have been formulated. Of particular interest to us is the "Lévy flight" model developed by Paul Lévy, a French mathematician. Lévy described a particular class of random walks, in which the distance l travelled between events (referred to in this chapter as "steps") is drawn from a "heavy-tailed" and scale-invariant probability distribution defined by Equation 4.1 (Viswanathan et al., 1999; Reynolds, 2018).

$$f(l) \sim l^{-\alpha} \quad for \quad l \in [l_{min}, \infty), \tag{4.1}$$

where l is the step length and α (referred to as the Lévy exponent) is in the range $1 < \alpha \leq 3$ (Viswanathan et al., 2011).

A Lévy walk exhibits three main properties: the probability distribution of step lengths l is heavy-tailed; the turning angles θ between steps are normally distributed, and the step lengths fit strongly into the power-law probability distribution (defined by Equation 4.1). Figure 4.1c illustrates the scale-invariant property of a Lévy walk (zooming into a part of a Lévy walk trajectory (in Figure 4.1b(ii)) reveals a statistically identical substructure), while the Figures in 4.1b illustrate the effects of varying values of α on the Lévy walk. Values closer to $\alpha = 1$ lead to ballistic (near-straight) paths (Figure 4.1b(iii)), while values closer to $\alpha = 3$ lead to more Brownian behaviour (Figure 4.1b(i)).

Subsequently, researchers adapted the Lévy walk theory to explain the movement behaviour of cognitively complex organisms in space when searching for patchily and randomly distributed resources (Reynolds, 2015). There is empirical evidence of Lévy walks in the movements of foraging birds such as albatross (Shlesinger, 2006); animals, such as deers (Viswanathan et al., 1999), spider monkeys (Ramos-Fernandez et al., 2003) and grey seals (Shlesinger, 2006); bees, such as bumblebees (Edwards et al., 2007); and humans i.e., hunter-gatherers (Raichlen et al., 2014), fishermen (Bertrand et al., 2007), and pedestrians (Jiang et al., 2009).

Exponents of the Lévy walk theory in the early 2000s sought to explain how organisms optimise their search for sparsely distributed resources (such as food), sometimes under predatory risks. Viswanathan et al. combined the Lévy walk model with the optimal foraging approach to formulate and test the Lévy flight hypothesis (Viswanathan et al., 1999). This hypothesis states that "since Lévy flights optimise random searches, biological organisms must have therefore evolved to exploit Lévy flights" (Viswanathan et al., 2008). This paved the way for several predictions of optimal values of α (Lévy exponent in equation 4.1) based on the density of the resources a walker is searching for, and how renewable the resources are. Consensus was reached that the optimal value of the exponent α in the Lévy probability distribution, and hence the predicted movement pattern, depends on the "depleting" and "non-depleting" nature of the resource and their density relative to the random walker (Viswanathan et al., 1999; Ferreira et al., 2012). Furthermore, the optimal values of α approach 1 (ballistic movement with little change in direction) for depletable resources. For non-depletable resources, α depends on the target density, for sparsely distributed resources α is closer to 2, and for highly dense resources α is closer to 3 (Brownian motion) (Ferreira et al., 2012).

Lévy walk behaviour observed in human movements occur in vast contexts ranging from hunting and foraging among preliterate societies to myriads of contexts among modern and often urban societies. Evidence of Lévy walks in humans predates history as shown in raw material transport distances in the archaeological records (Perreault and Brantingham, 2011). Motivated by the need to search for food, shelter and avoid predators (dangerous animals), preliterate human movements exhibited Lévy walks (Raichlen et al., 2014; Bertrand et al., 2007). Urbanisation, industrialisation, and higher cognitive abilities among humans have diversified the contexts in which Lévy walk behaviour can be studied. These contexts are often determined by the travel purpose (such as travel to shop, work, school or leisure), travel mode, and spatial scale (Rhee et al., 2011, 2008; Brockmann et al., 2006; Cao et al., 2011; Scafetta, 2011). In all these contexts, evidence of Lévy walks has been found. For example, Lévy walks were exhibited in GPS traces from five different outdoor sites (Rhee et al., 2011); the circulation of bank notes (Brockmann et al., 2006); city cabs in Beijing (Cao et al., 2011); and long-range human displacements (from 1 to 1000 km) (Scafetta, 2011).

However, we did not find any research providing evidence (or absence thereof) of Lévy walk behaviour in minibus taxis in a paratransit system. We contend that the minibus taxi movements represent another context in which we can study Lévy walk behaviour among humans when searching for non-depletable patchily located and sparsely distributed resources. On the basis of the visual inspection of known minibus taxi trajectories illustrated in Figure 4.1a, we hypothesise that while searching for, picking up and dropping off passengers, minibus taxi movement may be consistent with Lévy walk behaviour.

4.2.3 Spatial similarity analysis of movement trajectories

The empirical literature on quantifying and analysing movement trajectory similarity is sparse and scattered across application domains and classes of moving objects. Güting and Schneider identified two classes: objects that maintain a constant shape while moving, such as animals, human beings and vehicles, which they call moving point objects, and those that change their shape, such as a forest fire, which they represent as polygons (Güting et al., 2005). This chapter is concerned with the former.

The shape of the trajectory is significant. It illustrates how a moving object "winds" its way through a spatial reference system, and it is quantitatively represented in terms of tortuosity, curviness, and fractal dimension (Ranacher and Tzavella, 2014). In this study we use only tortuosity (a property of a curve being tortuous or twisted, having many turns, or degree of winding). Researchers use the term "tortuosity" to distinguish between a planned, oriented, and effective behaviour (low tortuosity), and random search behaviour (high tortuosity) (Benhamou, 2004). We found no studies that quantitatively describe the spatial similarities and dissimilarities between minibus taxis' movement trajectories in a paratransit system, hence the need to fill the gap.

The literature we reviewed on the current state of paratransit in Africa, the similarity or dissimilarity between animals and humans that use the Lévy walk search optimisation strategy, and the spatial similarity measures of movement trajectories, revealed gaps in paratransit movement-related studies. In addition, the absence of similar studies as applied to minibus taxi movements in a paratransit system, led us to undertake this empirical study. It expands on some of the few existing and limited studies of the operations of minibus taxis in a paratransit system in the Global South (du Preez et al., 2019; Klopp and Cavoli, 2019; Ndibatya et al., 2016).

4.3 Methods

Having acquired and pre-processed the data from the minibus taxis, we used three methods to characterise their movement patterns. First, we modelled their trajectories as "walks" composed of sequences of linear steps, defined rules for determining successive steps, and then tested the Lévy walk hypothesis. Second, we compared the spatial distances of different minibus taxi trajectories to confirm or refute our route evolution claim. Third, we analysed the tortuosity (degree of winding) of the trajectories in case it might explain the effort drivers use to search for passengers. We used data collected from a sample of Kampala's taxis. To maximise the accuracy of our results, we assumed that all GPS points were located on the earth's surface, and we computed the distance between them using Vincenty's formula.

4.3.1 Data acquisition and pre-processing

Kampala, Uganda's capital, is home to one and a half million people scattered throughout five administrative divisions: Central, Kawempe, Lubaga, Makindye and Nakawa. Commuters from the latter four and beyond converge mainly in the Central division for work, shopping, leisure, and school (ITP, 2010). Minibus taxis, which constitute 82% of the urban public transport, throng the streets of Kampala in a seemingly chaotic pattern, picking up and dropping off passengers at various stops in the city centre and the various settlements (ITP, 2010). Many of the stops are informal, i.e. they are not officially designated taxi or bus stops but they develop organically according to the demand in a particular area.

To study the movements of the minibus taxis in time and space, we used standard GPS receivers with a spatial accuracy of three meters and a temporal resolution set to 20 seconds when the vehicle's ignition is on and 10 minutes when the ignition is off. The data collected included unique identity, timestamp, longitude, latitude, speed, and direction.

We fitted 20 minibus taxis with GPS receivers that transmitted data to our servers for a period of eight months (Jan 2017 to August 2017). For analysis in this chapter, we used continuous movement data from nine receivers. Data from other receivers were omitted because of substantial discontinuities due to malfunction, vandalism of the receiver, or frequent mechanical problems with the taxi. Preliminary statistical analysis indicated that the nine minibus taxis under study were active for 155 to 235 days, 12 to 23 hours a day, with peak activity occurring between 4:00 and 9:00, 12:00 and 14:00, and 16:00 and 21:00. The mean and standard deviation of days active was 186 and 71 days, respectively, while the mean and standard deviation of hours active were 16.3 and 6 hours, respectively.

To improve performance during analysis, we used Massachusetts Institute of Technology(MIT)'s path simplification Python library (simplification) to reduce the GPS points while maintaining the integrity of the trajectories. Simplification is a robust high-level implementation of the Ramer-Douglas-Peucker algorithm. Figure 4.2a shows a sample minibus taxi trajectory for one day, Figure 4.2b the trajectory simplification process as applied to a small section of the original trajectory, Figure 4.2c the simplified trajectory, and Figure 4.2d the pause-based model that we used to extract Lévy walk steps and turning angles between subsequent steps.

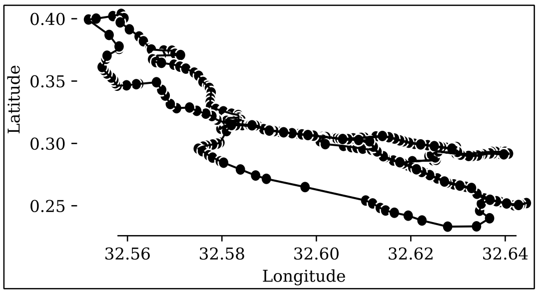
4.3.1.1 The pause-based model

In a pause-based model (illustrated in Figure 4.2d), we define a step as a straight-line movement between two positions P1 and P4 (regarded as pauses) given that the instantaneous velocities at positions P2, P3, and P5 are higher than the threshold velocity. The step length l is the sum of all individual segment distances that make up a step, whereas the turning angle θ is the bearing of the next pause (P6) from the current pause (P4).

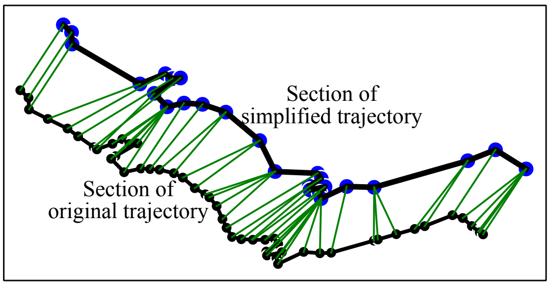
From the GPS traces, we extracted the taxis' steps, step lengths, turning angles, and average velocities during the step. To get these data, we re-sampled the trace data every five minutes and re-computed the relative position in space, cumulative distance and average velocity. Using the re-sampled data, we then extracted steps using a pause-based model. Figure 4.2e shows the spatial distribution of steps and pauses extracted from a minibus taxi trajectory sample in Figure 4.2a.

4.3.2 The Lévy walk as a descriptor for minibus taxi mobility

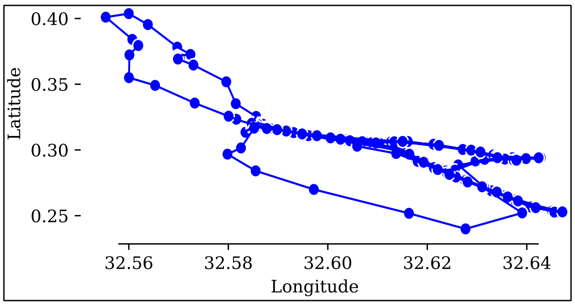
To test whether minibus taxis movements exhibit Lévy walks, we divided the minibus taxi trajectories' data into steps and pauses, as described in the pause-based model. We then fitted the step lengths to a power-law distribution defined by a probability density function: $f(l) \sim l^{-\alpha}$ where l is the step length, and α is the Lévy exponent. We then performed a logarithmic transformation on the data and estimated the Lévy exponent α from a power-law fit for each minibus taxi using the power-law python package (Alstott et al., 2014).



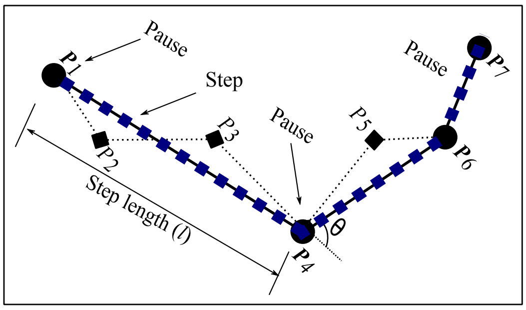
(a) MBT trajectory (UTX-19: 1 day)



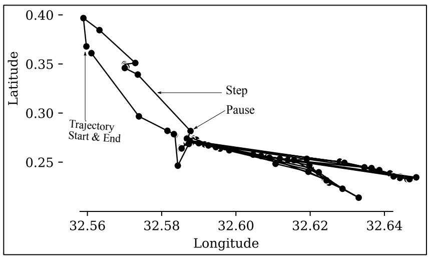
(b) Trajectory simplification process



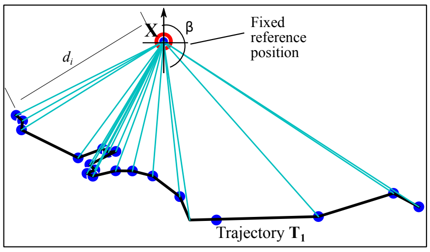
(c) Simplified MBT trajectory



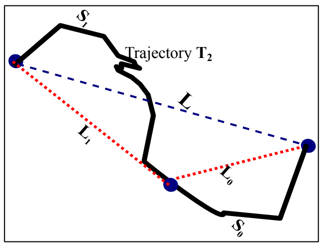
(d) Levy walk extraction: Pause-based model



(e) Spatial distribution of steps



(f) Estimating spatial distance



(g) Estimating tortuosity

Figure 4.2: a) Minibus taxi (MBT) trajectory sample for UTX-19 for one day; b) Trajectory simplification process; c) Simplified minibus taxi trajectory of the trajectory in a; d) Pause-based model to extract steps from a section of the minibus taxi trajectory in a; (e) Spatial distribution of steps and pauses extracted from the simplified trajectory in c; f) Illustration of trajectory spatial distance estimation; and g) Illustration of trajectory tortuosity estimation.

4.3.2.1 Methods to test the minibus taxi Lévy walk behaviour

To test the Lévy walk behaviour in minibus taxi movements, we performed three different tests on the probability distributions of step lengths and step turning angles. First, we examined the step length distribution's mean spread (standard deviation) and skewness to check if it was heavy-tailed. Figure 4.3a shows the general distribution of step lengths for all minibus taxis. Second, we examined the distribution of turning angles between steps to establish whether they were normally distributed. Figure 4.3b shows the general distribution of step turning angles for all minibus taxis and Figures 4.5a(iv), 4.5b(iv), and 4.5c(iv) show the probability distribution of turning angles of selected minibus taxis. Third, we fitted the step lengths' data to a power-law distribution after a logarithmic transformation and estimated the Lévy exponent α to see if it was within the range $1 < \alpha \leq 3$ (the third property of a Lévy walk (Jiang et al., 2009)). Figures 4.5a, 4.5b, and 4.5c show the log-log fit to step lengths and turning angles data. Furthermore, we tested the goodness of the fit using the maximum likelihood estimation method, as suggested by Alstott et al. (2014). We did this by comparing the R (log-likelihood) and p (significance) values from the comparison of the best fit of power-law distribution with other distributions, such as exponential distribution. Figures 4.5a(ii), 4.5b(ii), and 4.5c(ii) compare the goodness of fit of an exponential fit with the power-law fit. Table 4.1b shows the statistical summary of the Lévy walks analysis results.

4.3.3 Comparing trajectories

We analysed the similarities and dissimilarities between minibus taxi trajectories by computing their spatial distances from a common fixed position X. We defined a trajectory as the evolution of a minibus taxi's position in space (on the earth's surface) for 24 hours; space as the surface of the spheroid earth; position (0.314921, 32.578705) as the latitude and longitude coordinates of the fixed position X, which is a central location at the city square in Kampala; and spatial distance as the unit measure of how far (in space) a one day trajectory is from a given reference position. Given a trajectory T_1 (illustrated in Figure 4.2f) for a day D_1, we computed its spatial distance ℓ with respect to an arbitrary fixed reference position $X_{(lon,lat)}$ in space using the equation:

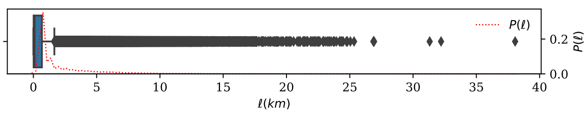
$$\ell = \sum_{i=0}^{n} \sqrt{d_i}\sqrt{\beta_i} \tag{4.2}$$

where d is the Vincenty distance between the arbitrary fixed GPS position X and the i^{th} GPS point on the trajectory T, and β is the bearing angle between two coordinates $A(A_{lat}, A_{lon})$ and $B(B_{lat}, B_{lon})$ on the earth's surface, given by the equation

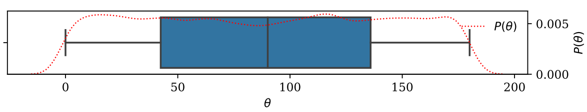
$$\beta = atan2(\gamma, \theta) \tag{4.3}$$

where, $\gamma = cos(B_{lat})sin(|B_{lon} - A_{lon}|)$ and $\theta = cos(A_{lat})sin(B_{lat}) - sin(A_{lat})cos(B_{lat})cos(|B_{lon} - A_{lon}|)$.

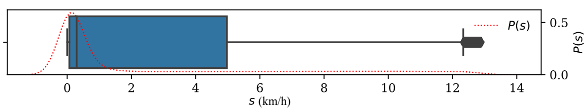
For each minibus taxi, we normalised the values of all trajectories' spatial distances to fall in the range 0 to 1 to simplify the analysis and interpretation of results. Table 4.1a and Figure 4.4b(ii) show the spatial distance distribution for the minibus taxis, sampled per day.



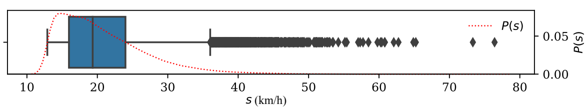
(a) Step lengths l ($\mu = 0.83$, x̃$= 0.1$, $\sigma = 1.94$)



(b) Turn angle θ ($\mu = 89.36$, x̃$=90.08$, $\sigma = 53.33$)



(c) Step speed s (Heads) ($\mu = 2.7$, x̃$=0.3$, $\sigma = 38.9$)



(d) Step speed s (Tails) ($\mu = 20.72$, x̃$=19.34$, $\sigma = 6.21$)

Figure 4.3: Summary distributions for all minibus taxis' step lengths, turning angles and steps speeds (speeds were calculated for distances where heads <= 0.676 km, and tails > 0.676 km).

Table 4.1: Trajectory characteristics and power-law analysis results

(a) Trajectory characteristics (distance ℓ, tortuosity τ) and Lévy step speeds (heads ≤ 0.67, tails>0.67).

Taxi IDs	#Trjs	Trajectory characteristics				Lévy steps speeds (km/h)						
		SD ℓ		Tortuosity τ		Heads			Tails			
		μ	δ	μ	δ	Max	μ	δ	Min	μ	δ	Max
UTX-04	1983	0.39	0.19	0.68	0.17	76.41	9.53	7.65	3.31	22.73	7.89	52.67
UTX-11	4191	0.40	0.10	0.96	0.04	48.54	0.36	1.03	2.01	14.45	8.09	47.45
UTX-12	4177	0.37	0.11	0.96	0.04	38.43	0.29	0.73	2.00	15.52	9.12	73.33
UTX-13	4806	0.67	0.19	0.90	0.09	53.49	10.51	6.95	2.65	21.30	5.51	57.73
UTX-15	1421	0.50	0.23	0.93	0.06	64.88	10.21	7.34	4.98	21.33	7.00	55.18
UTX-16	3483	0.57	0.24	0.93	0.08	65.23	9.71	6.41	2.05	20.32	5.09	49.30
UTX-17	185	0.44	0.40	0.95	0.06	41.51	10.30	6.35	9.42	20.86	5.58	62.76
UTX-18	1885	0.63	0.28	0.92	0.06	53.42	10.71	6.83	3.04	21.61	6.22	51.07
UTX-19	1728	0.74	0.20	0.85	0.10	51.36	9.71	6.67	2.02	21.11	7.52	55.42

(b) Step lengths, turning angles, Lévy exponent α and model goodness of fit results.

Taxi ID	#Steps	Lévy step lengths l (km)				Step turning angles θ^o			Power-law parameters		Goodness of fit: power-law with			
											exponential		log-normal	
		Max	μ	δ	Skew	μ	δ	Skew	α	σ	R	p	R	p
UTX-04	7887	17.99	1.03	1.79	3.84	89.93	54.19	-0.04	1.98	0.016	9.01	0.070	9.76	0.0
UTX-11	45924	24.74	0.47	1.61	6.13	91.93	53.83	-0.04	1.53	0.003	101.11	0.120	17.07	0.0
UTX-12	45947	25.05	0.51	1.84	5.77	89.68	54.38	0.02	1.51	0.003	127.44	0.080	10.54	0.0
UTX-13	18387	23.81	1.27	2.04	3.30	88.93	57.24	-0.01	1.96	0.009	19.93	0.240	14.89	0.0
UTX-15	7541	18.14	1.21	1.76	3.02	87.71	53.12	-0.02	4.18	0.176	-1.87	0.062	2.27	0.02
UTX-16	13008	26.92	1.12	1.70	3.34	88.72	47.84	-0.05	4.75	0.159	1.55	0.122	0.65	0.51
UTX-17	2026	20.36	1.52	2.42	2.80	88.86	47.40	-0.04	1.79	0.021	7.50	0.000	6.83	0.0
UTX-18	9515	32.21	1.22	1.94	3.93	87.79	51.81	-0.01	3.86	0.127	1.08	0.279	1.43	0.15
UTX-19	12780	38.07	1.63	2.92	3.22	85.04	46.19	0.07	7.36	0.521	0.52	0.606	0.18	0.86

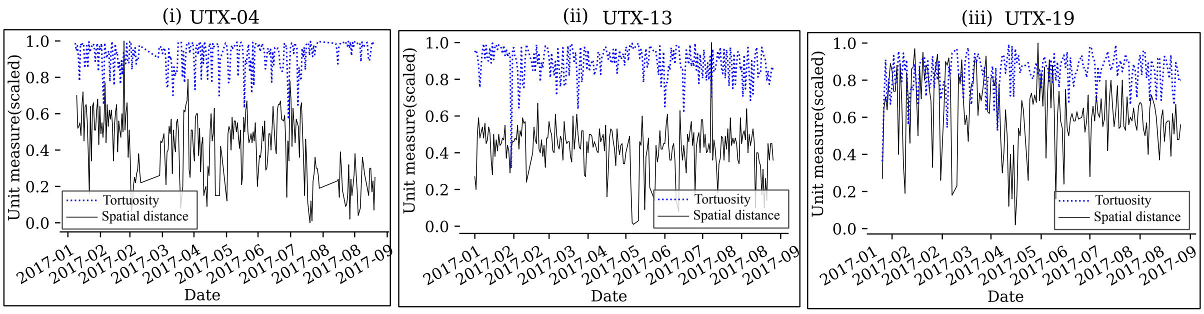
4.3.4 Trajectory tortuosity

To quantify and analyse the shapes of individual trajectories to describe how the minibus taxis wind their way through the spatial reference system, we computed their respective tortuosity values.

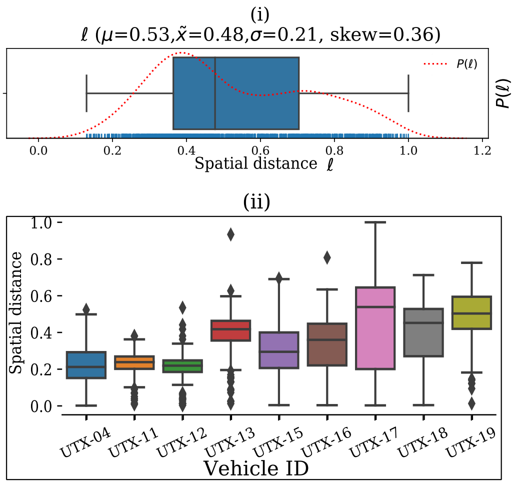
We estimated the tortuosity of trajectories as the ratio of a beeline distance between the start and end of the trajectory L to the length of the travelled trajectory S (Grisan et al., 2003) as illustrated in Figure 4.2g. For N minibus taxi sub-trajectories in a day's trajectory, the tortuosity τ is computed as:

$$\tau = \frac{N-1}{L} \sum_{i=1}^{N} \left(\frac{L_i}{S_i} - 1 \right) \tag{4.4}$$

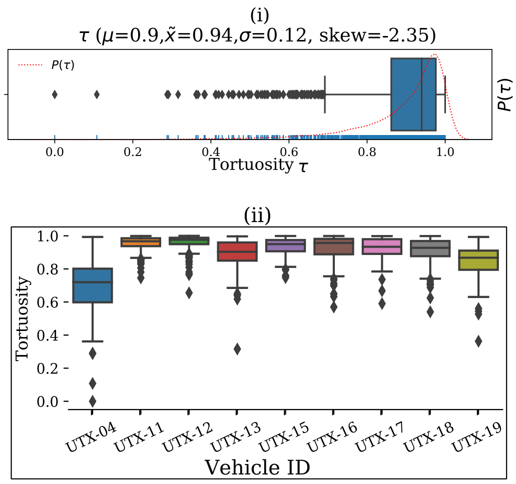
where N is the number of sub-trajectories, L is the beeline distance between the start and end of the day's trajectory, L_i is the beeline distance of the i^{th} sub-trajectory, and S_i is the cumulative length of the travelled sub-trajectory.



(a) Daily variation of normalised spatial distance and tortuosity.

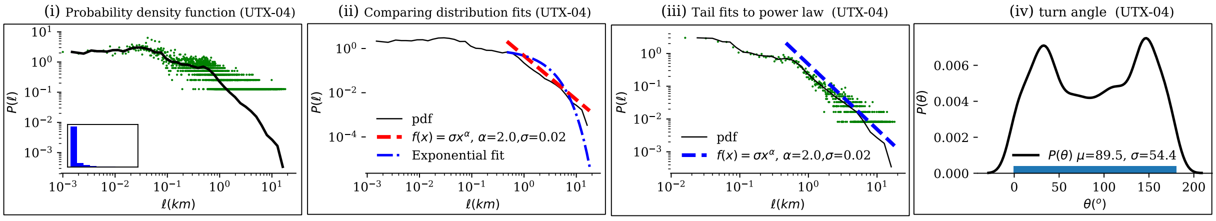


(b) Normalised spatial distance ℓ

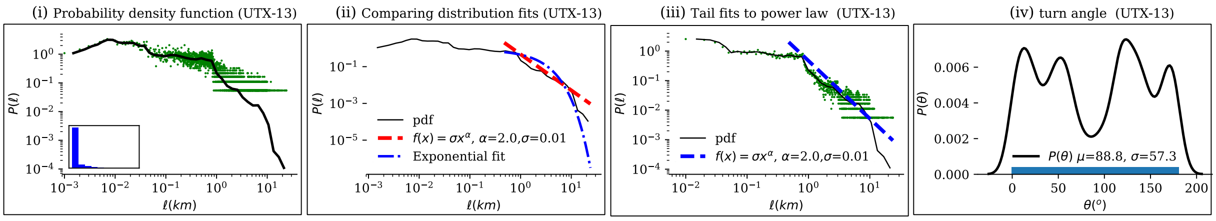


(c) Normalised tortuosity τ

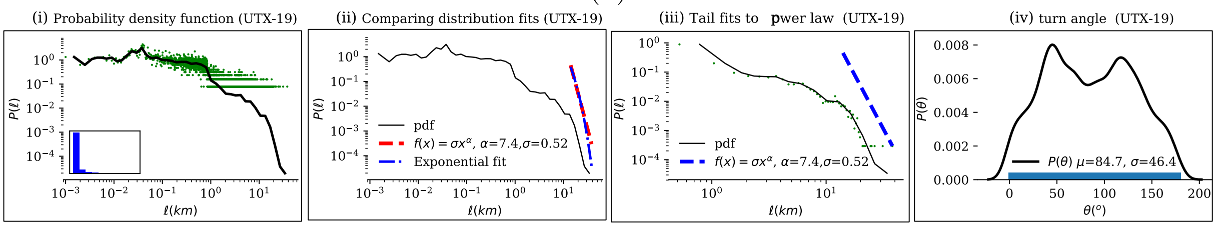
Figure 4.4: (a)Daily variation of spatial distance and tortuosity for selected minibus taxis, i.e., UTX-04, UTX-13 and UTX-19; (b) Aggregate distribution of normalised spatial distance; (c) Aggregate distribution of normalised tortuosity.



(a)



(b)



(c)

Figure 4.5: Power-law analysis of Lévy walk steps.

4.4 Results

4.4.1 Minibus taxi movements and Lévy walk behaviour

From the extracted steps, we found that the probability distribution of step lengths is heavy-tailed, as shown in Figure 4.3a, with a mean μ and a standard deviation σ of 0.83 and 1.9 kilometres, respectively. Its positive skewness of 4.52 shows a significant bulge on the distribution "head" and some rare long walks ("tails") of up to 39 kilometres. This is the first identifying characteristic of Lévy walk behaviour (Viswanathan et al., 2011). We also found the second identifying characteristic: turning angles between steps are normally distributed with a mean of 89.4^o and a standard deviation of 53.3^o, as shown in Figure 4.3b.

Micro-level analysis and fitting of individual minibus taxi walks' data to the power-law function revealed a strong power-law behaviour for minibus taxis UTX-04, UTX-11, UTX-12, UTX-13 and UTX-17, with an estimated Lévy exponent α in the range $1 < \alpha \leq 3$ (see power-law parameters in Table 4.1b). We further confirmed the power-law behaviour in the steps data by comparing the goodness of fit with other distributions and computing the log-likelihood ratio R between the candidate distributions. We also noted the significance value p. Table 4.1b shows the corresponding values of R and p from the goodness-of-fit comparison between power-law and two other distributions (exponential and log-normal). The positive R values and p values greater than 0.05 further confirmed a stronger fit to the power-law than to the exponential and log-normal distributions. Figure 4.5 shows that the step lengths data for UTX-04, UTX-13 (4.5a and 4.5b) fit the power-law better than the data for UTX-19 do (6c). Figures 4.5a(iii), 4.5b(iii), and 4.5c(iii) further illustrate the strength of power-law fit to the tails (>0.676) of the data. Figures 4.5a(iv), 4.5b(iv) and 4.5c(iv) exhibit a Gaussian mixture of turning angles with multiple Gaussian distributions where each peak represents a major hub visited by the minibus taxi, such as a formal taxi rank, and then makes a sharp turn. We also noted that the multi-Gaussian nature of the turning angle distributions is responsible for the generally high standard deviation of 53.3^o observed in Figure 4.3b.

We concluded that five of the nine minibus taxis under study exhibited Lévy walk behaviour. Jiang et al. say that to identify a Lévy walk pattern, all that is needed is to detect power-law behaviour and then estimate the exponent α to see whether it is within the range $1 < \alpha < 3$ (Jiang et al., 2009). For minibus taxis, UTX-04, UTX-11, UTX-12, UTX-13 and UTX-17 the Lévy exponent α for step lengths was in the range $1.51 \leq \alpha \leq 1.98$. The R values, when we compared the power-law function fit with exponential and log-normal model fits were in the range $9.10 \leq R \leq 127.44$. Furthermore, the p values for those five taxis were in the range $0.62 \leq p \leq 0.24$. This indicated a strong fit to the power-law, and hence a significant presence of Lévy walk behaviour in minibus taxi movement trajectories.

4.4.2 How similar are the minibus taxis' trajectories?

The spatial distance ℓ_T of a trajectory T, given by equation 4.2, is the distance between a fixed position X and the trajectory T, as illustrated in Figure 4.2f. We used the observed spatial distances to describe how trajectories from the same minibus taxi differ from each other in space. We took one day as the time interval of each taxi trajectory under discussion. Most of the spatial distances of the minibus taxi trajectories were normally distributed with a mean μ of 0.53 and moderately spread with a standard deviation σ of 0.21. Figures 4.4a(i-iii) shows the daily variations of normalised spatial distances and normalised tortuosity of three selected minibus taxis, i.e., UTX-04, UTX-13, and UTX-19, respectively. Figure 4.4b(i) shows the distribution of normalised spatial distances for all the minibus taxis studied. Figure 4.4b(ii) shows the distributions of distances for each of the nine minibus taxis and Table 4.1a gives a more detailed breakdown.

We observed a moderate spread in the distribution of spatial distances. This strongly suggests that minibus taxis often divert from the most frequented routes, leading to the discovery ("evolution") of new routes. This is also an indicator of growing passenger demand in areas where the new route passes. Practically, if the taxi travelled on the same route all the time, the spatial distances would be less spread.

4.4.3 The significance of the observed trajectory tortuosity

The shape of a trajectory illustrates how a moving object winds or twists its way through a spatial reference system. The similarity of shapes can be expressed qualitatively (topologically), or quantitatively, using parameters such as tortuosity (curviness), and fractal dimension (Ranacher and Tzavella, 2014). We used equation 4.4 to estimate the tortuosity τ of the minibus taxi trajectories. The tortuosity values were normalised to fall in the range 0 to 1 and are summarised in Table 4.1a. Generally, the tortuosity of minibus taxi trajectories is high. This is shown by the general distribution in Figures 4.4c(i), and at the individual taxi level (4.4c(ii)). We argue that this tortuosity distribution is suggestive of minibus taxi drivers' extreme determination to search for passengers to make the trips profitable.

4.5 Discussion

The results from this study show that minibus taxis movements in Kampala – which represent searching for, acquiring, loading passengers, and transporting them to their destinations – tend to follow a Lévy walk pattern similar to movements observed in a wide range of less cognitively complex species (Shlesinger, 2006; Viswanathan et al., 1999), and more recently, in humans (Jiang et al., 2009; Raichlen et al., 2014; Bertrand et al., 2007). The Lévy walk is evident in the many short steps interspersed with rare long steps, and in the Lévy exponent α values (in the range $1 < \alpha \leq 3$) for the greater number of the minibus taxi trajectories (five of the nine sampled taxis). Based on the findings (Viswanathan et al., 1999; Ferreira et al., 2012) of optimal Lévy walk for undepletable, heterogeneous and patchily located resources, we can claim that two taxis (UTX-04 and UTX-13) had adopted near-optimal search strategies, because they had values of $\alpha \approx 2$ (refer to Table 4.1b). The near-ballistic behaviour (α values 1.53 and 1.51) exhibited by taxis UTX-11 and UTX-12 might indicate that the drivers were influenced by previous knowledge of passenger demand (positive memory influence), or it could simply indicate that they often loaded passengers from the taxi ranks. Usually, taxis that load from the taxi ranks take longer to fill up. However, they only load "direct route passengers" who are going to areas closer to the final destination of the taxi, and they charge a fixed fare equivalent to the maximum amount for the passenger going furthest. We can suggest three possible reasons for the near-Brownian ($\alpha > 3$) movement behaviour of minibus taxis UTX-15, UTX-16, UTX-18 and UTX-19. First, the drivers could be new (to the routes, or to taxi driving) and, having not yet figured out a better passenger search strategy, were operating a very inefficient loss-prone strategy. Second, they could be town-service taxis operating in areas with densely distributed informal stops and uniformly distributed short trips demand, leading them to adopt a random search strategy. Third, they might be perennial traffic rule offenders adopting an evasive strategy to avoid encounters with traffic officers on the main routes. This comes at the cost of never being certain of the demand on their "by-pass" routes.

Furthermore, the high tortuosity and moderate spread of spatial distances (in Table 4.1a) suggests that, the drivers' search for passengers extremely energetically and some of them even aggressively. If the trips are not profitable, one strategy drivers use to improve profitability is to stay at a stop (hold back) until the taxi is full or almost full. Another strategy is to explore new routes, leading to route evolution. Though we could not verify how profitable the trips were because we lacked data on minibus taxi occupancy, we did a visual inspection of the geospatial layout of minibus taxis routes (using quantum geographical information system (QGIS) software) from individual minibus taxis. From the geospatial layout of the routes, we confirmed the visible change in the shapes of significant routes over several months, and thus concluded that the taxis' routes evolved. Another possible reason for the route evolution is the urban sprawl mentioned earlier. With the proliferation of informal settlements, and poor planning for the locations of amenities like schools, hospitals and shopping centres, the passenger demand is sparsely distributed among sparsely populated patches around the city, hence the unstable transport supply characteristics visible in route changes.

4.6 Summary

In this chapter we have demonstrated, using a sample of nine minibus taxis in Kampala, that minibus taxi movements in a quasi-demand-responsive paratransit system exhibit features statistically similar to those of a Lévy walk. We argue that even the minibus taxis that showed features outside the Lévy walk parameters will eventually subconsciously adopt the Lévy walk strategy. This, we suspect, is because of memory influence (ability to learn, memorise and respond to passenger demand), and the need to optimise profits. Our research further showed that a significant number of minibus taxi routes evolved (changed topology and shape) with time. This is suggestive of the dynamic demand patterns, and the demand-responsive nature of the minibus taxi paratransit system. Finally, we found that minibus taxi routes were extremely tortuous, indicating a determined, energetic, and even aggressive search for passengers.

We were unable to verify the effectiveness of the Lévy walk strategy in minibus taxis because we lacked data on minibus taxi occupancy. However, based on the Lévy exponent α, we can conclude that, overall, the minibus taxi search strategies revealed by our data are inefficient. This is because only two minibus taxis had a close to optimal Lévy exponent (UTX-04, UTX13, with values of α=1.98 and α=1.96). We can further conclude that the other seven minibus taxis passenger search strategies were not adequate, pointing to a rather inefficient paratransit system.

Consequently, we have characterised the movement patterns of minibus taxis in Kampala's paratransit system and estimated the system efficiency based on the Lévy exponent α. Therefore, we have achieved research objectives 1.2 and 1.3 and answered research questions RQ1 and RQ2.

Part II:

Research Stage II

Chapter 5

Agent-based modelling of urban minibus taxis

Chapter 5 Objectives
This chapter aims to achieve Objective 2.1 of the dissertation.
• ⇒**Research objective 2.1**

> Design and describe an agent-based model (ABM) of minibus taxis and passengers in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

To achieve the objective of this chapter, passenger travel by minibus taxi (MBT) in Kampala's paratransit system was modelled using agent-based modelling as illustrated in Figure 5.1. This agent-based model (ABM) models travel by minibus taxi in Kampala as a collection of autonomous decision-making entities called agents. Each agent (e.g., passenger or driver in control of a minibus taxi) individually assesses its situation and makes travel decisions based on a set of rules. The agents repetitively interact (with self, with other agents and with its surroundings) in a common environment, executing various actions (such as boarding a taxi and searching for passengers). Agents in the ABM are capable of learning and adapting their behaviour to achieve desired goals, thereby, allowing new and sometimes unanticipated behaviour to emerge. The overall goal is to provide a natural description of the minibus taxi transport dynamics in a quasi-demand-responsive paratransit system similar to the one in Kampala, Uganda.

5.1 ABM model overview

This section describes the agent-based model developed to study the dynamics of minibus taxi transport in Kampala. The description partially follows the ODD+D protocol (Müller et al., 2013), an extension of the ODD protocol Grimm et al. (2006). The ODD (Overview, Design Concepts and Details) protocol provides a standardised way of describing agent-based models, and the ODD+D extension to the protocol offers an elaborate structure for describing the decision-making process of agents in agent-based models.

The ABM (illustrated in Figure 5.1) was developed to represent the minibus taxi transport system in Kampala, Uganda and is used to simulate the movement dynamics of passengers and minibus taxis (or taxi drivers) in a quasi-demand-responsive paratransit system in a developing city setting. The model and the subsequent simulation (described in Chapter 6) are used to study micro-level autonomous agents' interactions and emergent behaviour. Subsequently the model and simulation results will enable the author to answer research questions two and three, and to achieve objectives two and three, respectively, as stated in Sections 1.3.1 and 1.3.3.

The agent-based model consists of three major components, namely, the agents, the broker, and the environment. The agents are autonomous decision-making entities (passengers and minibus taxis). The broker facilitates agent-to-agent, and agent-to-environment data exchange. The environment is the simulated space where agents reside and interact autonomously, occasionally exchanging data through the broker. Agents are intelligent: they can perceive their environment through sensors and act upon that environment based on the predefined rules. Figure 5.1 shows the framework for the developed agent-based model for minibus taxi transportation in Kampala's paratransit system.

5.1.1 Entities and state variables

The model is composed of two active entities (passengers and minibus taxis) and two passive entities (attractors and road network). Each entity is located at its own layer of abstraction as illustrated in Figure 5.1. Figure 5.2 shows the UML diagram of selected model classes. Below is the description of the layers and the associated entities.

5.1.1.1 *Layer 1: Passengers (active agents)*

The passengers layer represents the minibus taxi commuters in Kampala's paratransit system. The passenger agents on this layer plan their journeys in time and space and maintain a journey diary that is updated as the agent interacts with other agents and the

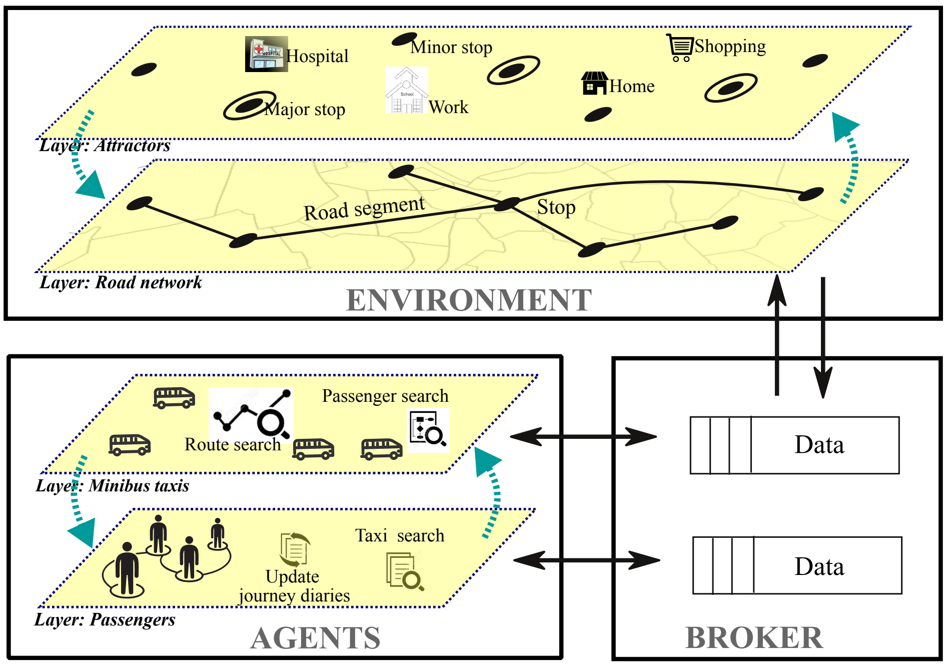


Figure 5.1: Conceptual agent-based model of minibus taxi transport in Kampala's paratransit system.

surrounding in the same environment. Figures 5.3a, 5.3b and 5.3c illustrate the passenger agent journey plan, the journey diary, and the nomenclature of the journey, respectively. Agents in the passenger layer have states (i.e., passive, waiting, onboard and arrived), and behaviour (i.e., journey planning, searching for a taxi, executing journey, and updating the journey diary). During journey execution, the passenger agents interact with other agents (of similar and different types) within the same environment. Table 5.1a describes the passenger agent states and the detailed passenger class definition is given in Appendix A.1 (Algorithm 4).

5.1.1.2 *Layer 2: Minibus taxis (active agents)*

The minibus taxi layer represents minibus taxis operating in Kampala's paratransit system. The minibus taxi agents move autonomously in the environment, occasionally changing states (passive, stopped, routing, loading, and moving) and exhibit certain behaviour (passenger touting, route searching and dynamic route abandonment). Table 5.1b describes the minibus taxi agent states. The detailed minibus taxi class definition is given in Appendix A.1 (Algorithm 5).

5.1.1.3 *Layer 3: Road network (passive objects)*

The ABM environment contains a layer of the roads network, which consists of interconnected road segments that provide a path constraint for the minibus taxi agents to move. The road network is important to this agent-based model. It represents the street network where minibus taxis operate to fulfil trips from origin to destination.

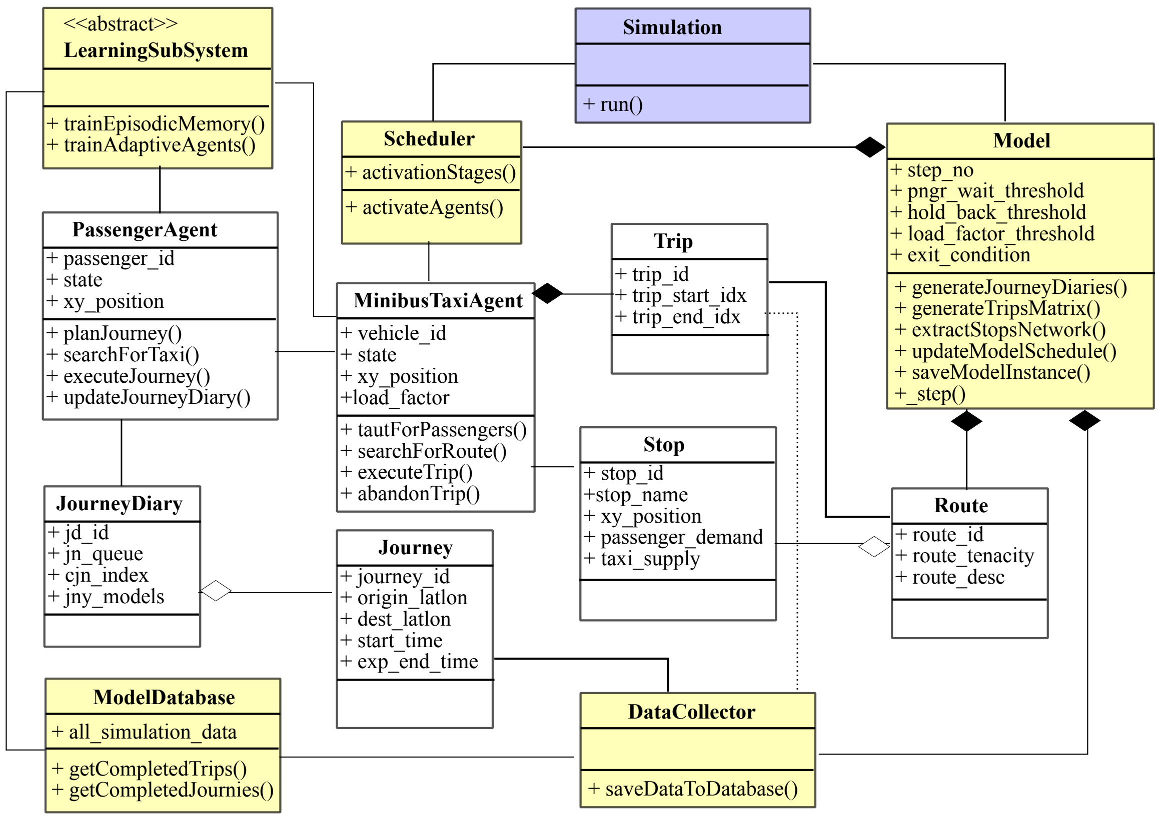


Figure 5.2: UML class diagram showing the most relevant classes of the model.

5.1.1.4 *Layer 4: Attractors (passive objects)*

Attractors in this ABM are locations or sets of locations in the environment space that act as points of interest to the passenger agents or minibus taxi agents. They "attract" purposeful commuter journeys and thus minibus taxi trips to their locations. The attractors layer consists of locations of passive objects in the environment such as schools, hospitals, shopping centres, and workplaces that act as sources and destinations of purposeful journeys. Another class of attractors in this layer is the stops. The stops represent locations on the road network – usually in proximity to the attractors where minibus taxis pick up and drop off passengers.

5.1.2 Scales

In this model, one time-step represents five minutes. Seventy days of minibus taxi and passenger interactions were simulated.

5.1.3 Process overview and scheduling

The model initialises when passengers make daily travel plans consisting of one or many purposeful journeys, as illustrated in Figure 5.3a. The journeys are queued and stored in a journey diary, as shown in 5.3b. At each time step, the agents examine their journey diaries to find if there is a journey to execute depending on the time of the day. The candidate journey is then scheduled for execution, and the journey diary updated accordingly.

The minibus taxis are scheduled into the model based on a Gaussian submodel (described in Section 5.3.3.1). Every minibus taxi agent searches for the best route to a predetermined destination zone, and then instantiates a trip to the destination through the selected route. En route, the minibus taxi searches for passengers based on a profit maximisation utility function. They occasionally hold back at random stops in anticipation of passengers. If the trip is persistently not profitable, it is abandoned. A new trip– anticipated to be profitable– is initiated and executed by the minibus taxi.

At the end of every passenger journey and minibus taxi trip, data relevant to how effective the journey or trip was executed is stored and used to train agent-specific submodels every five days for improved future decision making. Figure 5.4a illustrates the conceptual framework of the model, while Figure 5.4b shows snapshots of the model environment at different time steps. Environment E_1 (at time step t) represents a minibus taxi executing a trip on a pre-selected route. Along the route, there are three active passengers. Two of the passengers whose journeys can be fulfilled by the minibus taxi are picked up as shown in environment E_2 (time step $t + dt$). In environment E_3 (time step $t + n$), one of the passengers arrives at the journey destination. The passenger who was not picked up drops the journey after the threshold waiting time is exceeded.

A scheduler (shown in Figure 5.4) manages the time evolution of the model and the orderly execution of all actions and activities of the system. For this ABM, a *Staged Activation* type of scheduler was used. In this scheduler, the simulation goes through four stages. The first is *sensing* to get data about surrounding agents and the environment. The second is *cognitive action*, where agents select appropriate action to take, often based on a utility maximisation method discussed in Section 5.2.2. The third is the *physical action*, where the actual physical action takes place such as moving to a new location, boarding a taxi, or abandoning an active trip. The fourth is the *update* stage, where the agents and the environment are updated with the new state.

(a) Passenger agent states

State	***State description***
Passive (P)	Has a pending trip to make or complete
Waiting (W)	Waiting at a stop
On board (OB)	On board a minibus taxi
Arrived (A)	Arrived

(b) Minibus taxi agent states

State	***State description***
Passive (P)	Available but not active
Stopped (S)	Stopped to pick up or drop off passengers
Routing (R)	Searching for a route to take
Loading (L)	Loading passengers
Moving (M)	Moving to the next stop

Table 5.1: Passenger minibus taxi and agent states

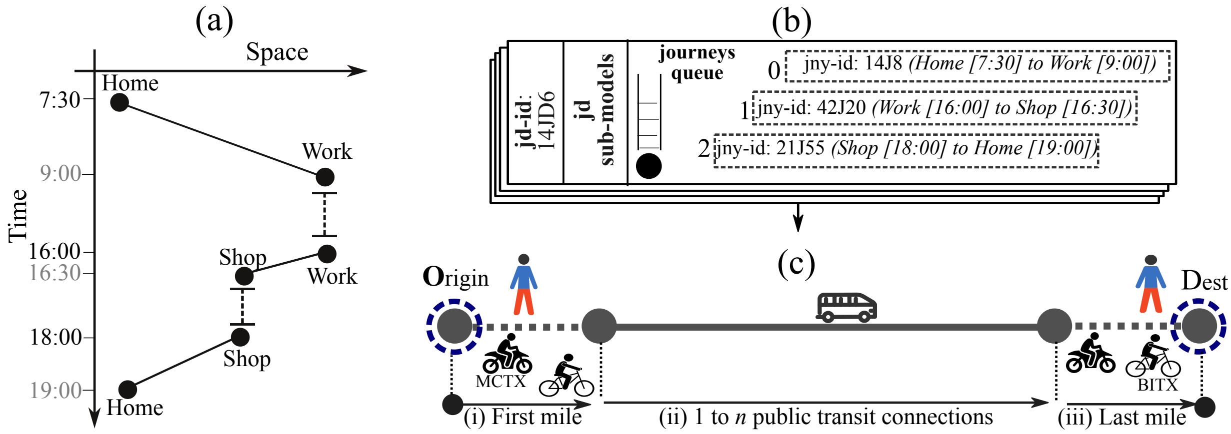
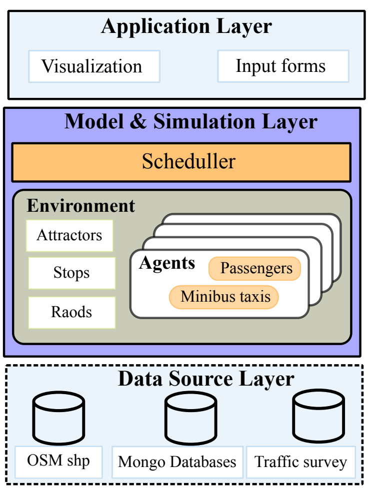
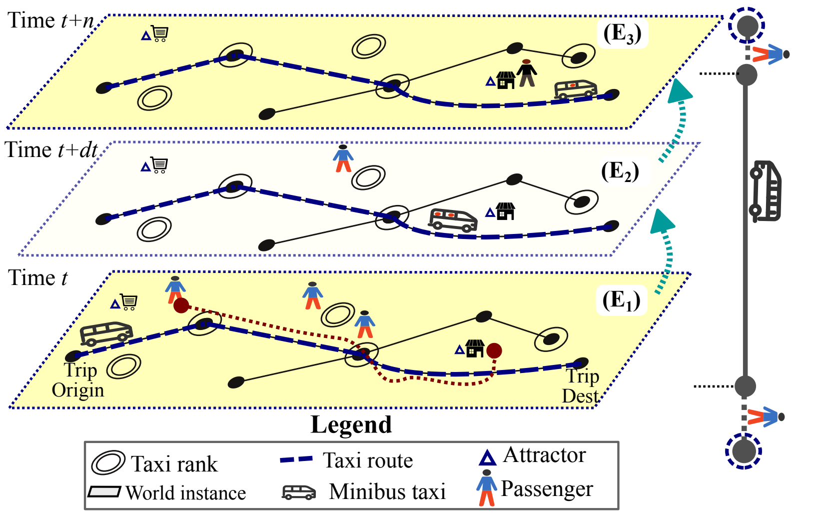


Figure 5.3: (a) Passenger agent journey plan in time and space representing three purposeful, planned journeys – Home-to-Work, Work-to-Shop and Shop-to-Home – scheduled for execution on the same day at 7:30, 16:00 and 18:00, respectively; (b) Journey diary showing the journey queue containing the three journeys; and (c) Nomenclature of a single passenger journey from origin to destination.



(a) ABM framework



(b) Agent-based model simulation overview

Figure 5.4: ABM conceptual framework and model simulation environment snapshots.

5.2 Design concepts

5.2.1 Theoretical background

The conceptual model of agents and interactions was created and refined based on the participatory observation of minibus taxi travel, unstructured interviews with minibus taxi drivers, passengers, managers of taxi associations and the director of Kampala City Council Authority.

The decision model of the agents is based on random utility theory (Cascetta, 2009), which is the basis of several models and theories of decision-making. According to random utility theory, an intelligent agent (such as a person) selecting from several alternatives, chooses one with the highest utility, where the utility function is defined as $U = V + \varepsilon$, with U being the total unobservable utility, V the deterministic observable component of the agent behaviour, and ε a random component representing the non-measurable factors of an individual's decision.

The random utility theory is based on the hypothesis that every agent is a rational decision-maker, maximising utility relative to its choices (Cascetta, 2009). Specifically, the following assumptions are made.

a) When making a choice, an agent (decision maker) i, considers m_i mutually exclusive alternatives from its choice set I^i. The choice set may be different for different agents;

b) Agent i assigns to each alternative j from its choice set a perceived utility, or "attractiveness" U^i_j and selects an alternative maximising the utility;

c) The utility assigned to each choice alternative depends on several attributes (measurable characteristics) of the alternative itself and of the agent, $U^i_j = U^i(X^i_j)$, where X^i_j is the vector of the attributes relative to alternative j and to agent i;

d) The utility assigned by the agent i to alternative j is not known with certainty and therefore represented by a random variable ε^i_j.

Thus, based on the above assumptions, it is usually not possible to predict with certainty the alternative that a generic agent will select. However, we can express the probability of choosing alternative j conditional on its choice set I^i, as the probability that the perceived utility of alternative j is greater than that of all the other available alternatives.

$$p^i[j/I^i] = Pr[U^i_j > U^i_k \qquad \forall k \neq j, k \in I^i] \tag{5.1}$$

The perceived utility U^i_j can be expressed by the sum of systematic utility value V^i_j, which represents the expected value of the utilities perceived by all agents having the same choice context as agent i and a random residual ε^i_j which is the deviation from the utility perceived by agent i.

$$U^i_j = V^i_j + \varepsilon^i_j \qquad \forall j \in I^i \tag{5.2}$$

5.2.2 Individual decision making

Agents of different types pursue different objectives as discussed in Sections 5.2.2.1 and 5.2.2.2. In general, passenger and minibus taxi agents' behaviour are controlled by a series of submodels described in Sections 5.3.3.2 and 5.3.3.3, respectively. The agents' decision model is based on the assumption that they have only partial information about their surroundings, hence they are boundedly rational (Simon, 1957). The agents use a form of inductive reasoning (Deadman et al., 2000) and rely on a general random utility model combined with an additive form of the Cobb-Douglas function that utilises both scores and weights to choose the appropriate behaviour rationally. The general Equation 5.3 guides the agents' behaviour.

$$Utility_i = \left(\sum_{k=1}^{n} score_{ki} \times weight_k \right) + \varepsilon_i \tag{5.3}$$

where, i is the agent, k is the dimension, and ε is the random noise (random variable, $\mu = 0$, $\alpha = 0.05$) to represent bounded rationality. Each choice context is (also known as "choice dimension") is defined by available alternatives, evaluation factors and decision procedures.

The dimensions selected for this model are not universal; they were selected depending on the type of agent and the decision to be made by the agent. For instance, if a passenger agent wants to determine the stop where to wait for a minibus taxi, the dimensions for the model are three (i.e., the distance to the candidate stop, the availability of routes through the candidate stop, and the likelihood of getting picked up).

The scores for each dimension were normalised to a scale of 0 to 1, where 1 was the most preferred score in the given dimension. They thus had meaning only relative to each other. The values of the weights were determined through iteratively testing and varying model rules. For a given alternative, the weights of all dimensions add up to 1.

To understand the agents' decision rules, consider a situation in which three alternatives–x, y and z–vary along four dimensions D_1, D_2, D_3 and D_4. Their scores and weights along these dimensions are given by the payoff matrix in Table 5.2. The utility is computed and the alternative with the maximum utility is selected. Tables 5.4 and 5.5 describe the dimensions, scores, and weights used by agents to determine the utility during decision making.

5.2.2.1 Decision making: Passengers

The passenger agents' objective is to complete scheduled journeys as optimally as possible individually. Passenger agent decisions are made based on three submodels. First is the initial stop model (ISM) that determines the best stop within a threshold radius to wait for a minibus taxi. Second is the boarding choice model (BCM) that decides which minibus taxi to board. The BCM objectives are the following: to minimise the number of connections necessary to complete

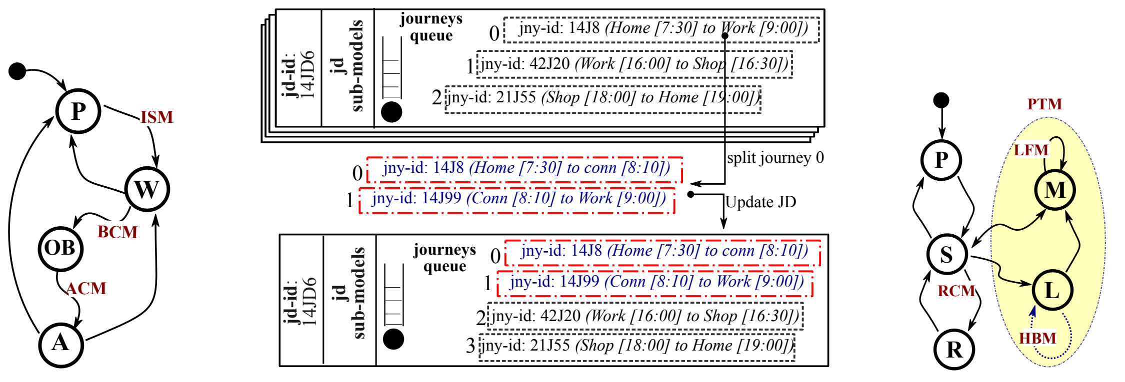
Table 5.2: Illustration of utility payoff determination by an agent given three alternatives, each with four dimensions. Note: (i) $\sum_{d=1}^{4} w_{ad} = 1, for\ a \in \{x, y, z\}$; (ii) $s_{ad} \in (0, 1]$ for $a \in \{x, y, z\}$ and $d \in \{1, 2, 3, 4\}$; (iii) ε is random variable with $\mu = 0$; and $\alpha = 0.05$.

		Dimensions				Noise (ε)	Utility (U)
		D_1	D_2	D_3	D_4		
Alternatives	x	$s_{x1}w_{x1}$	$s_{x2}w_{x2}$	$s_{x3}w_{x3}$	$s_{x4}w_{x4}$	ε_x	$(\sum_{d=1}^{4} s_{xd}w_{xd}) + \varepsilon_x$
	y	$s_{y1}w_{y1}$	$s_{y2}w_{y2}$	$s_{y3}w_{y3}$	$s_{y4}w_{y4}$	ε_y	$(\sum_{d=1}^{4} s_{yd}w_{yd}) + \varepsilon_y$
	z	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	ε_z	$(\sum_{d=1}^{4} s_{zd}w_{zd}) + \varepsilon_z$

a single journey, to minimise the travel distance, and to minimise the last leg distance. In case the available taxi cannot fulfil the entire journey but can fulfil at least half of the journey, the journey is split into sub-journeys (legs or connections), and the journey queue of the journey diary is updated as illustrated in Figure 5.5b. The third is the arrival choice model (ACM) that determines if the “last leg of commute” has been reached, or else the passenger agent waits at a stop for a connecting minibus taxi trip. Therefore, to achieve their objectives, they continuously learn from experience and update their short-term memory referred to as episodic memory in this thesis. Figure 5.5a shows a state transition diagram for passenger agents and the various submodels executed during the transitions. Section 5.3.3.2 provides a detailed discussion of select submodels and the associated dimensions (Table 5.4) that define passenger agents’ behaviour in the model.

5.2.2.2 Decision making: Minibus taxis

The minibus taxi agents’ objective is to maximise trip profitability by filling up the minibus taxi – thus maximising the occupancy – and completing as many trips as possible. To achieve their objectives, minibus taxi agents execute two submodels, the route choice model (RCM) and the passenger touting model (PTM). The RCM determines the route a minibus taxi takes for the current active trip. The PTM implements three strategies used by minibus taxi agents to search for passengers effectively. The PTM implements three passenger search strategies (discovered from our previous studies discussed in Chapter 3). First is the Lévy walk (LW) search behaviour where the distance between consecutive stops is drawn from a Lévy probability distribution. Second is random back off (holding back), referred to in Kampala as “*okukyebakamu*”, where the minibus taxi randomly holds back at a stop waiting for passengers. The third is strategic demand estimation, referred to as “*okubala gap*”, where a minibus taxi with occupancy less than the threshold scans the demand on the route and moves ahead in anticipation for passengers waiting at the stops ahead. To achieve their objectives, they continuously learn from experience and update their episodic memory accordingly. Section 5.3.3.3 provides a detailed discussion of select submodels and the associated dimensions (Table 5.5) that define minibus taxi agents’ behaviour in the model.



(a) (b) (c)

Figure 5.5: Illustration of agents’ state transitions and sub-models. (a) Passenger state diagram, (b) Passenger journey splitting during ISM, (c) Minibus taxi state diagram.

Note: P=Passive, W=Waiting, OB=On board, A=Arrived, S=Stopped, R=Routing, L=Loading, M=Moving, H=Holding back, ISM= Initial stop model, BCM=Boarding choice model, ACM=Arrival choice model, RCM=Route choice model, LFM=Lévy flight model, HBM=Hold-back model, PTM=Passenger touting model.

5.2.3 Learning

In this model, there are two types of learning for passengers and minibus taxi agents. The first is learning from experience at an individual level in which every agent implements a Gaussian mixture learning algorithm that occasionally updates its episodic memory. The second is collective learning where data about passenger demand and transport supply at stops, together with historical data about completed journeys and trips, is modelled and used as a basis for decision making at the stops.

5.2.4 Individual sensing

Every passenger agent in passive or waiting states knows its current location, and the locations of stops within an awareness radius from its position in the environment. Under some conditions, the passenger agent can also read information about a stop such as demand (the number of passengers waiting at a stop); and supply (the number of minibus taxis loading at a stop). Furthermore, the passenger agent can read the routes and occupancy status (load factor) of minibus taxis waiting at a stop.

Before starting a trip, every minibus taxi agent knows about the demand, supply, and candidate routes on stops within a threshold radius. After starting a trip, the minibus taxi knows about the passengers on board, and the demand and supply on stops within a one kilometre distance on the same route.

5.2.5 Stochasticity

During journey planning, once the number of journeys originating from a zone at a given time is determined, the originating positions are randomly distributed in the same zone.

5.2.6 Observation

We used several approaches to test, analyse, evaluate, and finally validate the model. They included setting up and running a controlled simulation experiment with agents semi-autonomously making individual decisions and often learning from individual experience (see Chapter 6 Section 6.4). The control experiment results showed statistically close distributions to the distribution results obtained from the field study described in Chapter 3. The model thus closely represents the transportation dynamics of minibus taxis in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

5.3 Implementation details

The model was implemented in Python. A simplified UML class diagram of the model is shown in Figure 5.2 and the UML activity diagram is shown in Figure 6.1. A model simulation environment based on MESA architecture (Masad and Kazil, 2015) was set up. The ABM structure, components, agents and their associated rules and behaviour were implemented using Python's object-oriented paradigm. The model was tested and validated based on face validity and mainly involved the interpretation of graphs (Ormerod and Rosewell, 2009; Institute of Medicine, 2015).

5.3.1 Input Data

The model used inputs from external data. They included:

a) A custom General Transit Feed Specification (GTFS) file containing Kampala's minibus taxi stops and taxi ranks.

b) Open Street Map (OSM) shapefile containing primary, trunk, secondary and tertiary roads, as well as division and parish administration levels.

c) The 2014 population census data for Kampala from the Uganda Bureau of Statistics (UBOS) (UBOS, 2014).

d) Traffic count data from the transport improvement study in Kampala conducted by the Japan International Cooperation Agency (JICA) in 2010 (JICA, 2010).

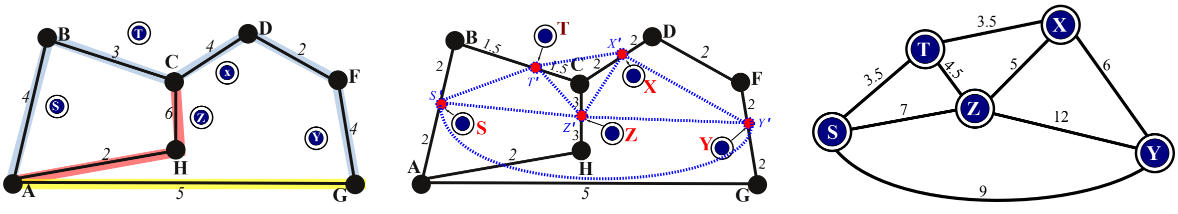
5.3.2 Initialisation

The initialisation of the model involves the acquisition of user-defined model parameters, input datasets, and creation of environment entities such as stops, road network, as well as generating passenger journey diaries. During this stage, the environment is loaded, and the model generates a network of stops whose internal structure is a connected undirected Networkx graph (described in Section 5.3.2.1 and illustrated in Figure 5.6c). From the Networkx graph, a Maximum Spanning Tree (MST) is computed by negating the weights of each edge and applying Kruskal's algorithm. From the MST, an origin-destination (OD) matrix of all graph nodes is generated. From this point on, the OD matrix is used as a basis for looking up and ranking minibus taxi routes.

Passenger journeys between zones – also known as parishes – are generated based on Kampala's 2010 population census data, traffic count data from a study conducted by JICA (2010), and the universal gravity model as discussed in Section 5.3.2.2. The temporal distribution of journeys throughout the twenty-four hours of the day is based on a Gaussian model trained and fitted on hourly minibus taxi trips count in and out of the central division of Kampala. Section 5.3.3.1 discusses the Gaussian model fit to temporal minibus taxi trips count data.

5.3.2.1 *Stops network extraction*

Consider a road network as a connected graph $G = (V, E)$ consisting of a set of *vertices* (also called *nodes*) V and a set of undirected edges $E \subseteq \{\{u, v\} | u, v \in V \wedge u \neq v\}$. Without loss of generality we will assume that elements of V are labelled by letters A, B, C, D, F, G, H as



(a) (b) (c)

Figure 5.6: Stops network extraction; (a) Road network and corresponding stops, (b) Alignment (snapping) of stops to nearby road segments, (c) Internal representation of the extracted stops network.

shown in Figure 5.6a. The road network consisting of roads, road segments and edges that can be uniquely identified.

Consider a road network shown in Figure 5.6a composed of three roads

$$RD_1 : E\{ab, bc, cd, df, fg\},\ RD_2 : E\{ch, ha\},\ RD_3 : E\{ga\} \text{ and four segments}$$
$$S_{1RD1} : E\{ab, bc, cd\},\ S_{2RD1} : E\{df, fg\},\ S_{1RD2} : E\{ch, ha\},\ S_{1RD3} : E\{ga\}$$

Also consider the minibus taxi stops S, T, X, Y and Z located along different road segments. We align the stops to the corresponding closest road segments and compute the stops' relative positions on the segment edges as described in Figure 5.6b. The segments are then transformed to include the stops' relative positions as new nodes, e.g.,

$$S_{1RD1} \rightarrow S'_{1RD1} : E\{as, sb, bt, tc, cx, xd\},$$
$$S_{1RD2} \rightarrow S'_{1RD2} : E\{df, fy, yg\},$$
$$S_{1RD2} \rightarrow S'_{1RD2} : E\{cz, zh, ha\}$$

For purposes of this illustration, after alignment we assumed the stops are located in the centre of the respective edges. We then condensed the road network into a smaller network of stops by aggregating edges between the stops using distance as a metric, i.e.,

$$E\{sb, bt\} \rightarrow E\{st\};\ E\{tc, cx\} \rightarrow E\{tx\};\ E\{sb, bt\} \rightarrow E\{st\};\ E\{xd, df, fy\} \rightarrow E\{xy\};$$
$$E\{yh, fa, ag, gz\} \rightarrow E\{yz\};\ E\{zc, cx\} \rightarrow E\{zx\};\ E\{zc, ct\} \rightarrow E\{zt\}$$

Figure 5.6b shows how the network was condensed to a network of stops shown in Figure 5.6c. The network of stops illustrated in Figure 5.6c was used as the reference network during the model simulation. However, the original network was preserved to give us the ability to alter road characteristics in the future.

5.3.2.2 Generating minibus taxi trips and passenger journeys

To generate minibus taxi trips, consider a discrete set of spatial locations (Zones) $Z=\{z_j, ..., z_J\}$. The number of minibus taxi trips T generated from zone z_i to zone z_j per day are T_{ij}, and the number of trips attracted to zone z_i from zone z_j per day are T_{ji}. Then the total trips generated from zone z_i to all other zones is $\sum_i^I T_{ij} D_j$ and the total trips attracted from all other zones to zone z_i in $\sum_j^J T_{ij} O_i$ as described in the trips matrix structure in Table 5.3a.

Minibus taxi traffic count data from a study on Kampala's road network improvement by JICA (2010), together with Kampala's 2014 parish-level population census data from UBOS (2014), were used to fit a universal Gravity model defined by Equation 5.4 and estimating the fitting parameter γ for minibus taxi travel in Kampala. Table 5.3b shows the trips' matrix (trips for all motor vehicles) generated from application of the gravity model. Equation 5.5 was then used to generate the origin-destination (OD) minibus taxi passenger journeys matrix shown in Table 5.3c.

$$T_{ij} = \frac{m_i n_j}{r_{ij}^{\gamma}} \tag{5.4}$$

where,

T_{ij} is the number of minibus taxi trips from zone i to zone j

m_i is the source zone population size

n_j is the destination zone population size

r_{ij} is the Euclidean distance between zones i and j

γ is the fitting parameter

The data used from the traffic survey included trips from all vehicle types. We however, were interested in minibus taxis trips. Therefore, we calculated commuter journeys based on the percentage minibus taxi mode share (27.7%) relative to other vehicles, and the mean minibus taxi occupancy (10.38 passengers for selected roads observed during three time segments, i.e., morning, midday, and evening, representing 7:00 to 8:00, 12:00 to 13:00 and 18:00 to 19:00, respectively) recorded during the survey as defined in Equation 5.5. Table 5.3c shows the aggregate division-level synthetic commuter journeys for the five divisions of Kampala computed according to Equation 5.5.

$$J_{ab} = T_{ab} \times \mathcal{O} \times \omega \tag{5.5}$$

Table 5.3: Illustration of (a) OD trips matrix structure; (b) Trips matrix achieved from applying the gravity model to trips count data from JICA (2010); OD matrix of minibus taxi passenger journeys.

(a) Minibus taxi trips' matrix structure.

		Trips attraction zone							
		1	2	3	...	j	...	J	$\sum_j T_{ij}$
Trips generation zone	1	T_{11}	T_{12}	T_{13}	...	T_{1j}	...	T_{1J}	O_1
	2	T_{21}	T_{22}	T_{23}	...	T_{2j}	...	T_{2J}	O_2
	:	:	:	:		:		:	:
	i	T_{i1}	T_{i2}	T_{i3}	...	T_{ij}	...	T_{iJ}	O_i
	:	:	:	:		:		:	:
	I	T_{I1}	T_{I2}	T_{I3}	...	T_{Ij}	...	T_{IJ}	O_I
	$\sum_i T_{ij}$	D_1	D_2	D_3	...	D_j	...	D_J	$\sum_i \sum_j T_{ij}=T$

(b) Trips matrix generated by the general gravity model Equation 5.4.

		Destination					
		Central	Kawempe	Makindye	Nakawa	Rubaga	Total
Orign	Central	69,229	16,150	30,304	12,363	22,516	150,562
	Kawempe	42,672	42,555	3,687	4,784	21,787	115,485
	Makindye	42,411	4,644	58,512	1,900	15,681	123,148
	Nakawa	21,688	3,538	2,427	71,104	2,385	101,142
	Rubaga	41,955	11,143	13,505	2,390	33,267	102,260
	Total	217,955	78,030	108,435	92,541	95,636	

(c) Passenger journeys OD matrix generated by Equation 5.5.

		Destination					
		Central	Kawempe	Makindye	Nakawa	Rubaga	Total
Orign	Central	199,051	46,435	87,132	35,547	64,739	432,905
	Kawempe	122,693	122,357	10,601	13,755	62,643	332,049
	Makindye	121,943	13,353	168,237	5,463	45,087	354,083
	Nakawa	62,359	10,173	6,978	204,442	6,857	290,810
	Rubaga	120,632	32,039	38,830	6,872	95,651	294,024
	Total	626,677	224,357	311,779	266,079	274,978	

where,

$J_{ab} =$ Passenger journeys from division a to b

$T_{ab} =$ Minibus taxi trips from division a to b

$\omega =$ Percentage share of minibus taxis relative to other transport modes

$\mathcal{O} =$ Average minibus taxi occupancy.

Note: $\omega = 0.277$, $\mathcal{O} = 10.38$ passengers

5.3.3 Submodels

5.3.3.1 Gaussian model for temporal distribution of trips

The trips generated in Section 5.3.2.2 represent aggregated passenger trips throughout the day. We used a separate detailed hourly minibus taxi trips survey carried out by JICA in Kampala on the 14^{th} and 15^{th} of January 2010 to train a Gaussian model and fit it to the sparse hourly minibus taxi trips count data collected between 7:00 and 18:00. Figure 5.7a shows the trips distribution inbound to and outbound from the central division whereas Figure 5.7b shows the daily temporal profile prior to training the Gaussian model.

To train the Gaussian model, we divided the data into a ratio of 7:3 with respect to training and test data sets. We used GPy (a Gaussian Process (GP) framework from the Sheffield machine learning group) (GPy, 2012) with a radial basis function (RBF) kernel, and a bias function defined in Equations 5.6 and 5.7, respectively. Figure 5.7c shows the results of the Gaussian model fit with a double belly for inbound trips, and triple belly for the outbound trips. Figure 5.7d shows the predication test results.

This submodel is used for estimating the temporal distribution minibus taxi trips and passenger journeys.

$$K(\mathbf{x}, \mathbf{x}') = \left(-\frac{\| \mathbf{x} - \mathbf{x}' \|^2}{2\sigma^2}\right) \tag{5.6}$$

where, $\mathbf{x}$ and $\mathbf{x}'$ are two samples, represented as feature vectors in some input space.

$$K_{ij} = \kappa(X_i, X_j) \tag{5.7}$$

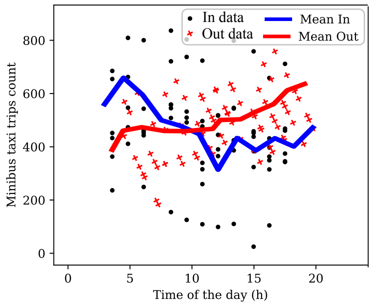
where, X is the first set of inputs to the kernel, and X_2 (optional) is the second set of arguments to the kernel.

5.3.3.2 Passenger agent submodels: ISM, BCM and ACM

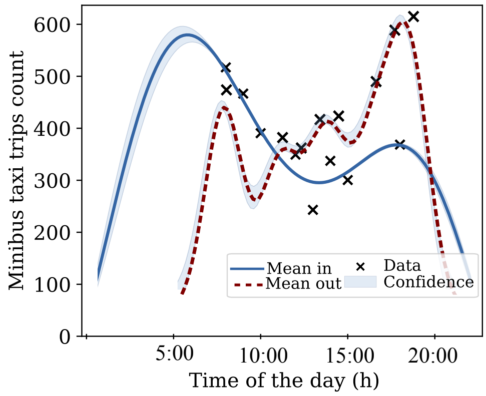
This section describes the sub-models used to manage passenger agent behaviour. The state diagram in Figure 5.5a shows the abstract description passenger agent behaviour as a series of transitions between discrete states described in Table 5.1a. For each submodel, we discuss its dimensions and weight choices to fit the general random utility model described in Equation 5.3. Algorithm 1 shows the implementation snapshot of passenger journey execution behaviour.

i) **Initial stop model (ISM)**
The initial stop model determines the stop within a threshold radius where a passenger agent can wait for a minibus taxi. The model has three dimensions described in Table 5.4.

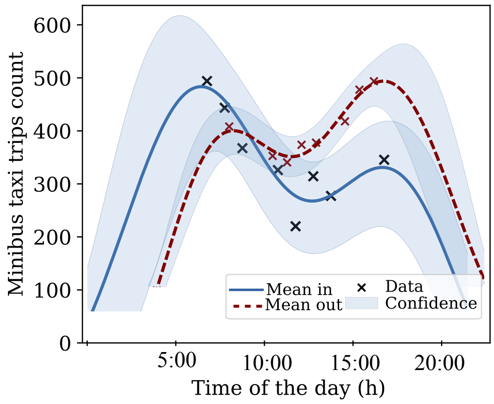
ii) **Boarding choice model (BCM)**
The BCM submodel has three dimensions described in Table 5.4. It is used by passenger agents to select a minibus taxi to board. If the available taxis do not completely fulfil the journey, it is split into connecting legs (see Figure 5.5b) and the journey diary updated accordingly.



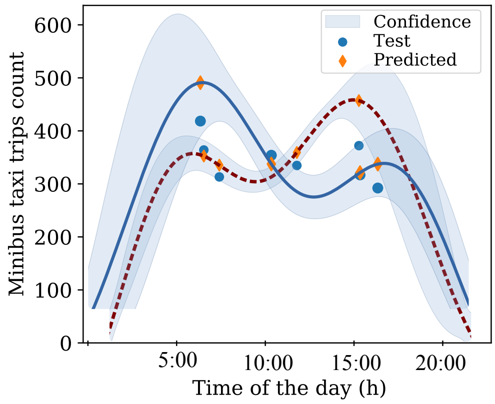
(a) Temporal distribution



(b) Prior



(c) Posterior



(d) Prediction test

Figure 5.7: Gaussian model fit to sparse minibus taxi trips data, prior and posterior trips profile from Gaussian model training and prediction test results for a 24-hour day.

iii) **Arrival choice model (ACM)**
The ACM is used to determine if the passenger agent should exit the minibus taxi paratransit system after executing a scheduled journey. An agent is considered arrived if it has been moved as close as possible to its intended destination; otherwise, it waits for a connecting minibus taxi trip to move it closer to its destination.

Table 5.4: Passenger agent submodels, dimensions, dimension metrics and associated weights. Note: ISM= Initial stop model, BCM=Boarding choice model, ACM=Arrival choice model.

Model	Dimension description	Metric	Weight
ISM	***First leg distance:*** The spatial distance from the passenger agent location to the candidate stop (see illustration in Figure 5.3c(i)).	km	0.2
	Candidate routes: The number of routes through the candidate stop to the passenger destination. Note: Each route is scored separately depending on its tenacity and how close it would take an agent to its final destination. Then the aggregate score for this dimension is obtained.	*	0.4
	Pickup likelihood: The probability of getting picked up from the candidate stop. This dimension simulates episodic memory. When the simulation starts, a default value is assumed by every agent, and it is occasionally updated by the agent based on its experience at a given stop.	*	0.4
BCM	***Last leg distance:*** The spatial distance from the candidate drop-off stop to the final journey destination (see illustration in Figure 5.3c(iii)). The lower the last leg distance, the higher the score for this dimension. Note: There is a threshold last leg distance, beyond which, a journey is split into two connection sub journeys as illustrated in Figure 5.5b.	km	0.6
	Connection journey distance: The distance from the current stop to the last leg connection stop (see illustration in Figure 5.3c(ii)).	km	0.25
	Instantaneous minibus taxi load factor: The percentage of minibus taxi occupancy at the current time step.	*	0.15
ACM	***Destination distance:*** The spatial distance to the final journey destination.	km	0.8
	Connection probability: The probability of getting a connection trip from the current stop.	*	0.2

Algorithm 1: ExecutePaxJourney Manages passenger agent behaviour during journey execution

```
foreach TIME STEP do
    Input: oPax, m   // oPax: object class in Algorithm 4 (Appendix A.1); m: model objects
    if oPax.state ='P' then
        sData ← oPax.sensor.getData()                  // Get agent surrounding data
        initStop ← oPax.runISM(sData)
        oPax.moveTo(initStop.pos)                      // Move to initial stop
        oPax.state ←'W'
    else if oPax.state ='W' then
        sData ← oPax.sensor.getData()
        oMv ←oPax.runBCM(sData)
        if oMv ⇒ NOT NULL then
            oPax.boardMBT(oMv)
            oPax.state ←'OB'
        else if oPax.wait_time > m.wait_threshold then
            oPax.dropJourney()
            oPax.state ←'P'
        end
    else if oPax.state ='OB' then
        oPax.updateLocation(oMv.pos)
        if oPax.pos = oPax.jny.dest_pos then
            oPax.disembark()
            oPax.state ←'A'
        end
    else if oPax.state ='A' then
        if oPax.runACM ⇒ TRUE then
            oPax.endJourney()
            oPax.state ←'P'
        else
            oPax.loadNextJourney()     // Load the next journey in the journey leg queue
            oPax.state ←'W'
        end
    end
end
```

5.3.3.3 Minibus taxi agent sub-models: RCM and PTM

This Section describes the submodels used to manage minibus taxi agent behaviour. The state diagram in Figure 5.5c shows the abstract description minibus taxi agent behaviour as a series of transitions between discrete states described in Table 5.1b. For each submodel, we discuss its dimensions and weight choices to fit the general random utility model described in Equation 5.3. Algorithm 2 show the implementation snapshot of minibus taxi trip execution behaviour.

i) Route choice model (RCM)
Before starting a trip, a minibus taxi agent decides where to go and the route to use for the trip. The RCM uses three dimensions (described in Table 5.5) to choose the route with high probability of making the trip profitable.

ii) Passenger touting model (PTM) Minibus taxi agents tout (or search) for passengers based on the PTM sub-model. The PTM consists of three dimensions, which are also sub-models as described in Table 5.5. The dimensions include the hold back model (HBM), the strategic demand estimation and the Lévy flight model defined by Equation 5.8 below.

$$f(l) \sim l^{-\alpha} \quad for \quad l \in [l_{min}, \infty), \tag{5.8}$$

where l is the step length and α (referred to as the Lévy exponent) is in the range $1 < \alpha \leq 3$ (Viswanathan et al., 2011).

Table 5.5: Minibus taxi agent submodels, dimensions, dimension metrics and associated weights. Note: RCM=Route choice model, PTM=Passenger touting model

Model	Dimension description	Metric	Weight
RCM	***Zone trips supply:*** This dimension is based on a system-wide origin-destination active trips matrix. The matrix monitors under-supplied zones based on the transport supply threshold (as learned from the Gaussian model in Section 5.3.3.1).	*	0.2
	Route tenacity: This dimension deals with the trustworthiness of route to have a threshold passenger demand and transport supply. Route tenacity is categorised into three (0-temporary, 1-candidate and 3-designated). New routes are introduced into the systems with the tenacity of zero, and if the usage of the route exceeds a threshold level, they are promoted to a higher tenacity. If the usage persistently declines, their tenacity is demoted until they are aged out of the system.	Route tenacity	0.3
	Demand & supply status: This dimension simulates the minibus taxi driver's episodic memory – the ability to recall previous short-term experiences. At the start of the simulation, a low default score is assigned to this dimension, and it is updated as the agent experience improves.	*	0.5
PTM	***Hold back:*** The hold back dimension score is determined by the Hold-back model (HBM) which the agent uses to randomly back off and wait for anticipated demand replenishment on the active route.	km	0.45
	Lévy flight: This dimension determines the distance between consecutive stops when touting for passengers. It is based on the Lévy flight hypothesis (Equation 5.8) discussed in Chapter 3. The objective is to generate step lengths the optimises the Lévy exponent α.	km	0.45
	Strategic demand estimation: This dimension helps the minibus taxi agent to determine whether to proceed with the unprofitable trip or not by strategically estimating the demand ahead based on a few clues such as the number of taxis on the same route and their respective load factors.	*	0.1

Hold back in this dissertation refers to the time a minibus taxi stays at a stop waiting in anticipation for passengers. Minibus taxis use a hold back model to determine how long to hold back at a stop to make the trip profitable.

Algorithm 2: ExecuteMbtTrip Manages minibus taxi agent behaviour during trip execution

```
Input: oMv, m                                          // Minibus taxi and model objects
foreach TIME STEP do
    if oMv.state ='P' then
        sData ← oMv.sensor.getData()                   // Get agent sorrounding data
        initStop ← oMv.findInitStop(sData)
        oMv.moveto(initStop.pos)
        oMv.state ←'S'
    else if oMv.state ='S' then
        if oMv.trip ⇒ NULL then
            oMv.state ←'R'
        else
            if oMv.trip.end ⇒ TRUE then
                oMv.trip ← NULL, oMv.state ←'P'
            else if oMv.loadCond ⇒ TRUE then
                oMv.state ←'L'
            else
                oMv.state ←'M'
            end
        end
    else if oMv.state ='R' then
        sData ← oMv.sensor.getData()
        if oMv.runRCM(sData) ⇒ TRUE then
            oMv.state ='S'
        end
    else if oMv.state ='L' then
        sData ← oMv.sensor.getData()
        if oMv.runHBM(sData) ⇒ FALSE then
            oMv.state ←'M', oMv.hbtm ← 0               // Reset holdback timer
        else
            oMv.hbtm ← oMv.hbtm + 1
            if oMv.hbtm > m.hbtm_threshold then
                oMv.abandonTrip()
                oMv.state ←'P', oMv.hbtm ← 0
        end
    else if oMv.state ='M' then
        sData ← oMv.sensor.getData()
        if oMv.stopInterrupt ⇒ TRUE then
            oMv.state ='S'
        else
            next_step_length ← oMv.runLFM()
            oMv.moveToNextStop(next_step_length)
            oMv.dist_travelled ← oMv.dist_travelled + next_step_length
            if oMv.tot_trip_dist >= oMv.dist_travelled then
                oMv.state ='S'
                oMv.trip.end ← TRUE
            end
        end
    end
end
```

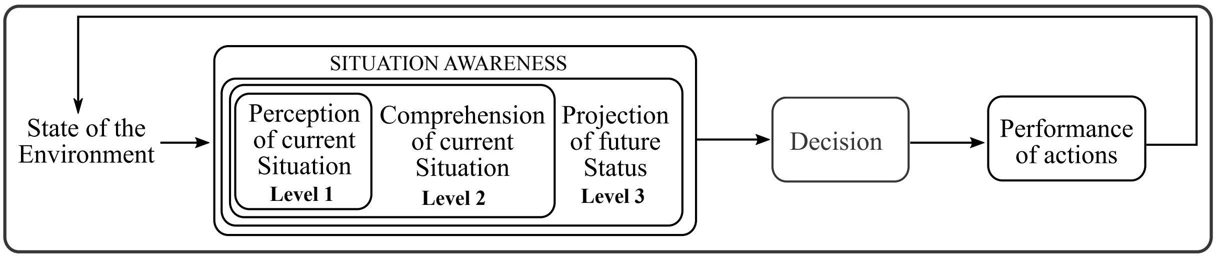


Figure 5.8: The cognitive model of situational awareness.

5.3.3.4 The cognitive model of situational awareness

Figure 5.8 describes a framework for situation awareness. An intelligent agent senses the environment, evaluates the current environment status, it projects the future environment state before making a decision. Once a decision is made, it performs the action and then updates its state.

5.4 Summary

In this chapter, we have designed and described an agent-based model of minibus taxi transportation in an organically-evolved, quasi-demand-responsive paratransit system similar to that in Kampala city.

Chapter 6

ABM model simulation and results

Chapter 6 Objectives

This chapter aims to achieve the research Objective 2.2 and Objective 2.3 of the dissertation to answer research questions three and four, respectively.

- ⇒**Research objective 2.2**

Implement and validate the designed agent-based model in a simulator. This includes studying the micro-level semi-autonomous interactions between Kampala's minibus taxis and passengers and analysing emergent behaviour at the macro level of the system.

- ⇒**Research objective 2.3**

Establish user-centric metrics for evaluating the efficiency of minibus taxi transportation in Kampala's organically-evolved, quasi-demand-responsive paratransit system.

This chapter presents the implementation and simulation of the minibus taxi agent-based model (ABM) developed and described in Chapter 5. The simulation results and validation are also presented at the end of this Chapter.

6.1 Simulating minibus taxi transportation dynamics

Simulation is the process of model execution that takes the model through discrete state changes over time. A model simulation environment based on MESA architecture (Masad and Kazil, 2015) was set up. The ABM structure, components, agents and their associated rules and behaviour were implemented using Python programming language. At runtime, the model goes through three phases. The first phase is *initialisation*, where the ABM loads the user defined model parameters (see Table 6.2); loads the external input data (such as minibus taxi stops); generates passenger journey diaries; and instantiates the model environment, together with the environment entities such as the stops, the scheduler and the data collector (refer to Section 5.3.2 of Chapter 5 for an in-depth discussion of model initialisation phase).

The second phase is the *runtime loop*, where scheduled agents repetitively interact (with self, with other agents and with their surroundings) in a shared environment, executing various actions, occasionally changing and updating their states and the environment. The simulation output data and screen are also updated during the second phase. The UML activity diagram in Figure 6.1 shows the general view of the three phases during the ABM simulation and the activity sequence at each phase. The third is the *exit* phase, where the model goes into the terminal state when the terminal condition is met. The rest of this chapter is about the second phase of the ABM simulation – runtime loop.

6.2 The runtime loop

After initialisation, the scheduler manages the orderly and repetitive execution of all model simulation activities – a process referred to as the *runtime loop*. During the runtime loop, the model executes through a series of time steps until it terminates, either normally through an exit function call or abnormally by an abort signal. The *Staged Activation* type of scheduler used for this ABM simulation consists of four stages. The first stage is *sensing* to get data about the surrounding agents and the environment. The second stage is *cognitive action*, where agents select appropriate action to take, often based on a utility maximisation method discussed in Section 5.2.2. The third is the *physical action*. This entails effecting the actual physical action such as moving to a new location, boarding a taxi, or abandoning an active trip. The fourth stage is the *update* stage. It involves updating the agents and the environment with the new state and the associated state data (refer to Table 5.1 for the description of agent states).

At each time step, the scheduler checks the exit condition. If the exit condition is not satisfied, the model advances to activate or deactivate passenger or minibus taxi agents. Passenger agents are activated based on the journeys available in the schedule that are pending execution at the appropriate time of the day. Minibus taxi agents are activated based on a supply matrix. To prevent oversupply or undersupply to different divisions of the simulated environment (Kampala city), the model maintains an up-to-date transport supply matrix that contains a system-wide view of division-level origin-destination minibus taxi volumes. The supply matrix is occasionally checked. If the transport supply volumes to any division falls below a pre-defined threshold, more minibus taxis are activated and assigned trips to the affected divisions. Or else, excess minibus taxi trips to such divisions are set to terminal such that the minibus taxi agent is deactivated at the end of the trip. Passenger agents are deactivated either after a successful journey execution or if it drops the journey due to failure to get a minibus taxi to the appropriate destination within the threshold waiting time.

6.2.1 Agent behaviour during runtime

The passenger and minibus taxi agents' behaviours were modelled based on our previous research findings documented in Chapters 3 and 4 of this dissertation and published in the Journal of Transport Geography (Ndibatya and Booysen, 2020a). In the research, as mentioned earlier, we found that:

1. In searching for, picking up and transporting passengers, Kampala's minibus taxis trajectory steps followed a heavy-tailed power-law distribution similar to a "*Lévy walk*".

2. Three passenger search strategies were used by minibus taxi drivers in Kampala, i.e.,

    i. *random passenger search*, where the minibus taxis start a trip with a few passengers in anticipation of passenger demand build-up along the route;

    ii. *random back off* or holding-back, where the driver interrupts the trip for a random period to allow for passenger demand replenishment on the route before continuing; and

    iii. *trip abandonment*, where trips deemed unprofitable by the drivers are either abandoned or the trip routes are changed to new destinations with anticipated high demand.

3. When passengers wait at a stop for a certain period without getting a taxi, they abandon the intended minibus taxi-based journey and often use motorcycle taxis to get to their destinations quickly.

We, therefore, modelled the agents' behaviour in the agent-based simulation to closely match the expected behaviour of passengers and minibus taxis in Kampala's paratransit system.

During the runtime loop, passenger agents examine their journey diaries to find if there is a journey to execute depending on the time of the day. The candidate journey is then scheduled for execution, and the journey diary updated accordingly. The minibus taxis are scheduled into the model based on a Gaussian submodel (described in Section 5.3.3.1). Every minibus taxi agent searches for the best route to a predetermined destination division, and then instantiates a trip to the destination through the selected route. En-route, the minibus taxi searches for passengers based on the three search strategies mentioned earlier, and a utility maximisation function discussed in Section 5.2.2. The minibus taxis occasionally hold-back at random stops in anticipation for passengers. Unprofitable trips are abandoned, and new trips anticipated to be profitable are initiated and executed.

Passenger and minibus taxi agents store metric data related to the fully or partially executed journeys and trips, which they use to train agent-specific submodels every five days for improved

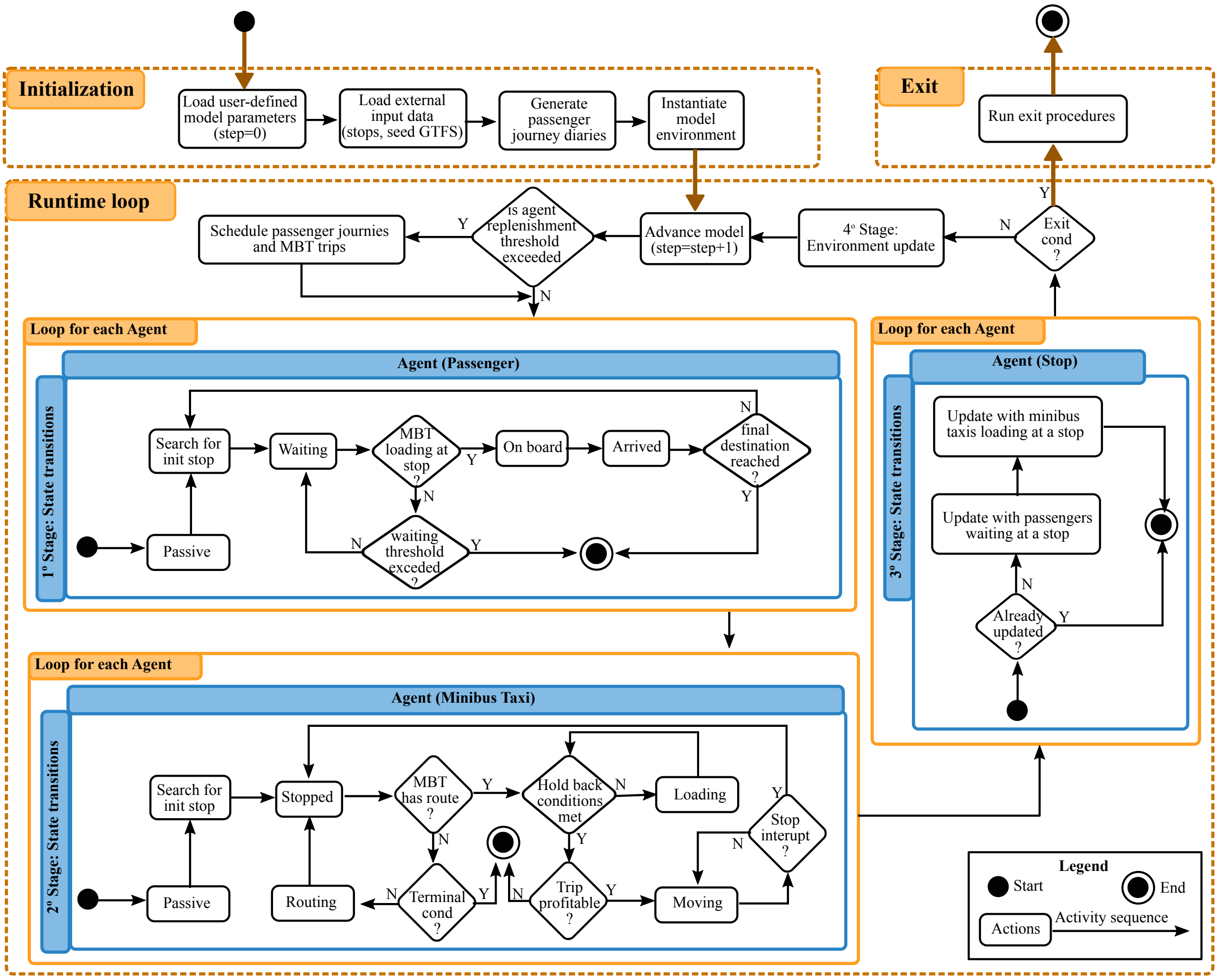


Figure 6.1: UML activity diagram of our ABM framework. The diagram shows the activity sequences of the entire model algorithm. The orange box represents a loop for each agent, and the blue boxes show the different agent steps.

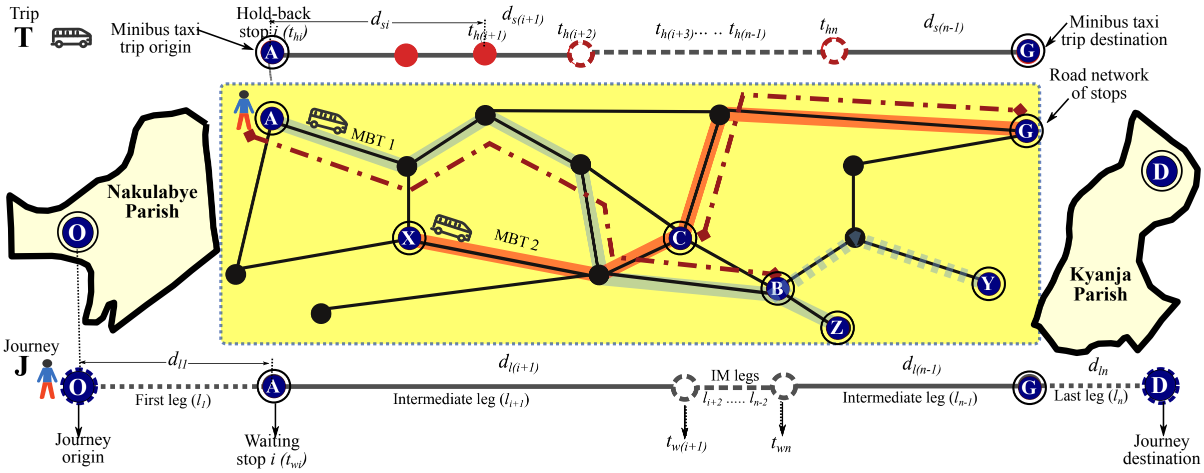
future decision making. Figure 6.1 shows the detailed activity sequence for each agent during the runtime loop.

6.2.2 Metrics evaluated during the runtime loop

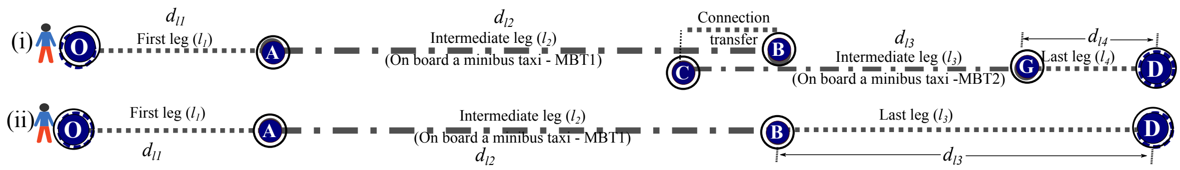
Conceptually, there are two main active and co-dependent entities at model runtime – the *journey* (managed and executed by the passenger agent) and the *trip* (managed and executed by the minibus taxi agent). At the beginning of the simulation, both the journeys and trips have a reciprocal but asymmetric relationship where the demand for journeys often exceeds the supply of trips and vice versa. Hence the simulation objective is to facilitate autonomous agent interactions in a common environment. During the agents' interactions, we measure and analyse selected metrics' values to establish emergent characteristics and other phenomena in the simulated minibus taxi transport system. This section presents the journey and trip metrics evaluated (see Table 6.1) and describes the relationship between the metrics (see Figure 6.2a). Figures 6.2b and 6.2c illustrate scenarios of passenger journeys and minibus taxi trips fully executed (6.2b(i) and 6.2c(ii)), partially executed (6.2b(ii)) and a case of trip abandonment by a minibus taxi driver (6.2c(i)).

Consider two spatially distant locations 'O' (representing a random origin in Nakulabye parish) and 'D' (representing a random destination in Kyanja parish) shown in Figure 6.2a. A network of stops (extracted and summarised based on the method described in Section 5.3.2.1) connects both parishes – the lowest territorial and administrative entities in Kampala city. Also consider an imaginary journey 'J' originating from O to D. The journey (J) can be sub-divided into three parts, namely: the *first leg* l_1 (sometimes referred to in transport studies as the *first mile of commute*); the intermediate legs $l : \{l_2, l_3, ..., l_{n-2}, l_{n-1}\}$, where n is the total number of legs; and the *last leg* l_n (sometimes referred to in transport studies as the *last mile of commute*). The passenger agent (initially in passive state) begins the journey by executing the first leg l_1 (O to A). Executing the first leg involves searching for and moving to the anticipated optimal stop (e.g., stop A) to get and board a minibus taxi (refer to Section 5.2.2.1 for the details about passenger decision making process). Stop A marks the beginning of the intermediate legs for journey J shown in Figure 6.2a. At the beginning of each intermediate journey leg, a passenger waits for a minibus taxi to take them as close as possible to the final destination D. The time taken waiting at the stop before boarding the taxi, t_w, is recorded. If the passenger fails to get a taxi from a given stop, they may opt to try and wait at another nearby stop, until the threshold waiting time is exceeded hence the passenger terminates the journey leg. If the leg termination happens at the beginning of the intermediate legs (l_2), the whole journey is considered to have failed. If the leg termination happens at any other leg after l_2, the journey is considered partially fulfilled, hence the last leg distance d_{ln} is measured from the previously successful intermediate leg destination. The journey is considered fully completed if the last successful intermediate leg destination is within a threshold last leg distance computed as a percentage of the total spatial distance from O to D. Finally, the passenger agent computes and stores all the journey-related metrics' values. Table 6.1a describes the metrics related to a passenger journey.

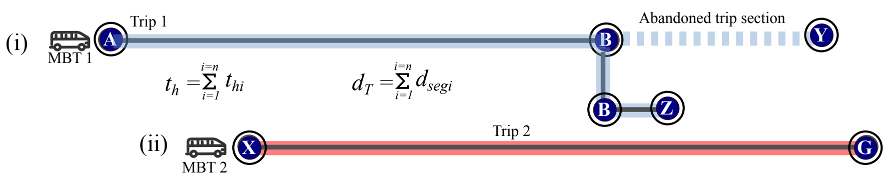
Also consider a minibus taxi agent (in passive state) at a random stop A. When the minibus taxi is activated, it determines its trip destination based on an inter-division transport supply matrix discussed earlier – a process we refer to in this dissertation as *self-selection* of origins and destinations. The minibus taxi then computes an optimal route through the stops network using route distance as a metric. Once the route to its destination is acquired, the minibus taxi executes a trip (e.g., trip T from stop A to stop G illustrated in Figure 6.2a) following the rules discussed in Section 5.3.3.3. During trip execution, the minibus searches for and loads passengers who are waiting along the route. It randomly holds back at various stops to improve its instantaneous occupancy (or load factor). Suppose the load factor is persistently



(a) Trip and journey general description.



(b) Two execution scenarios of a passenger journey from O to D.



(c) Two execution scenarios of minibus taxi trips.

Figure 6.2: (a) Model runtime journey and trip description; (b) Journey execution scenarios; (ci) Illustration of trip abandonment by a driver; (cii) Illustration of a connection trip, i.e., fulfilling the second intermediate leg of the journey in 6.2bi.

below the minimum threshold. In that case, the trip is abandoned by the driver and a terminal signal is sent to the passengers on board to disembark and search for connecting trips to their respective final destinations. The distance d_{si} between consecutive hold-back stops is generated from a Lévy probability distribution defined by Equation 4.1. The Lévy exponent α value in the equation is set to $\alpha = 2.2$ (an average value obtained from our research in Chapter 4). A minibus taxi trip is considered, (i) complete, if it reaches its final destination stop G, (ii) incomplete if it is abandoned before its final destination, and (iii) failed if it does not get enough passengers to leave the stop of origin A. Finally, the minibus taxi agent computes and stores all the trip-related metrics' values. Table 6.1b describes the metrics related to minibus taxi trips.

The Figures in 6.2b illustrate two independent execution scenarios for a journey from origin O to destination D and the associated metric measurements for each scenario. In the scenario in Figure 6.2b(i), the passenger moves to the initial stop A (executes the first leg) and waits for a minibus taxi for t_w hours. While at stop A, a minibus taxi destined to stop Y (executing a trip illustrated in Figure 6.2c(i)) passes by. The passenger decides to board the minibus taxi. However, the minibus taxi abandons the trip to Y in favour of stop Z as the final destination, thereby dropping off the passenger at stop B. Since the passenger's full journey length was not fulfilled by the minibus taxi trip (Trip 1 in Figure 6.2c(i)), the passenger splits the journey (see Figure 5.5b) into two intermediate legs, and updates the journey diary accordingly. The first intermediate leg (l_2 – from stop A to stop B) which was fulfilled by trip 1 (Figure 6.2c(i)). The passenger waits for a connection trip at stop B for a random period below the maximum waiting threshold. When the passenger fails to get a minibus taxi at stop B, the passenger moves to a nearby stop C and continues waiting. While at C, another minibus taxi executing trip 2 (X to G – Figure 6.2c(ii)) arrives at the stop and the passenger boards to fulfil the second intermediate leg (l_3 – from stop C to G). At stop G, the passenger disembarks from the minibus taxi, computes the last leg distance ($l_n = l_4$), and tags the journey as complete. The passenger further computes and stores other derived metrics values such as the intermediate legs count l_{count}, the intermediate legs distance d_l, the total waiting time t_w, the total waiting stops sw_{count}, the first d_{l1} and last legs distances d_{l4} distances.

Table 6.1: Metrics used for evaluating the simulated minibus taxi transportation system efficiency.

(a) Passenger journey metrics.

Metric	Description	Unit
Waiting time (t_w)	The total time a passenger agent waits (outside) for a minibus taxi at a stop (or stops) during execution of a single journey. Note: for a multi-leg journey, $t_w = \sum_{i=1}^{n} t_{wi}$, where, t_{wi} is the time spent waiting for a minibus taxi at a single stop.	*hours*
Stops at which waited (sw_{count})	The number of stops at which a passenger waits for a minibus taxi to arrive during a single journey execution (including a journey with multiple legs).	*count*
First leg distance (d_{l1})	The spatial distance between a journey origin and the stop where the passenger agent successfully boards a minibus taxi.	*km*
Last leg distance (d_{ln})	The spatial distance between the stop where a minibus taxi drops off a passenger and the final destination of a passenger journey.	*km*
Intermediate legs (l_{count})	The number of minibus taxi connections a passenger makes to fulfil a single journey.	*count*
Intermediate legs distance (d_l)	The sum of distances of all journey legs executed (while onboard a minibus taxi) during a single journey. $d_l = \sum_{i=1}^{m} d_{li}$, where, d_{li} is the distance of the i^{th} intermediate leg l_i.	*km*

(b) Minibus taxi trip metrics.

Metric	Description	Unit
Hold-back time (t_h)	The accumulated time a minibus taxi stays at stops along a route waiting for (or in anticipation of) passengers. It includes the time spent at the stop of origin loading passengers before embarking on a trip.	*hours*
Hold-back time per km	The average measure of hold-back time per unit distance travelled by a minibus taxi during the trip.	t_h/km
Occupancy ($\mathcal{O}$)	The instantaneous number of passengers onboard a minibus taxi as a percentage of the total taxi capacity. Note: we modelled only fourteen-seater minibus taxis, and the instantaneous occupancy was measured every time a minibus taxi left a stop (during the moving (M) to loading (L) state transition).	*ratio*
Total trip distance (d_T)	The sum of distances of all road segments from the trip origin (origin stop) to the trip destination (destination stop). $d_T = \sum_{i=1}^{n} dseg_i$, where $dseg$ is the distance of the road segment between two subsequent stops on the route where the trip was executed.	*km*
Trip operating speed (v_o)	The average speed at which a minibus taxi could travel from origin to destination without stopping en route. When computing the operating speed, we excluded the hold-back time. Note: conceptually, v_o was assumed to be constant and its value set during model initialisation (see Table 6.2). However, during the simulation execution, the values v_o varied because of the need to optimally execute the Lévy steps generated.	km/h
Trip commercial speed (v_c)	The overall average speed of the minibus taxis during the trip, including the time spent at the stops (hold-back time). $v_c = \frac{d}{t}$, where, d is the total trip distance, and t is the total time taken to complete a trip.	km/h

Figure 6.2b(ii) illustrates a scenario of an incomplete journey execution, where, after the first intermediate leg (l_2 – A to B), the passenger failed to get a connecting trip hence recording a longer last leg distance d_{l3} (in this case the last leg is l_3).

6.3 Design of simulation experiments

We designed and ran two simulation experiments – the *Controlled* and *Test* experiment – to more substantially study the minibus taxi transportation dynamics in an organically-evolved, quasi-demand-responsive paratransit setting as well as to overcome the scaling resource limitations associated with field studies. In both experiments, the passenger journeys' demand and the spatial and temporal characteristics of the journeys' demand were kept constant (see Section 6.3.1 for the journeys demand characteristics). We varied the factors that influence demand-responsiveness in the system such as, agents' situational awareness, passenger search behaviour and episodic memory. We then analysed the resultant metrics' data to assess the effect on the overall paratransit system's efficiency and eventually answer research questions RQ3, RQ4, and RQ5 as stated in Section 1.3.1 of Chapter 1.

During the controlled experiment, in addition to the constant passenger journeys demand characteristics, we tuned the minibus taxi trips distance d_T distribution to closely match the minibus taxi route lengths (with mean 5.85 km) observed during the field research (see Chapter 3). We iteratively adjusted the agent's behavioural logic, while analysing the resultant metrics values until the distributions of metrics values closely matched the known system values observed during the field research, indicating that the dynamics in both systems were approximately the same. We further performed a statistical analysis of the metrics values (see Tables 6.3b, 6.4b and Figures 6.5 and 6.7) to identify the related efficiency metrics.

During the test experiment (see Chapter 7), we trained the passenger and minibus taxi agents to adopt behaviour that optimises values of the dependent efficiency metrics identified during the controlled experiment simulation. Thus, the test experiment simulated agents with improved situational awareness, more optimal demand responsiveness, and considerable long-term memory. We then analysed the resultant metrics values and compared them with the values observed from the controlled experiment.

6.3.1 Passenger journeys daily demand characteristics

The Figures 6.3a to 6.3f show the daily spatial and temporal distributions of the journeys that were input into the simulation models. Figure 6.3a shows the journey diaries generated for each division in Kampala and the corresponding number of journeys in their journey queues. Correspondingly, 74% of the journey diaries generated had one journey in their journey queue (see Figure 6.3a), 22% had two journeys, whereas 4% had three journeys in their respective journey queues (refer to Figure 5.3 in Chapter 5 for the relationship between the journey, journey plan, journey diary, and journey queue). The journey diaries with more than one journey in their journey queue represent passengers who relied on the minibus taxi transport system for multiple journeys during the same day. For example, a passenger who planned to travel by minibus taxi for the journeys from home-to-work, work-to-shop, and shop-to-home would have a journey diary consisting of three journeys in the journey queue, as illustrated in Figure 5.3a. Figure 6.3b shows the division-level origin-destination heat-map of passenger journeys. Figure 6.3c shows the distribution of journeys' distances from the origin to destination. Figures 6.3d and 6.3e show the spatial distribution of passengers' journeys at the parish level in Kampala. Figure 6.3f shows the temporal hourly distribution of passengers' journeys throughout the simulated day.

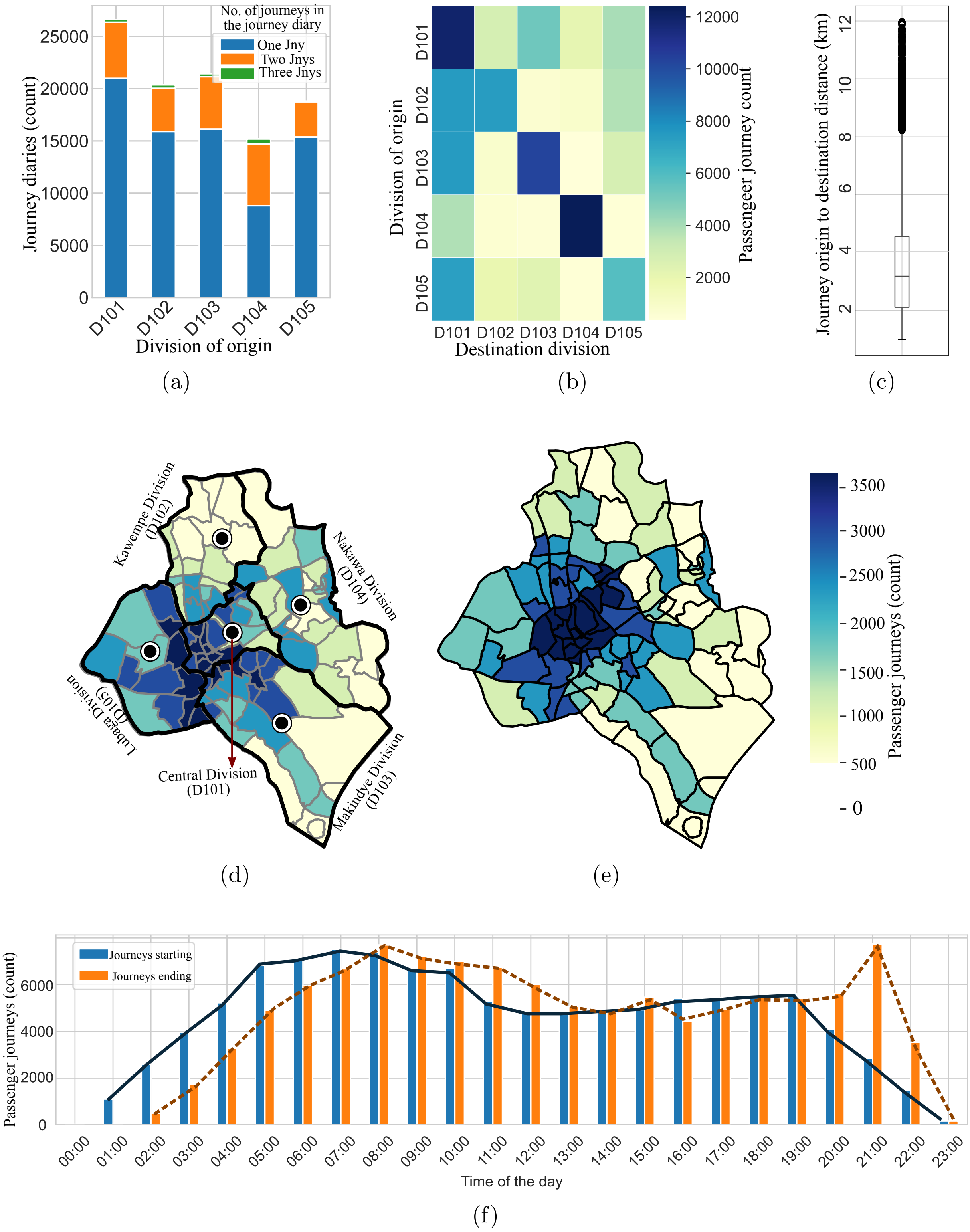


Figure 6.3: Description of daily passenger journeys' demand characteristics: (a) Journey diaries per division and their respective number of journeys in the journey queue; (b) Passenger journeys origin-destination divisions; (c) Distribution of journeys beeline distances from origin to destination; (d) Spatial distribution of passenger journeys' parish origins; (e) Spatial distribution of passenger journeys' parish destinations; and (f) Hourly temporal distribution of passenger journeys for one day.

6.4 Set up of the controlled experiment (CER)

This section describes the setup and execution of a controlled simulation experiment designed to simulate the minibus taxi transportation dynamics in Kampala's paratransit system. The simulation implements the agent-based model (ABM) developed in Chapter 5. The main goal of this controlled experiment is to replicate the minibus taxi paratransit system in Kampala as closely as possible. As such, we tuned some of the known system characteristics (see Section 6.2.1) and other constraining attributes such as minibus taxi trip lengths, locations of minibus taxi stops, the spatial and temporal distribution of minibus taxi trips and passenger journeys. We simulated the autonomous interactions of the minibus taxis and passenger agents, and we analysed the metrics values associated with their journeys and trips, respectively. Finally, we validated the simulation results by comparing the graphical result from the simulation with the results of similar metrics obtained from the previous field research.

6.4.1 Initialisation

The initialisation phase involves loading user-defined parameters, loading external input data, generating passenger journey diaries, determining the threshold trips matrix, and instantiating the model environment and the environment entities. Refer to Section 5.3.2 for the details about model initialisation.

6.4.1.1 Description of inputs

For model inputs, we used data from four primary sources. First, we developed a custom General Transit Feed Specification (GTFS) file and loaded it with data of 796 stops that were tagged during the field research (see Chapter 3). A standard GTFS consists of six required files, i.e., agency, stops, routes, trips, stop times, and calendar. However, as observed by Williams et al. (2015), semi-formal transit systems such as minibus taxis operate differently from traditional buses. Their GTFS thus requires customisation to cater for demand responsiveness and other peculiar characteristics observed and discussed in Chapter 3. In this simulation setup, we only populated the stops file. Other files such as trips, routes and stop times were dynamically built during model execution.

Second, we loaded the Open Street Map (OSM) shapefiles for Kampala (OpenStreetMap contributors, 2017). The files contained Kampala's primary, trunk, secondary and tertiary roads, as well as division and parish administration levels. We combined the roads shapefiles with the minibus taxi stops from the GTFS files to generate a stops network at the initialisation stage, as discussed in Section 5.3.2.1. Conceptually, the stops network is a weighted, undirected and connected Networkx Graph.

Table 6.2: Description of some of the selected initial simulation model parameters

Model parameter	Description	Value	Unit
LF_THRESHOLD	Load factor threshold	0.3	
HBM_THRESHOLD	Hold-back time threshold	1	Hours
WAIT_THRESHOLD	Waiting time threshold	2	Hours
OP_SPEED	Operating speed	20	km/h
SYS_CAPACITY	System capacity	20%	
STEP_TO_MIN	Time step to minute conversion	1:5	
MBT_CAPACITY	Minibus taxi maximum passenger capacity	14	pax

Third, we used Kampala's 2014 population census data, together with traffic count data from JICA's 2010 transport improvement study, to model the daily public transport demand (expected passenger journeys). The base minibus taxi supply in Kampala which was also used as input to the simulation model (JICA, 2010; UBOS, 2014). Finally, we specified the values of the model's initial parameters as shown in Table 6.2.

6.4.2 Results

This subsection presents the results from a controlled simulation experiment. The simulation represents seventy days of agents' interactions in the virtual quasi-demand-responsive minibus taxi paratransit system environment. The simulation scenario implements agents' behaviour similar to those we observed from our field study in Kampala (see Chapters 3 and 4). They include Lévy passenger search behaviour, random back off, trip abandonment, and limited situational awareness (episodic memory – an agent's unique memory of a specific event). The results are presented in two parts, namely: the results related to passenger journeys (Section 6.4.2.1), and the results related to a minibus taxi trip (Section 6.4.2.2). For purposes of presenting results in this section, we randomly sampled and analysed results from twenty days of the seventy-day simulation.

6.4.2.1 Results: Passenger journeys

Analysis of executed passenger journeys show low journey completion rates (40%) for the first twenty days; moderate improvement– to 70%– between the twenty-first and fiftieth day; and then stabilising at 80% after the fiftieth day as illustrated in Figure 6.4a. There was a substantial percentage of partially completed journeys throughout the model simulation. Trip abandonment by drivers contributed an insignificant share of the failed journeys compared with the journeys that failed due to exceeding the waiting threshold at the stops.

Table 6.3a gives the statistical summary of the passenger journey metrics values measured during the simulation execution. Figures in 6.4b and 6.4c show the distributions associated with the respective journey metrics. The long first and last leg distances with mean values of 1.42km and 1.72km, and standard deviation values 0.85km and 1.15km, respectively (see Table 6.3a) suggest that passengers struggled to find the right locations with high chances of getting a taxi. The long last leg distance could also be because of the high number of partially completed journeys. The heavy-tailed distribution of the intermediate legs distance is indicative of either presence of *circling* within the system (where passenger agents boarded minibus taxis repeatedly without reaching their final destination) or selecting and boarding taxis going through longer routes before getting to the passengers' destination. The mean waiting time of 1.17 hours and standard deviation of 1 hour was most likely caused by the observed over-fragmentation of journeys into multiple legs. Each journey was split into two-to-three intermediate legs. Fragmentation of journeys could further be responsible for the high number of "stops waited at" sw_{count} and thus the high waiting time. However, the high sw_{count} could also be because of difficulty in finding minibus taxis. Thus, the passenger agents had to change to multiple locations before eventually getting a taxi to their destination.

Correlation analysis of the passenger journeys metrics values (see Figure 6.5 and Table 6.3b) shows a strong positive correlation between waiting time and "stops waited at" ($r = 0.88$); stops waited at and legs count ($r = 0.86$); and waiting time and intermediate legs count ($r = 0.71$). There is a moderate correlation between intermediate legs distance and intermediate legs count ($r = 0.4$); stops waited at and intermediate legs distance ($r = 0.35$); and intermediate legs distance and waiting time ($r = 0.32$).

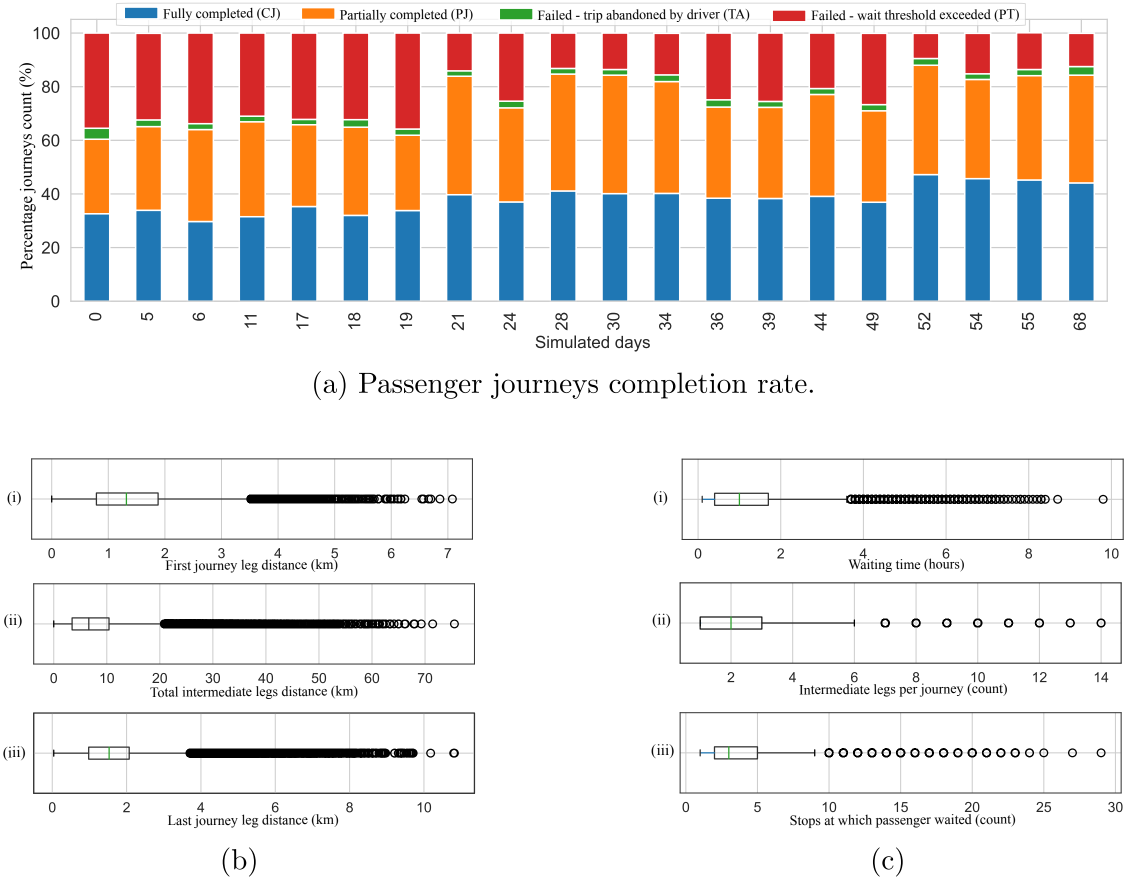


Figure 6.4: Controlled experiment results: (a) Daily passenger journeys' completion rates; and distributions of: (b)i First leg distance d_{l1}, (b)ii intermediate legs distance d_l, (b)iii last leg distance d_{ln}; (c)i Waiting time t_w, (c)ii Intermediate legs count l_{count} and (c)iii Stops at which passengers waited sw_{count}.

6.4.2.2 Results: minibus taxi trips

Results from the minibus taxi trips' analysis show a moderate variation of minibus taxi occupancy during the model runtime. Occupancy varied between 30% and 70% as illustrated in Figure 6.6a. Table 6.4a summarises the statistical values of the trip metrics values measured, whereas Figures 6.6 show the associated distribution plots.

Correlation analysis of metrics results values show a strong negative correlation between hold-back per kilometre and the commercial speed ($r = -0.74$); trip distance and commercial speed ($r = 0.61$); and moderate correlation between hold-back time and hold-back per km ($r = 0.53$), commercial speed ($r = -0.48$), occupancy ($r = 0.34$), operating speed ($r = -0.33$); and hold-back per km and operating speed ($r = -0.43$). Table 6.4b and Figure 6.7 show the correlation analysis results.

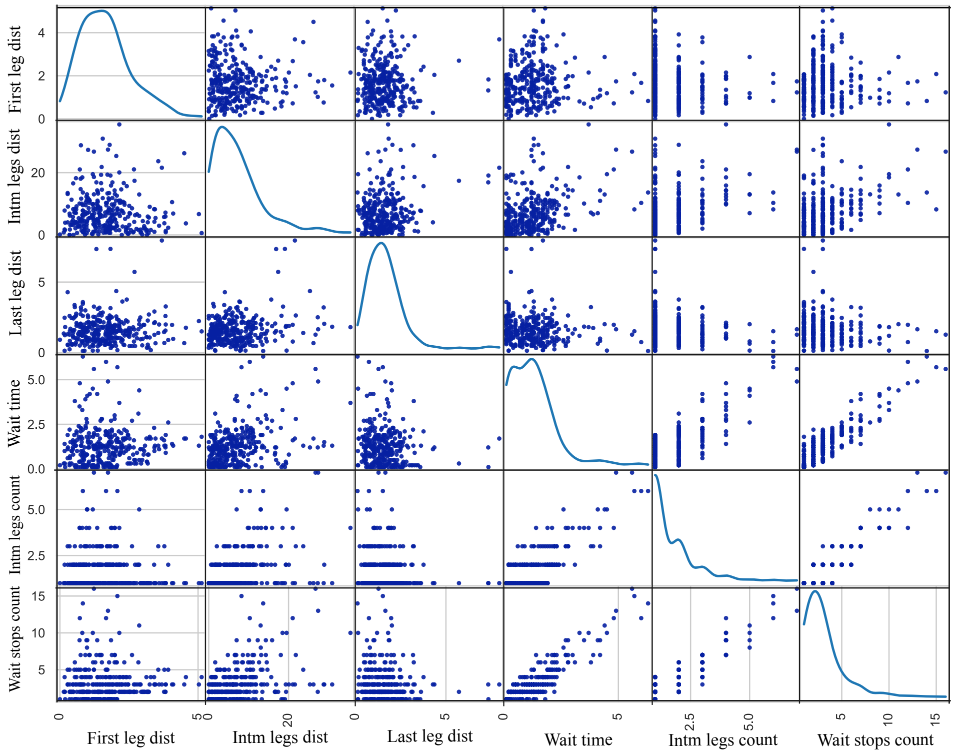


Figure 6.5: Scatter plot matrix visualising bivariate relationships between combinations of selected journey metrics values from the controlled experiment simulation

6.4.3 Model validation

We used a combination of two validation methods for agent-based modelling– *input validation*, and *descriptive output validation* (Xiang et al., 2005; Ormerod and Rosewell, 2009). Input

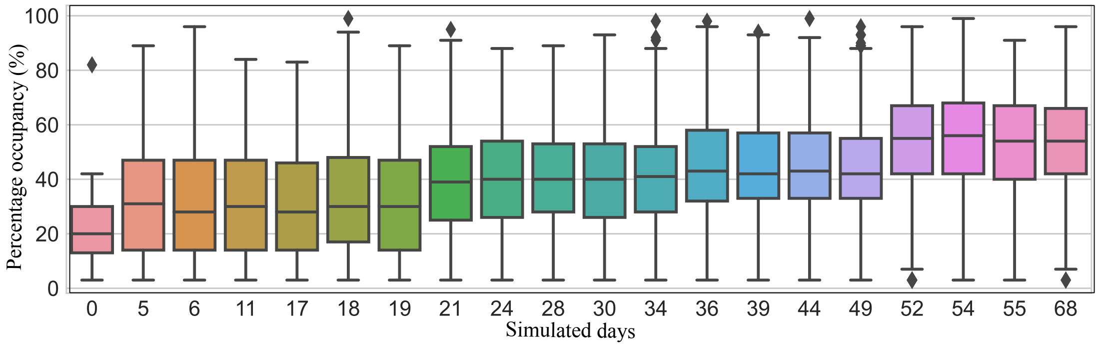
Table 6.3: Summary statistical analysis of selected metrics values of passenger journeys fully and partially completed during the controlled experiment simulation.

(a) Mean, standard deviation, and percentage quantile values of selected journey metrics.

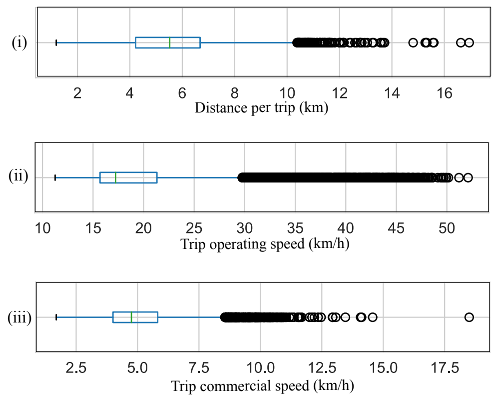
	First leg distance (d_{l1})	Last leg distance (d_{ln})	int legs distance (d_l)	Waiting time (t_w)	Stops waited at (sw_{count})	legs count (l_{count})
Mean	1.42	1.72	8.05	1.17	3.68	2.05
std	0.85	1.15	6.95	1.01	2.68	1.38
Q 25%	0.79	0.98	3.46	0.4	2	1
Q 50%	1.32	1.53	6.61	1	3	2
Q 75%	1.88	2.07	10.42	1.7	5	3
Max	7.08	10.82	75.54	9.8	29	14

(b) Correlation matrix of selected journey metrics values.

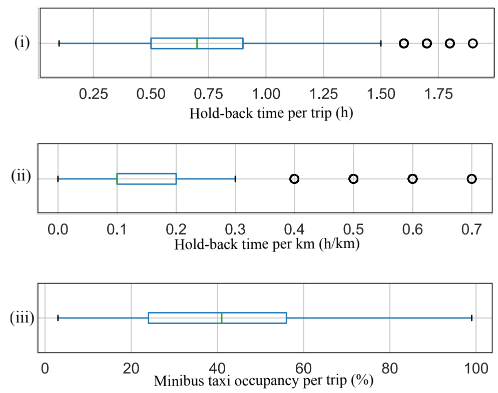
	First leg distance	Last leg distance	Int legs distance	Waiting time	Stops waited at	Legs count
First leg distance (d_{l1})	1	-0.01	-0.03	0.1	0.1	-0.05
Last leg distance (d_{ln})	-0.01	1	0.06	-0.1	-0.13	-0.14
Int legs distance (d_l)	-0.03	0.06	1	0.32	0.35	0.
Waiting time (t_w)	0.1	-0.1	0.32	1	0.88	0.71
Stops waited at (sw_{count})	0.1	-0.13	0.35	0.88	1	0.86
Legs count (l_{count})	-0.05	-0.14	0.4	0.71	0.86	1



(a) Variation of minibus taxi occupancy during controlled simulation experiment.



(b)



(c)

Figure 6.6: Controlled experiment results: (a) Varying minibus taxi occupancy per day; and distributions of: (b)i Total trip distance d_T,(b)ii trips operating speed v_o, (b)iii trips commercial speed v_c, (c)i hold-back time t_h, (c)ii hold-back per km and (c)iii occupancy $\mathcal{O}$.

validation uses information about parameter values that come from external knowledge of the system's microbehaviour. Descriptive output validation (or *face validity*) matches the computationally generated output with pre-existing data on the process being modelled (Institute of Medicine, 2015). Thus, for input validation, we used the known passenger journey demand characteristics (described in Section 6.3.1 and Figures 6.3) as inputs, as well as defining other known agents' behavioural constraints (described in Section 6.2.1) at runtime. Therefore, the inputs provide a degree of microbehaviour closely consistent with the Kampala's minibus taxi quasi-demand-responsive paratransit system that was modelled.

Furthermore, the simulation results were validated based on *face validity* as mentioned earlier: this involved interpretation and comparison of graphical results. We compared the distributions of several metrics' values that were arrived at during the field study (refer to the results in Chapter 3) with the control experiment results to ascertain if the two result sets closely match. The Figures in 6.8 show the comparisons of selected metrics values distributions of the field results (FR) with the controlled experiment results (CER). Table 6.5 shows a side-by-side comparison of the results from both studies, i.e., the FR and CER.

For purposes of input validation, the trips distances d_T were kept in close range with mean values of 5.01 km and 5.5km for FR and CER, respectively. The low CER occupancy observed during the first thirty simulation days (see Figure 6.6a) caused the moderate difference between the mean occupancy values ($\Delta\mu = 27\%$, see Table 6.5). This could be attributed to the time minibus taxi agents had not yet built up enough episodic memory to make semi-optimal decisions. The same reason could explain the $\Delta\mu$ values observed for the Waiting time

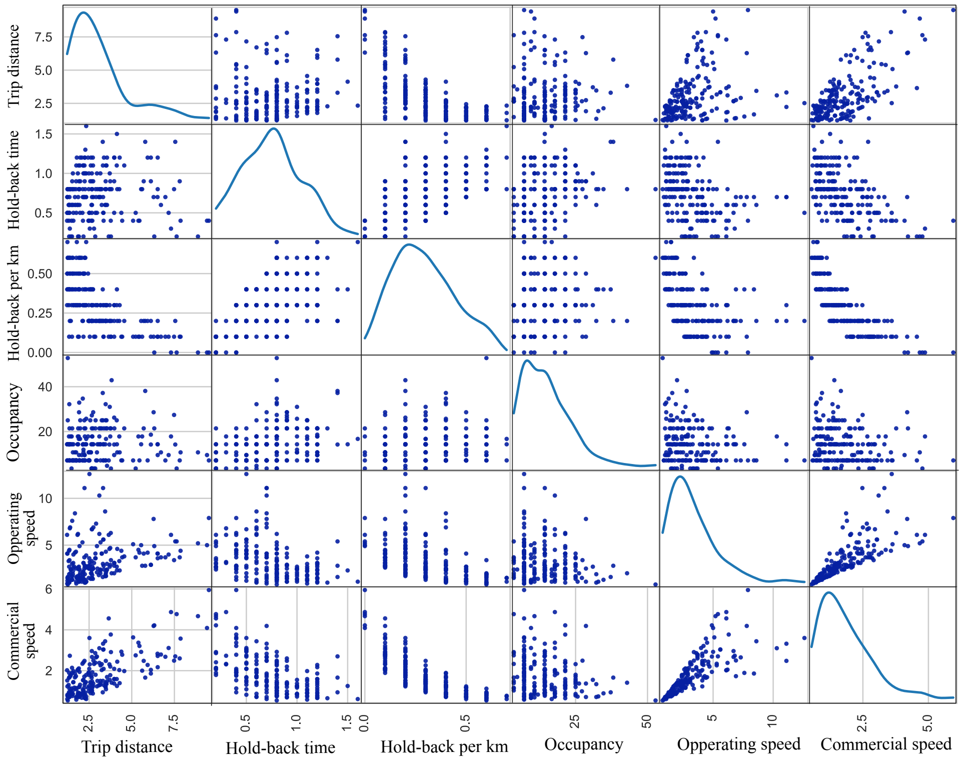


Figure 6.7: Scatter plot matrix visualising bivariate relationships between combinations of selected minibus taxi trips metrics values from the controlled experiment simulation

($\Delta\mu = -0.5$), Hold-back time ($\Delta\mu = 0.23$), hold-back per km ($\Delta\mu = 0.09$) and commercial speed ($\Delta\mu = -0.55$).

The graphical results from the controlled simulation experiment are closely related to the

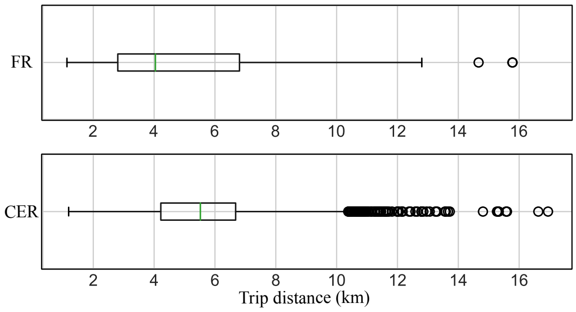
Table 6.4: Summary statistical analysis of selected metrics values of minibus taxi trips fully and partially completed during the controlled simulation experiment.

(a) Mean, standard deviation, and percentage quantile values of minibus taxi trips metrics.

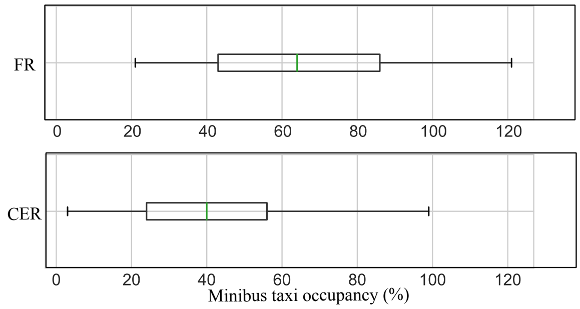
	Trip distance (d_T)	Hold-back time (t_h)	Hold-back per km (t_h/km)	Occupancy ($\mathcal{O}$)	Operating speed (v_o)	Commercial speed (v_c)
Mean	5.45	0.73	0.15	40.72	19.82	5.02
std	1.84	0.35	0.1	21.01	6.86	1.59
Q 25%	4.22	0.5	0.1	24	15.69	4
Q 50%	5.52	0.7	0.1	41	17.23	4.75
Q 75%	6.68	0.9	0.2	56	21.32	5.82
Max	16.95	1.9	0.7	99	52.1	18.5

(b) Correlation matrix of selected minibus taxi trips metrics values.

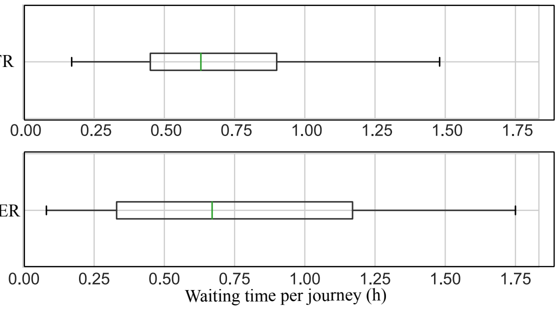
	Trip distance	Hold-back time	Hold-back per km	Occupancy	Operating speed	Commercial speed
Trip distance (d_T)	1	0.11	-0.61	0.01	0.26	0.61
Hold-back time (t_h)	0.11	1	0.53	0.34	-0.33	-0.48
Hold-back per km (t_h/km)	-0.61	0.53	1	0.18	-0.43	-0.74
Occupancy ($\mathcal{O}$)	0.01	0.34	0.18	1	-0.18	-0.2
Operating speed (v_o)	0.26	-0.33	-0.43	-0.18	1	0.8
Commercial speed (v_c)	0.61	-0.48	-0.74	-0.2	0.8	1



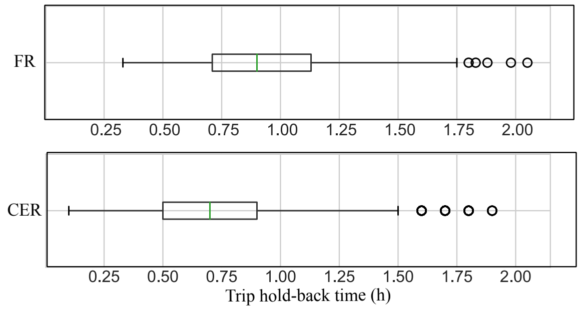
(a) Comparing trips distance d_T.



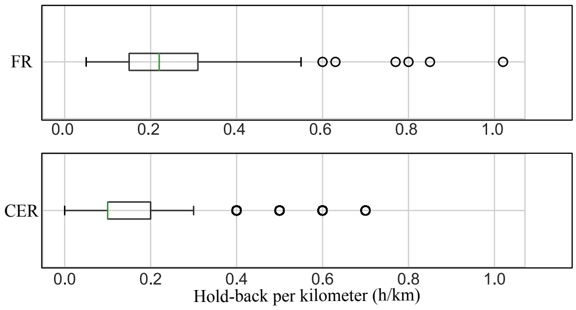
(b) Comparing minibus taxi occupancy $\mathcal{O}$.



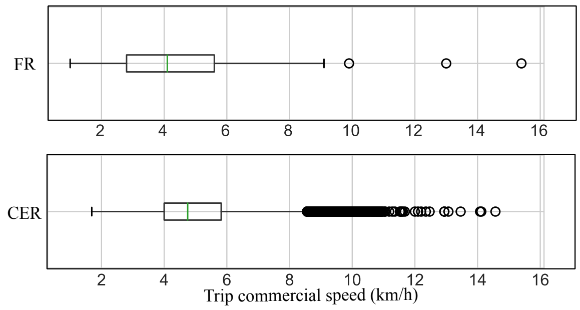
(c) Comparing passenger waiting time t_w.



(d) Comparing trips hold-back time t_h.



(e) Comparing hold-back time per km t_h/km.



(f) Comparing trips commercial speed v_c.

Figure 6.8: Model validation results: Side-by-side graphical comparison of distributions of selected journeys and minibus taxi trips metrics values from two studies, i.e., the field study results (FR) with the controlled experiment results (CER).

Table 6.5: Model validation: Side-by-side numeric comparison of field study results (FR) with the controlled experiment results (CER) for the mean μ, standard deviation σ, twenty-fifth quantile (Q 25%), fiftieth quantile (Q 50%) and the seventy-fifth quantile (Q 75%). $\Delta\mu$ is the difference between the mean values of the two results sets.

	Mean μ			Std dev σ		Q 25%		Q 50%		Q 75%	
	FR	CER	$\Delta\mu$	FR	CER	FR	CER	FR	CER	FR	CER
Trip distance d_T (km)	5.01	5.45	*-0.44*	2.95	1.84	2.8	4.22	4.04	5.52	6.8	6.68
Occupancy $\mathcal{O}$(%)	67%	40%	*27%*	28%	21%	43%	24%	64%	41%	86%	56%
Waiting time t_w (h)	0.67	1.17	*-0.5*	0.28	1.01	0.45	0.4	0.6	1	0.9	1.7
Hold-back time t_h (h)	0.96	0.73	*0.23*	0.37	0.35	0.7	0.5	0.9	0.7	1.13	0.9
Hold-back per km	0.24	0.15	*0.09*	0.15	0.1	0.15	0.1	0.22	0.1	0.31	0.2
Commercial speed v_c (km/h)	4.47	5.02	*-0.55*	2.18	1.59	2.8	4	4.1	4.75	5.6	5.82

graphical results obtained from our field study of minibus taxis transport dynamics in Kampala as illustrated in the side-by-side comparison of both studies in the Figures 6.8 and Table 6.5. Thus, the controlled agent-based simulation experiment closely represents Kampala's minibus taxis' organically-evolved, quasi-demand-responsive paratransit system.

6.4.4 Identifying efficiency metrics

Conventionally, public transport efficiency was evaluated based on *mobility*. This assumes that the faster a transport system moves passengers, the better (Litman, 2012). Levine et al. (2012) introduced a new paradigm of evaluating transport efficiency based on *accessibility* – the people's ability to access the transport service. The quasi-demand-responsiveness, combined with the highly atomised ownership structure in Kampala's minibus taxi paratransit system made us introduce the *profitability* dimension of evaluating efficiency from the driver's perspective (see Chapter 3). In the profitability paradigm, we are concerned with the driver's ability to make profitable trips.

We categorised the journeys and trips metrics to select the appropriate metrics for evaluating the mobility efficiency, accessibility efficiency and profitability efficiency. Table 6.6 shows the categorisation of metrics. Thus, in response to the research question RQ4 (see Section 1.3.1 of Chapter 1), the metrics for measuring efficiency in an organically-evolved, quasi-demand-responsive paratransit system are: passenger waiting time t_w; minibus taxi hold-back time t_h; passengers' first leg distance d_{l1}; and minibus taxi occupancy $\mathcal{O}$. The secondary efficiency metrics were selected based on their correlation coefficient strength with the primary efficiency metrics (refer to Tables 6.3b and 6.4b for the correlation coefficient values).

Table 6.6: Categorising metrics by mobility efficiency, accessibility efficiency and profitability efficiency.

	Mobility	Accessibility	Profitability
Primary efficiency metrics	Waiting time (t_w), Hold-back time (t_h)	First leg distance (d_{l1}), Last leg distance ($d_l n$)	Occupancy ($\mathcal{O}$)
Secondary efficiency metrics	Legs count (l_{count}), Stops waited at (sw_{count}), Commercial speed (v_c)		Hold-back time (t_h), Commercial speed (v_c)

6.5 Summary

In this chapter, we set up a controlled agent-based simulation experiment of minibus taxi transportation dynamics in an organically-evolved paratransit system setting. We designed minibus taxi agents with limited situational awareness, quasi-demand-responsiveness, random passenger search, Lévy walk behaviour and occasional abandonment of trips that were considered 'unprofitable' (with persistently low passenger occupancy). The passenger agents were designed with limited situational awareness, considerable persistence (able to move from stop to stop waiting for a taxi), and limited memory – they depended on episodic memory to make some decisions such as where to wait for a taxi.

Analysis and validation of the simulation results indicated distribution statistically close to the distributions obtained from the field study in Kampala. We can therefore conclude that the agent-based simulation closely represents Kampala's organically-evolved, quasi-demand-responsive paratransit system. We further identified four primary metrics for evaluating the efficiency of a paratransit system. These included: the passenger waiting time t_w; minibus taxi hold-back time t_h; passengers' first leg distance d_{l1}; and minibus taxi occupancy $\mathcal{O}$. Thus, answering research questions RQ3 and RQ4, and achieving Objective 2.2 and Objective 2.3.

Part III:

Research Stage III

Chapter 7

Test experiments and optimisation results

Chapter 7 Objectives
This chapter aims to achieve the research Objective 3.1 of the dissertation to answer research question RQ5.
• ⇒**Research objective 2.2**

> Optimise selected efficiency metrics of Kampala's simulated minibus taxi transport system and evaluate the associated gain in system efficiency at a macro level.

This chapter presents two test experiments that we set up to optimise selected efficiency metrics identified in Section 6.4.4. To support more optimal decision making by agents, we modified the general utility scoring Equation 5.3 defined in Section 5.2.2 to include a *situational awareness* (SA) dimension Φ as described in Equation 7.1.

$$U_i = \left(\sum_{k=1}^{n} s_{ki} \times w_k \right) \times \Phi + \varepsilon_i \tag{7.1}$$

where, U_i is the utility associated with the i^{th} alternative, s_{ki} is the score associated with dimension k, w_k is the weight associated with dimension k, Φ is the situational awareness score, and ε is the random noise (random variable, $\mu = 0$, $\alpha = 0.05$) to represent bounded rationality.

Thus, given a situation in which three alternatives x, y and z vary along four dimensions D_1, D_2, D_3 and D_4, and their scores and weights along these dimensions are given, the utility payoff associated with each alternative is computed based on Equation 7.1 as illustrated by the payoff matrix in Table 7.1. The alternative with the maximum utility is selected. Tables 5.4

Table 7.1: Illustration of utility payoff determination by agents given three alternatives, each with four dimensions. Note: (i) $\sum_{d=1}^{4} w_{ad} = 1, for\ a \in \{x, y, z\}$; (ii) $s_{ad} \in (0, 1]$ for $a \in \{x, y, z\}$ and $d \in \{1, 2, 3, 4\}$; (iii) ε is random variable with $\mu = 0$; and $\alpha = 0.05$; (iv) Φ is the situational awareness score.

		Dimensions				SA (Φ)	Noise (ε)	Utility (U)
		D_1	D_2	D_3	D_4			
	x	$s_{x1}w_{x1}$	$s_{x2}w_{x2}$	$s_{x3}w_{x3}$	$s_{x4}w_{x4}$	Φ_x	ε_x	$(\sum_{d=1}^{4} s_{xd}w_{xd})\Phi_x + \varepsilon_x$
Alternatives	y	$s_{y1}w_{y1}$	$s_{y2}w_{y2}$	$s_{y3}w_{y3}$	$s_{y4}w_{y4}$	Φ_y	ε_y	$(\sum_{d=1}^{4} s_{yd}w_{yd})\Phi_y + \varepsilon_y$
	z	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	$s_{z1}w_{z1}$	Φ_z	ε_z	$(\sum_{d=1}^{4} s_{zd}w_{zd})\Phi_z + \varepsilon_z$

and 5.5 describe the dimensions, scores, and weights used by agents to determine the utility during decision making.

The main difference between the two experiments is the methods they use to determine and score the situational awareness dimension given several alternatives. The test experiment one (TOR1) uses the cognitive model of situational awareness integrated with the random forest classification algorithm (Hoogendoorn et al., 2011; Buitinck et al., 2013; Pedregosa et al., 2011). The test experiment two (TOR2) uses a neural network-based ranking mechanism to generate the situational awareness dimension (Pasumarthi et al., 2019). The setup and results from the two experiments are presented in the following sections.

7.1 Test experiment #1 (TOR 1)

In this experiment, the situation awareness dimension was implemented based on the cognitive model of situational awareness developed by Endsley (1995) and improved by Matthews et al. (2001), Hoogendoorn et al. (2011), and Bosse et al. (2012). We described the model in detail in Section 5.3.3.4 of this thesis. The model consists of four components: Perception of elements or cues in the environment; comprehension and integration of information about the current situation; projection of information for future events; and updating the mental model (Hoogendoorn et al., 2011).

7.1.1 Description of experiment set up

The test simulation experiment (TOR1) was designed to simulate minibus taxi transportation dynamics in a quasi-demand-responsive paratransit system. The simulation implements the agent-based model (ABM) developed in Chapter 5 with agents trained to adopt behaviour based on improved situational awareness. The passenger journey characteristics and inputs were maintained as described in Sections 6.3.1 and 6.4.1.1, respectively.

7.1.1.1 Learning the situation awareness score for TOR1

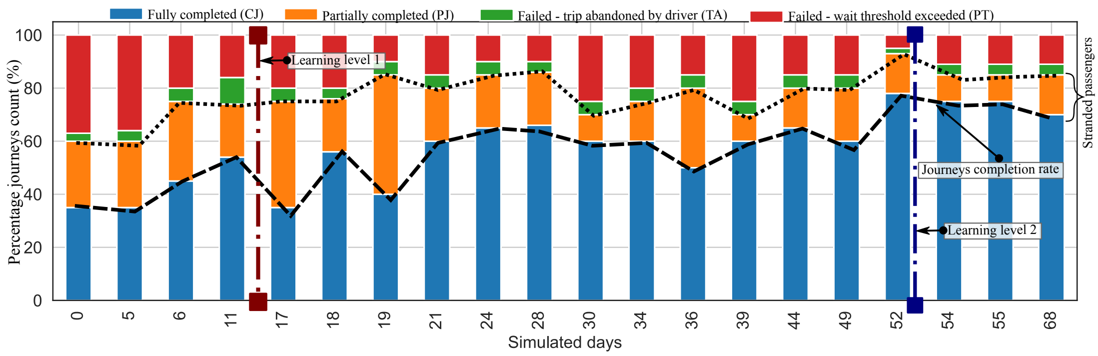
In this experiment setup, the agents can *observe* the current status of the world within a threshold distance. Hence, they can form a *belief* about the current situation. For example, a minibus taxi agent can view the demand and supply status of the stops two kilometres ahead of its current position. In addition to observing the current status, the agents can also infer *future beliefs* based on the belief value 'learned' cooperatively with other agents that executed a similar objective under considerably similar conditions. The *future belief value* in the simulation is associated with the stop and the target decision submodel to be executed, i.e., the initial stop model (ISM); the boarding choice model (BCM); the route choice model (RCM); and the passenger touting model (PTM) (see Sections 5.3.3.2 and 5.3.3.3). The future belief value is learned using a random forest classification algorithm that is trained with data from every ten-day window during the simulation runtime for each stop and each target decision sub model. The current belief and future belief are combined to form a relative score referred to as the *situational awareness score* Φ that is used to compute the utility associated with the particular alternative (refer to Equation 7.1 and Table 7.1).

7.1.2 Experiment #1 results

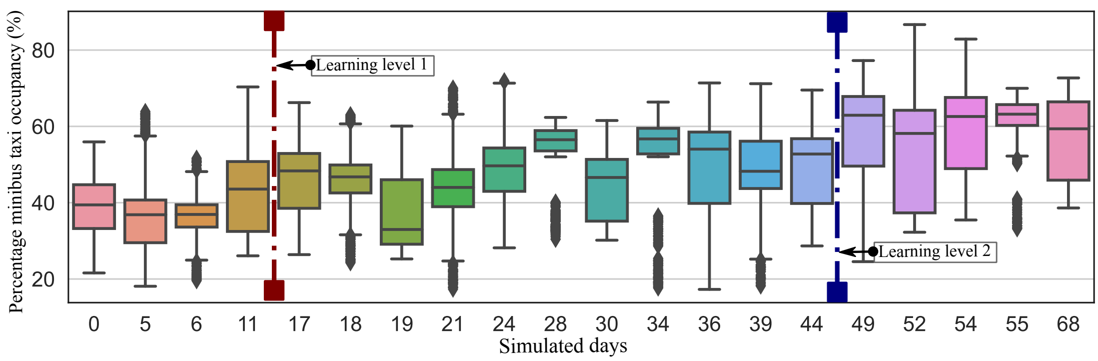
The results from test experiment TOR1 are presented in two categories. First, we present the general macro view of test experiment one results (see Figures 7.1). Second, we present a comparative analysis of the results for selected metrics for the controlled experiment (CER) results, test experiment one (TOR1), and test experiment two (TOR2) in Section 7.3 (see Figures 7.4, 7.5 and Table 7.2).

The results show two distinct learning levels (learning levels one and two), representing the significant points during the simulation runtime where the agents adopted significantly optimal behaviour, thus gaining a degree of objective execution efficiency (see Figure 7.1). In general, the journey completion rate for passenger agents in TOR1 improved from 40% before learning level one on day 17, to 70% after learning level two on day 52 (see Figure 7.1a). The rate of partially completed and failed journeys also reduced. Likewise, the average minibus taxi agent occupancy improved from 50% to 65% after learning levels one and two, respectively (see Figure 7.1b). Hence the number of passengers stranded in the system because of failure to get transport reduced, which is an indicator of system efficiency improvement.

More test experiment one results for selected metrics will be presented in the comparative analysis in Section 7.3



(a) Journeys completion rates during test simulation experiment 1 (TOR1).



(b) Minibus taxi occupancy during test simulation experiment 1 (TOR1)

Figure 7.1: Test experiment one variations of: (a) Passenger journeys completion rates; (b) Percentage minibus taxi occupancy.

7.2 Test experiment #2 (TOR 2)

In this experiment, we used a convolutional neural network (CNN)-based ranking mechanism to implement situational awareness. This was done by training a list-wise ranking submodel (in other words, learning-to-rank model) using Google's TensorFlow ranking (tf-ranking) application programming interface (API) (Pasumarthi et al., 2019). The tf-ranking API uses a multivariate scoring function and a listwise loss ranking function to order a set of multifeatured items. Equation 7.2 defines a listwise Softmax loss function (Ai et al., 2018), and Equation 7.4 defines the metric scoring and evaluation function – Normalised Discounted Cumulative Gain (NDCG) (Valizadegan et al., 2009; Qin et al., 2010).

$$\hat{\ell}(y,\hat{y}) = -\sum_{j=1}^{n} y_j log(\hat{y}_{jg}) = -\sum_{j=1}^{n} y_j log(\frac{exp(\hat{y}_j)}{\sum_{j=1}^{n} exp(\hat{y}_j)}) \tag{7.2}$$

where, $y_j \in \{0,1\}^k$ represents the lebel of the i^{th} sample in one-hot encoded representation and $\hat{y}_j \in [0,1]^k$ is the predicted probability with Softmax. $\hat{y}_{ig}$ represents the predicted probability of the ground-truth class for the i^{th} sample.

$$DCG(\pi f, y) = \sum_{j=1}^{n} \frac{2^{y_{j-1}}}{log_2(1+\pi f(j))} \tag{7.3}$$

where, πf is the ranking function, y is the dataset.

$$NDCG(\pi f, y) = \frac{DCG(\pi f, y)}{IDCG(y)} \tag{7.4}$$

where, $IDCG$ is the ideal DCG value of the best ranking function on the dataset y.

The goal of the learning-to-rank model is to learn a scoring function f such that, given a list of multi-features items, it produces an optimal ordering of the items in their order of relevance. So, the most optimal item (with higher utility) will be on top as illustrated in Figure 7.2a. Figure 7.2b illustrates how the neuro-network layers interact during groupwise multivariate scoring. Figure 7.2c summarises the five-step workflow used to build the ranking submodels for test experiment two. Algorithm 3 further describes a step-by-step process of building, training, and serialising the ranking neural network submodel for test experiment two (TOR2).

7.2.1 Description of experiment setup

The test simulation experiment (TOR2) was designed to simulate minibus taxi transportation dynamics in a quasi-demand-responsive paratransit system. The simulation implements the agent-based model (ABM) developed in Chapter 5 with agents trained to adopt behaviour based on optimised situational awareness. The passenger journey characteristics and inputs were maintained as described in Sections 6.3.1 and 6.4.1.1, respectively.

7.2.1.1 Optimising situation awareness of agents in TOR2

In this experiment setup, we modelled the passenger and minibus taxi agents' decisions made during the execution of the four submodels– ISM, BCM, RCM, and PTM– as ranking problems. For example, given a set of stops within threshold radius relative to its spatial position, the passenger agent's objective during ISM is to get the ordering of the stops such that the first leg distance d_{l1}, waiting time t_w, legs count l_{count} and last leg distance d_{ln} are optimised (see illustration in Figure 7.2a). Thus, we trained a CNN-based ranking model with three dense layers, an NDCG ranking evaluation metric, and a Softmax loss function. The ranking model was implemented using tf-ranking API form Google's TensorFlow 2.0 library. Figure 7.2c and Algorithm 3 describes the ranking model development and training workflow.

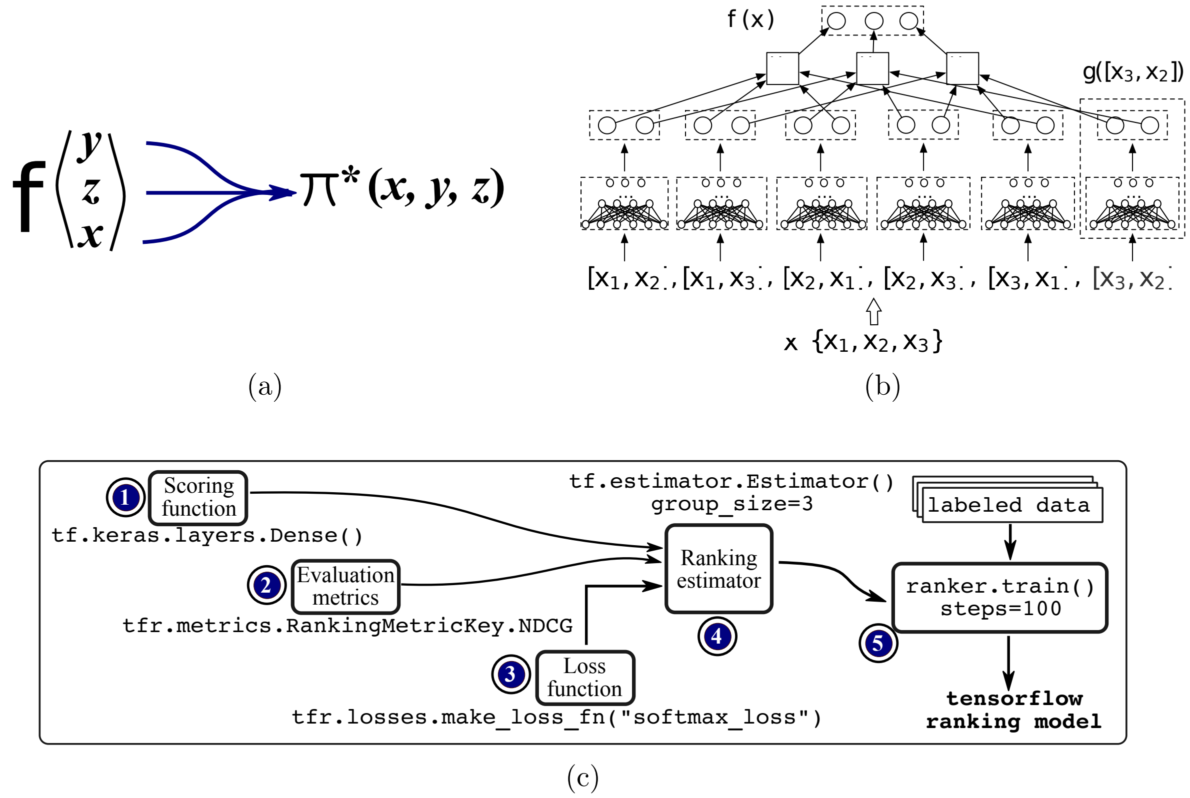


Figure 7.2: Illustrating the multi-item ranking process: (a) Function f takes a list of unordered items, and returns an optimal ordering π^*; (b) How groupwise multivariate scoring functions work (Ai et al., 2018); (c) Ranking model development workflow in tensorflow 2.0.

Once deployed, the model inputs a set of origin-destination parameter pairs and returns a ranked list according to relevance (or utility value). Depending on the decision submodel, the origin parameters include parameters related to the current status of the system, e.g., current demand, supply, occupancy status, waiting time, and hold-back time of active agents in the neighbourhood.

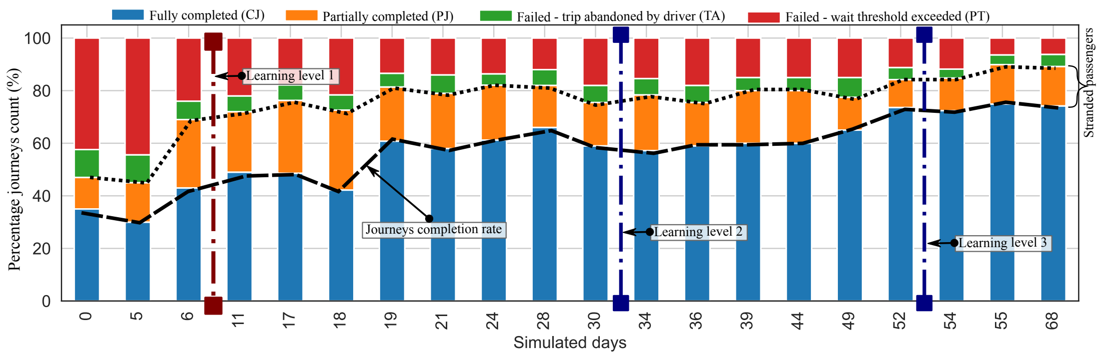
Algorithm 3: BUILDANDTRAINRANKINGMODEL Build and train a multi-item scoring ranking model

```
Data: paxJourneys, mbTrips, dataLabels          // executed journeys and trips labelled
foreach AGENT_DECISION SUBMODEL do
                                                   // i.e., ISM, BCM, RCM and PTM
    Specify a scoring function                     // (1)Three hidden dense layer scoring function.
    tf.keras.layers.Dense()
    Specify the evaluation metrics to optimise for  // (2) NDCG tfr.metrics.RankingMetricKey.NDCG
    Specify the loss function                      // (3) Softmax loss, tfr.losses.make_loss_fn("softmax_loss")
    Build the ranking estimator                    // (4) Build multi-item scoring ranking estimator
    tf.estimator.Estimator()
    Load data from a mongo database
    Extract features and labels associated with a submodel
    Train the ranking estimator                    // (5) ranker.train()
    Serialise and save ranking model
end
```

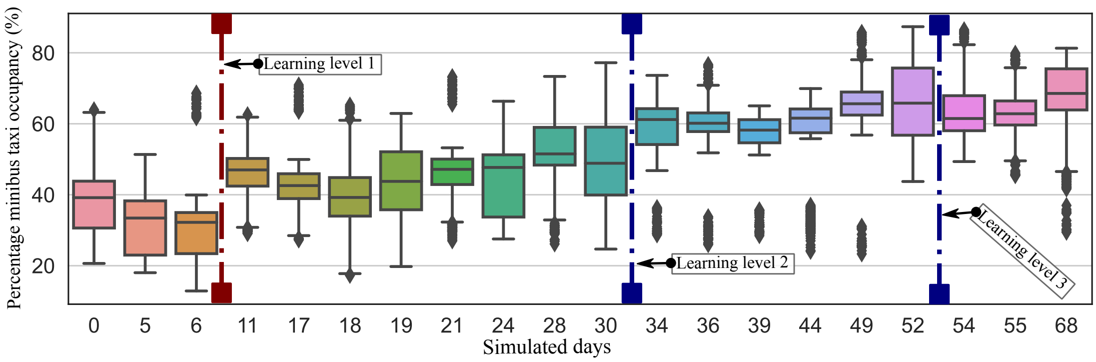
7.2.2 Experiment #2 results

The results from test experiment TOR2 are presented in two categories. First, we present the general macro view of the results (see Figures 7.3). Second, we present a comparative analysis of the results for selected metrics for the controlled experiment results (CER), test experiment one (TOR1), and test experiment two (TOR2) in Section 7.3 (see Figures 7.4, 7.5 and Table 7.2). The results show three distinct learning levels (learning level one, two, and three), representing the significant points during the simulation runtime where the agents adopted significantly optimal behaviour, thus gaining a degree of objective execution efficiency.

In general, the journey completion rate for passenger agents in TOR2 improved from 45% before learning level one on day 6, to 50% after learning level two on day 34, then to 75% on day 52 after learning level 3 (see Figure 7.3a). The rate of partially completed and failed journeys also reduced significantly between learning level two and three. Likewise, the average minibus taxi agent occupancy improved from 30% before learning level one, to 60% after learning level two, then to 70% after learning level three (see Figure 7.3b). Hence the number of passengers stranded in the system because of failure to get transport significantly reduced, which is an indicator of system efficiency improvement.



(a) Journeys completion rates during test simulation experiment 2 (TOR2).



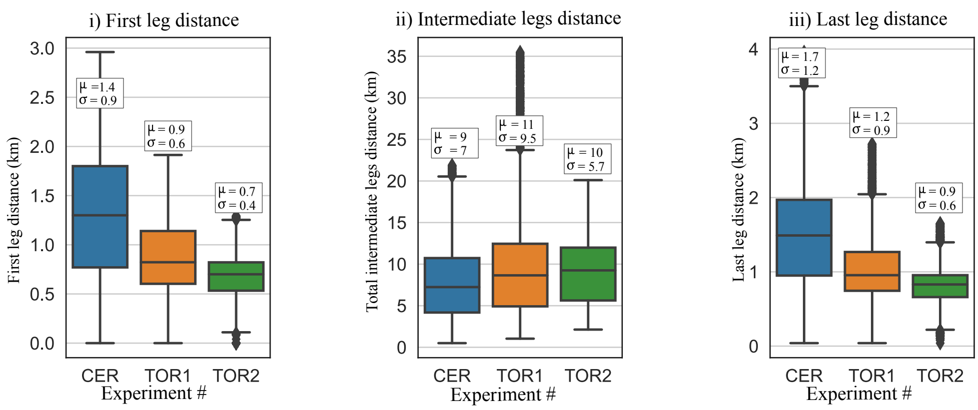
(b) Minibus taxi occupancy during test simulation experiment 2 (TOR2)

Figure 7.3: Test experiment two variations of: (a) Passenger journeys completion rates; (b) Percentage minibus taxi occupancy.

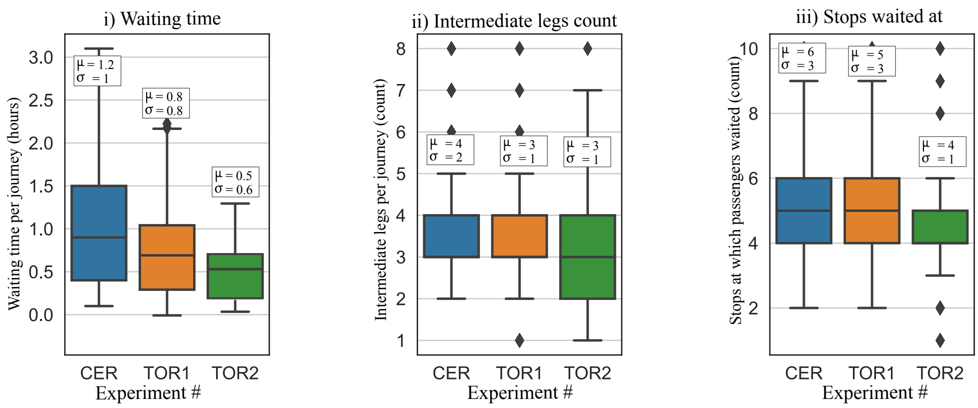
7.3 Comparative analysis of CER, TOR1 and TOR2 results

This section presents a side-by-side comparative analysis of results from three simulation experiments, i.e., the controlled simulation experiment (CER in Section 6.4), test experiment one (TOR1 in Section 7.1), and test experiment two (TOR2 in Section 7.2). The results represent aggregate macro-level measurements of selected metrics associated with passenger agents' journeys (i.e., the first leg distance d_{l1}, intermediate legs distance d_l, last leg distance d_{ln}, waiting time t_w, intermediate legs count l_{count}, and the count of stops where passengers waited during a single journey sw_{count}), and minibus taxi agents trips (the trip distance d_T, operating speed v_o, commercial speed v_c, hold-back time t_h, hold-back per km, and minibus taxi occupancy $\mathcal{O}$). Figures 7.4 and 7.5 and Tables 7.2a and 7.2b show the statistical analysis of the journeys and trips metrics, respectively.

Analysis of passenger agents journey data shows that there is a 33% and 49% reduction in the mean of first leg distances between the controlled simulation experiment, and test experiments TOR1 and TOR2, respectively (refer to Table 7.2a and Figure 7.4ai). The mean journeys last leg distances also reduced from 1.7 km in the CER, to 1.2 km in the TOR1, then to 0.9 km in the TOR2, representing a 31% and 46% decrease between CER, and test experiments TOR1



(a) Comparing first leg d_{l1}, intermediate legs d_l, and last leg d_{ln} distances.

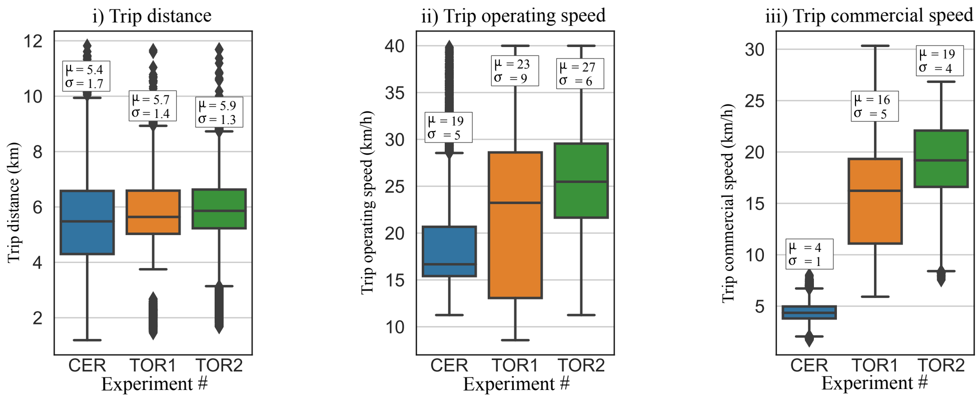


(b) Comparing waiting time, intermediate legs count and count of stops waited at.

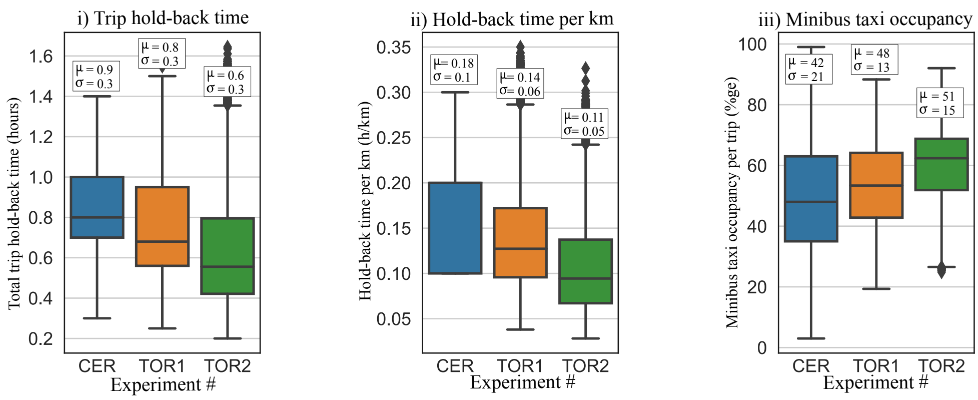
Figure 7.4: Passenger journeys metrics: Comparing controlled simulation experiment results (CER) with test optimisation results from two simulation experiments (TOR1 and TOR2).

and TOR2, respectively (Figure 7.4a*iii*). Furthermore, from CER, TOR1, and TOR2 results, there was a substantial reduction in the journeys' waiting time from 1.3 hours, to 0.8 hours and then to 0.5 hours (Figure 7.4b*i*); intermediate legs count from 6 legs, to 5 legs then to 4 legs (Figure 7.4b*ii*); and count of stops where passengers waited for minibus taxis, from 4 stops, to 3 stops, then to 3 stops (Figure 7.4b*iii*). Finally, there was an increase in the intermediate legs' distance from 9 km in CER, to 11 km in TOR1, to 10 km in TOR2 (Figure 7.4a*ii*). This represents a 23% and 15% increase in the intermediate legs' distance between CER and test experiments TOR1 and TOR2, respectively, as shown in Table 7.2a.

Analysis of the minibus taxi trips data from the simulation experiments shows an increase in operating speed, commercial speed, minibus taxi occupancy, and a reduction in hold-back time and hold-back time per km. Figure 7.5a*iii* shows an increase in the mean commercial speed from 4 km/h in CER to 16 km/h in TOR1, to 19 km/h in TOR2. There was a moderate reduction in hold-back time from 0.9 hours in CE, 0.8 hours in TOR1, to 0.6 hours in TOR2 (see Figure 7.5b*i*). The hold-back per kilometre significantly reduced from 0.18 h/km in CER to 0.14 h/km in TOR1, to 0.11 h/km in TOR2 (see Figure 7.5b*ii*). The mean minibus taxi occupancy increased from 42% in CER to 48% in TOR1, to 51% in TOR2 (see Figure 7.5b*iii*). This represents a 14% and 21% increase between CER and test experiments TOR1, and TOR2, respectively, as shown in Table 7.2b.



(a) Comparing trip distances, trip operating speed, and trip commercial speed.



(b) Comparing hold-back time, hold-back per kilometre and minibus taxi occupancy.

Figure 7.5: Minibus taxi trips metrics: Comparing controlled simulation experiment results (CER) with test optimisation results from two simulation experiments (TOR1 and TOR2).

Table 7.2: Summary statistics analysis and comparisons for passenger journeys and minibus taxi trips for the controlled and test simulation experiments. NOTE: CE–Controlled experiment, T1=TOR1 – Test simulation experiment 1, T2=TOR2 – Test simulation experiment 2.

(a) Comparing passenger journeys metrics' statistical values for mean, standard deviation, and percentage quantiles Q25, Q50, and Q75.

	Mean μ			%ge $\Delta\mu$ CE-		Std dev σ			Q 25%			Q 50%			Q 75%			Max		
	CE	T1	T2	CE-T1	CE-T2	CE	T1	T2	CE	T1	T2	CE	T1	T2	CE	T1	T2	CE	T1	T2
First leg distance (d_{l1})	1.4	0.9	0.7	33%	49%	0.9	0.6	0.4	0.8	0.6	0.6	1.3	0.9	0.7	1.9	1.2	0.9	7.1	6.2	5.7
Intermediate legs distance (d_l)	9	11	10	-23%	-15%	7	9.5	5.7	4.4	5.2	5.9	7.6	9.2	9.6	11	13	12	76	55	40
Last leg distance (d_{ln})	1.7	1.2	0.9	31%	46%	1.2	0.9	0.6	1	0.8	0.7	1.5	1	0.8	2.1	1.4	1	10.8	9.8	8.7
Waiting time (t_w)	1.2	0.8	0.5	34%	55%	1	0.8	0.6	0.4	0.3	0.1	1	0.7	0.5	1.7	1.1	0.8	9.8	9	7.5
Stops waited at (sw_{count})	4	3	3	21%	24%	2	1	1	3	3	2	4	3	3	4	4	4	17	15	11
Legs count (l_{count})	6	5	4	5%	27%	3	3	1	4	4	4	5	5	4	7	7	5	30	30	24

(b) Comparing minibus taxi trips metrics' statistical values for mean, standard deviation, and percentage quantiles Q25, Q50, and Q75.

	Mean μ			%ge $\Delta\mu$ CE-		Std dev σ			Q 25%			Q 50%			Q 75%		
	CE	T1	T2	-T1	-T2	CE	T1	T2	CE	T1	T2	CE	T1	T2	CE	T1	T2
Trip distance (d_T)	5.4	5.7	5.9	-6%	-9%	1.7	1.4	1.3	4.3	5	5.2	5.5	5.7	5.9	6.6	6.6	6.6
Operating speed (v_o)	19	23	27	-21%	-42%	5	9	6	15	13	22	17	24	26	21	30	30
Commercial speed (v_c)	4	16	19	-300%	-375%	1	5	4	4	11	17	4	16	19	5	19	22
Hold-back time (t_h)	0.9	0.8	0.6	11%	33%	0.3	0.3	0.3	0.7	0.6	0.4	0.8	0.7	0.5	1.1	1	0.8
Hold-back per km (t_h/km)	0.18	0.14	0.11	22%	39%	0.1	0.06	0.05	0.1	0.1	0.07	0.2	0.13	0.09	0.2	0.17	0.14
Occupancy ($\mathcal{O}$)	42%	48%	51%	-14%	-21%	21%	13%	15%	25%	38%	40%	42%	47%	49%	57%	58%	63%

7.4 Summary

In this chapter we set up two test agent-based simulation experiments of minibus taxi transportation dynamics in an organically-evolved paratransit system setting. In test experiment one (TOR1), we improved the agents' decision making based on a situation awareness dimension. The passengers and minibus taxi agents observe the status of the world, they form their current and future beliefs based on a supervised learning algorithm (Random forest). The supervised learning algorithm generates situational awareness scores that are used to evaluate the alternative with high utility. Test experiment two further improved the agents' decision making and situational awareness based on a deep learning method, i.e., convolutional neural network (CNN). The CNN was trained to optimally rank (or order) a set of choices from which an agent must choose, such that the option with the highest utility is on top. Thus, the agent chooses one with higher utility.

Results from optimising and analysing selected metrics from the two test experiments (TOR1 and TOR2) indicate significant improvement in minibus taxi transport system efficiency at macro-level. Thus, we have answered research questions RQ5, and achieved Objective 3.1 of this dissertation.

Chapter 8

Discussion

This chapter is presented in two contexts: First, the research questions identified in Chapter 1 are revisited and discussed in the light of accepted knowledge and the results obtained. Second, a general discussion of the methods and outcomes of the dissertation is provided in a broader perspective, and in relation with the stated hypotheses.

The contents of this chapter are organised into four broad themes that are logically presented to provide the synthesis of the whole dissertation. The first theme is about the *operations*. The operations aspects of minibus taxis as presented in the dissertation are summarised and discussed. The second theme deals with *efficiency*: certain aspects of the paratransit efficiency are revisited and broadly discussed. The third theme looks at the paratransit modelling and *complex adaptive system*. Here we justify the reasons for modelling minibus taxi paratransit as a complex adaptive system. The final theme focuses on *paratransit efficiency improvement*. Here we discuss how we used distributed intelligence and situational awareness to improve the simulated paratransit system efficiency. The relationships between the selected themes, the research questions and the research hypotheses are also clearly explained.

8.1 Minibus taxi operations in a paratransit system

- ⇒**Research Question 1**

> How do minibus taxis operate in organically-evolved, quasi-demand-responsive paratransit systems?

In response to RQ1, we formulated research Objectives 1.1 and 1.3, and we carried out two independent studies described in Chapters 3 and 4. As mentioned earlier, part of each chapter addresses the Objectives 1.1 and 1.2, respectively.

The operational aspects of minibus taxis in organically evolved paratransit systems studied and presented in this dissertation broadly encompasses four main facets. These are: minibus taxi management; routes; passenger search strategies; and movement characteristics. The paratransit system operates an owner-driver model characterised by extreme ownership fragmentation with little or *no centralised management*. In this system, several low-occupancy vehicles (often 14-seater minibuses) privately owned (by a single or a few individuals) operate semi-autonomously to fulfil the mobility needs of the population. Various drivers and owners organise themselves into associations for purposes of licensing and self-regulation, though the few existing regulations are seldom adhered to by the drivers, and only loosely enforced by the authorities.

The minibus taxi routes, and their associated stops are not clearly established and labelled. They often vary according to demand, traffic conditions, competition, and drivers' preference.

Though there were previous attempts to map paratransit (minibus) routes in Kampala (Ndibatya et al., 2016), our results (presented in Chapter 3) give us reason to believe that paratransit routes *evolve*. For instance, KCCA developed and published a static paratransit route map for Kampala (Ndibatya and Booysen, 2020a). However, we found evidence of new routes and stops in addition to the absence of several routes and stops that appeared on KCCA's static route map. The noticeable difference in the routes' trajectory profiles (measured using spatial distance, see Table 4.1a, p. 54) further suggests route evolution. Broadly, this indicates: (1) the *quasi-demand-responsiveness* of paratransit, i.e., the routes' profiles and stops locations might have shifted due to change in demand in those areas; (2) the *self-organizing* nature of paratransit; and (3) the *adaptation* and *emergence* of new behaviour among the paratransit systems.

Minibus taxis in Kampala use three main strategies when searching for passengers to maintain a profitable business. First is the *random passenger search*, where the driver starts a trip with a few passengers anticipating finding more passengers en-route to fill up the taxi. This often leads to losses by the drivers when they do not get the anticipated passengers en-route. Second is the *random back-off* or *holding-back*, where the driver interrupts the trip for a random period to allow for passenger demand replenishment on the route before proceeding with the trip. This method often works to the disadvantage of the passengers in the taxi. It dramatically increases the passenger total travel time if the taxi they boarded frequently "holds back" during the trip, and it leads to the general system's inefficiency. The third is the *trip abandonment*, where the trips deemed unprofitable by the drivers are either abandoned, or the trip routes are changed to new destinations where drivers anticipate high demand. When trip abandonment occurs, passengers disembark, and they wait for connecting trips to their destinations. This strategy also disadvantages the passengers. Because the total waiting time and total "legs" required to complete a single journey also increase, there is an increase in the total travel time. Sometimes, passengers do not get connecting trips and the drop the journeys, which affects the general efficiency of the paratransit transport system.

We discovered that during searching for, picking up and transporting passengers, Kampala's minibus taxis adopt a scale-invariant super diffusive movement pattern where the taxi trajectory steps follow a heavy-tailed power-law distribution similar to the "Lévy walk" pattern defined by $f(x) \sim l^{-\alpha}$ where l is the step length, and α (referred to as the Lévy exponent) is in the range $1 < \alpha < 3$. In the reviewed literature, the Lévy walk strategy was found to optimise random searches for randomly and patchily distributed replenishable resources among cognitively complex organisms (Reynolds, 2015; Raichlen et al., 2014). As the Lévy exponent values get closer to two ($\alpha \approx 2$), the search strategy becomes optimal (Viswanathan et al., 2011; Reynolds, 2018). Our Lévy walk strategy investigation of minibus taxi movements revealed that though the Lévy walk search strategy was confirmed in the sample of minibus taxi trajectories we analysed (i.e., $1 < \alpha < 3$), only 22.2% had the α values close optimal ($\alpha \approx 2$). The others were slightly away from the optimal (refer to Table 4.1b, p. 54). This could be an indication of generally ineffective search strategies by minibus taxi drivers that often resulted in the system inefficiency discussed in the following section.

8.2 Minibus taxi paratransit system efficiency

We looked at efficiency from two perspectives – the passengers', and the drivers' perspectives. The passengers were concerned with executing the planned journeys and getting to their destinations within the least possible time and using the least possible effort. The drivers, on the other hand, were concerned with making the trips as profitable as possible.

- ⇒**Research Question 2**

  Are the paratransit operations efficient?

- ⇒**Research Question 4**

  What metrics can be used to measure efficiency in such a paratransit system?

In response to RQ2, we formulated research Objective 1.2, that we achieved in the two independent studies described in Chapters 3 and 4. In response to RQ4, we formulated research Objective 2.3 that we achieved in Chapter 6.

With respect to the passengers, we used two metrics to estimate the system efficiency, i.e., the passenger waiting time and the minibus taxi hold-back time. Results show that travel by minibus taxi is inefficient for the passengers. It is characterised by a long waiting time (22 to 59 minutes), and a long hold-back time (35 to 110 minutes). This means that passengers waste a lot of time either waiting for a taxi or seated in a stationary taxi that is holding back en-route to fill up with more passengers before proceeding with the trip. Figures 3.4b*vi* and 3.4b*vii* on page 40, respectively, show the distributions of waiting time and hold-back time recorded during the field study period.

With respect to the minibus taxis, we used the taxi commercial speed and profitability index to estimate the efficiency. The observed commercial speed of minibus taxis in Kampala is low (3.1 to 15.4km/h) the profitability index is 0.76, indicating low profitability. The low commercial speeds are partially because of high hold-back time which in turn affects the number of trips executed per day and hence the low profitability index. Refer to Figure 3.4b*x* (p.40) and Table 3.3 (p.38) for the commercial speed distribution and profitability index estimation, respectively. To break even, drivers often work for long hours (over 15 hours) and sometimes overload the taxis. Figures 3.4b*v* page 40 shows the percentage distribution of minibus taxi occupancy of the selected minibus taxis during the field study period. Note the occurrence of overloading (above 100% occupancy).

In response to RQ4, after modelling and simulating Kampala's minibus taxi transport dynamics (see Chapter 6), we identified five primary metrics for evaluating the efficiency of minibus taxi transportation in a paratransit system. The five metrics were broadly categorised into three, i.e., mobility, accessibility, and profitability, as shown in Table 6.6. The mobility-related efficiency metrics evaluated how fast a paratransit system moves passengers. They included *waiting time*, *hold-back time*, and *commercial speed*. The accessibility-related metrics evaluated how accessible a paratransit system is from the origin and the final destination, and they include: *first leg distance* and *last leg distance*. The profitability-related metrics evaluate how profitable the paratransit system is to the driver, and they include minibus taxi *occupancy*. The waiting time, first leg distance and last leg distance, are all related to the passenger journey, while the hold-back time, occupancy and commercial time are related to the minibus taxi trips. Accordingly, the metrics are used during optimisation to improve the efficiency of the journey and trips executed by passengers and minibus taxis, respectively.

8.3 Simulation and validation of paratransit as a complex adaptive system

In this subsection, we discuss the intrinsic characteristics of paratransit and, in particular, minibus taxi transportation in relation to the complex adaptive systems theory. We then align our findings in Chapters 5 and 6 with the general emergentists' school of thought that we introduced in Sections 2.1 (pg. 13) and 2.6 (pg. 24). We also discuss the features of paratransit that are consistent with the complex adaptive systems, thus justifying our reason for using agent-based modelling to study Kampala's minibus taxi system. We conclude this subsection by discussing the observed emergent properties from the agent-based model simulation and the macro-level analysis results for the controlled experiment (CER) presented earlier in Chapter 6. First, we re-echo the research question we answered, and the hypothesis examined by this subsection.

• ⇒**Research Question 3**

> How do individual-level operations and autonomous interactions between minibus taxis and passengers shape the higher-level (macro-level) system behaviour in an organically evolved, quasi demand-responsive paratransit system?

In response to RQ3, we formulated research Objectives 2.1, and 2.2, that we achieved in Chapters 5 and 6.

⊙ ⇒***Hypothesis 1***

> The transportation dynamics of organically evolved paratransit systems in Sub-Saharan Africa are shaped by local interactions of autonomous agents at micro-level of the system giving rise to a stable (often inefficient) state at macro-level through demand and supply.

Earlier in Section 8.1, we discussed five of the minibus taxi operational characteristics that we discovered from our field research in Kampala's paratransit system. The first is *vehicle ownership fragmentation*. This creates a large collection of small, competing entities in the paratransit system. The second is the *absence of centralised management*: this allows for and enables the fragmented entities to operate semi-autonomously. We refer to them as semi-autonomous because we are cognizant of the role played by the few paratransit associations such as UTODA and KOTISA that may limit their full autonomy to some extent. The third is *self-organisation*. Despite the loose regulations, lack of centralised management, and lack of a scheduling system, paratransit has a unique way of responding to demand. It often fulfils the mobility needs of the population, despite the inefficiencies associated with it. The fourth characteristic is *evolution* and *adaptation*. As observed in the paratransit routes and stops (see Chapter 3), unprofitable routes are believed to have been phased out in preference to new more profitable routes and new passenger demand dynamics. The fifth characteristic is the *emergence* of new behaviour because of micro-level interactions. We discovered a new movement behaviour that emerged (the "Lévy walk"). The *Lévy walk* behaviour (discussed in-depth in Chapter 4) could have emerged due to the minibus taxi drivers' desire to optimise their search. It is also worth noting that the Lévy walk behaviour is observable only at larger scales (macro-level). It may not be predicted by observing individual minibus taxi's trajectories at lower scales (micro-level) such as one minibus taxi's trip trajectory. The emergent behaviour at larger scales that may not be observable from individual components at lower scales is a key feature of complex adaptive systems. Thus, following Abbott and Hadžikadić's definition of a complex adaptive system, we concluded that minibus taxi transportation dynamics in Kampala's paratransit system exhibited characteristics that were closely related to those of a complex adaptive system (Abbott and Hadžikadić, 2017).

Abbott and Hadžikadić (2017) defined a CAS as "*a system composed of a large number of independent simple components that locally interact in an independent and non-linear fashion, exhibit self-organisation through interactions that are neither completely random nor completely regular and are not influenced by some central or global mechanism and yield emergent behaviour at large scales that are not predictable from observation of the behaviour of the components*"

Having established (with some level of certainty) that Kampala's minibus taxi transportation system is closely related to a complex adaptive system, we designed an agent-based model (ABM) of Kampala's minibus taxi paratransit system (see Chapter 5). Blume (2015) defined an ABM as a computational instantiation of a complex adaptive system. Therefore, we set up the minibus taxi ABM to overcome the sample-size limitation and huge costs associated with field data collection. We sought to take advantage of the available memory and computing resources to scale up and further study the minibus taxi paratransit transportation dynamics in a multi-agent simulated environment.

In Chapter 6 we set up a controlled ABM simulation experiment (CER) which replicates Kampala's minibus taxi transportation system as closely as possible by iteratively tuning the selected simulated system properties to match those observed in Kampala's paratransit system during the field study. During the simulation runtime, there were two main active and co-dependent entities i.e., the journey (managed and executed by the passenger agent), and the trip (managed and executed by the minibus taxi agent). We then measured several metrics' values associated with the journeys and trips. The results from running the controlled ABM simulation experiment (CER) are presented in Section 6.4.2 (pg 88). Tables 6.1 (pg. 84), 6.3 (pg. 90), and 6.4 (pg. 92), respectively, show the description of metrics used, statistical results of journeys' metrics value, and statistical results of trips' metrics value.

Furthermore, the simulation results, showed some levels of self-organisation. Through local agents' interactions, without external influence, the rates of journey completion were as follows: 60% on day one to a stable 80% on day 68 (see Figure 6.4a). This may have been due to self-organisation within the system, resulting in a slightly better state. Emergent properties such as *first leg distance*, *last leg distance* and *intermediate legs per journey* were also measured (see Figures 6.4b*i* pg. 89, 6.4b*iii* pg. 89, and 6.4c*ii* pg. 89, respectively). These properties could hardly be studied from the field. The figures show the macro-level analysis selected metrics.

8.4 Improving paratransit efficiency through distributed intelligence

This subsection presents our attempts to improve paratransit efficiency in a simulated environment. We based our experiments on complex adaptive systems (CAS) theory concepts. These concepts include: *self-organisation*, where a seemingly "chaotic" or disorganised system at micro-level achieves some level of order at the macro-level as a result of micro-level agents interactions; *emergency*, where new systems behaviours emerge at macro-level; and *adaptation*, where agents in the system adopt new and better emerging behaviour. The cycle continues, and as observed by Odell (2002), what emerges is always better than the sum of individual components. We sought to influence more optimal/improved macro-level behaviour by improving the *intelligence* and *situational awareness* of agents at a micro-level in the simulated paratransit system. We believed that *distributed intelligence* among agents would have a positive effect on the system *efficiency*. First, the research question and the hypothesis associated with this section are re-echoed below.

• ⇒**Research Question 5**

What is the macro-level effect on system efficiency of intelligent routing of autonomous and situationally aware minibus taxi agents with self-selected origins and destinations in an organically evolved, quasi demand-responsive paratransit system?

• ⇒***Hypothesis 2***

Improving the intelligence and situation awareness of autonomous agents in organically evolved paratransit systems leads to agents adapting more optimal travel behaviour resulting in improved macro-level paratransit system efficiency.

In response to RQ5, we formulated the research objectives 3.1 and 3.2 that we achieved in Chapters 7 and 9, respectively.

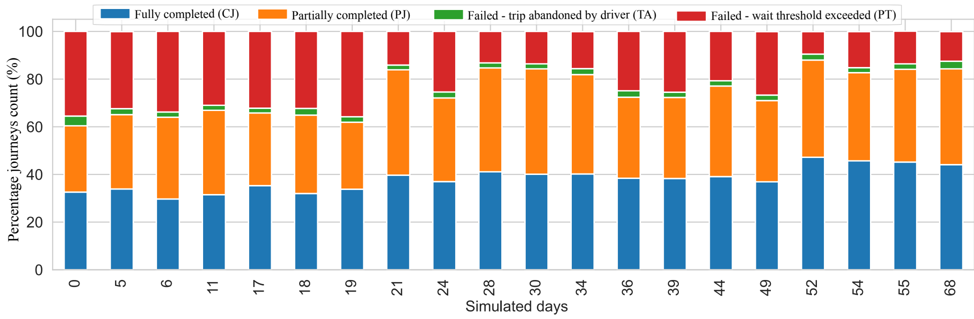
In Chapter 7, we set up two test experiments: test experiment one (TOR1) and test experiment two (TOR2). The test experiments used two closely related supervised machine learning methods (random forests and convolutional neural networks) to train agents at the micro-level to make more intelligent decisions based on previous experiences of self, and of others. Selected efficiency metrics' data is stored at the end of each journey and trip. Environmental status data – such as zone-based demand and supply – is also stored at different predefined intervals.

The comparative statistical analysis of macro-level efficiency metrics' values recorded during the test simulations runtime indicated a positive gain in the overall system efficiency as a result of improved micro-level agents intelligence and situational awareness (refer to Section 7.3 pg. 103). The Figures 8.1 and 8.2 (extracted from Chapter 7, pages 99 and 102) show a side-by-side comparison of the daily variations of the passenger journeys completion rates and the minibus taxi occupancy, respectively. From the figures, we identified the points in time where the simulated systems showed significant changes in states. We referred to such points as "learning levels". Alternatively, we can think about them as points of significant self-organisation within the simulated systems. During TOR1, we observed two such points (learning level 1, and learning level 2) (see Figures 8.1b and 8.1c). During TOR2, we observed three such points, i.e., learning levels one, two, and three (see Figures 8.2b and 8.2c). We drew conclusions about efficiency based on the metrics' results collected during the simulations runtime after the fiftieth simulation day, because we believed, at that time, that the system had attained a certain level of self-organisation.

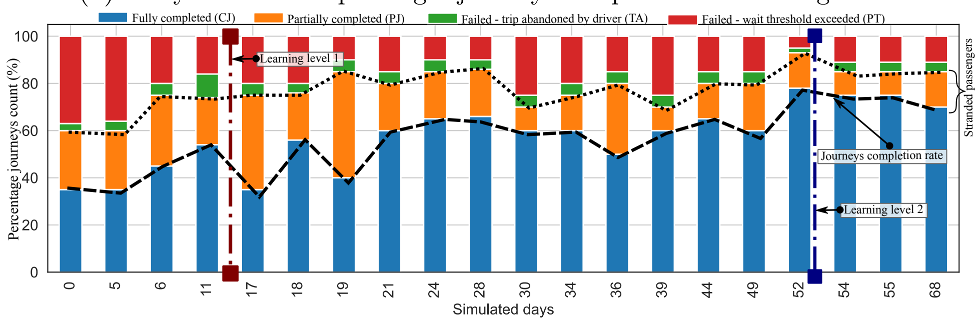
From Figure 8.1, the percentage of fully completed journeys increased from approximately

40% during the controlled experiment CER, to 60% during the test experiment TOR1, then to 75% during test experiment TOR2 as shown in Figures 8.1a, 8.1b, and 8.1c, respectively. There is also an observed decrease in the percentage number of partially completed and failed journeys from CER, to TOR1, then to TOR2. Correspondingly, there was an improvement in the percentage minibus taxi occupancy from 60% in CER, to 65% in TOR1, to 65% in TOR2, as shown in the Figures 8.2a, 8.2b, and 8.2c, respectively. The observed increase in percentage passenger journeys' completion, the corresponding reduction in partial and failed journeys, and the increase in minibus taxi occupancy are all indicators of efficiency improvement in the simulated paratransit system.

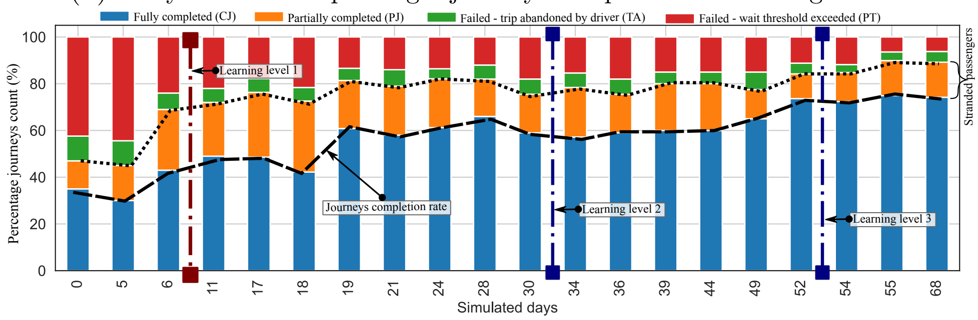
In Chapter 6 we identified and categorised paratransit efficiency metrics according to *accessibility*, *mobility*, and *profitability* (see Section 6.4.4 pg. 94). Subsequently, in Chapter 7, the test experiments' objectives were geared towards optimising these efficiency metrics at the system's micro-level with the hope of achieving macro-level efficiency globally. In Table 8.1 we presented a partial extract of the statistical analysis results presented in Chapter 7 (Table 7.2 pg. 105). The results (in Table 8.1) also strongly suggest that there was macro-level efficiency gain in the



(a) Daily variation of passenger journeys completion rate during CER.



(b) Daily variation of passenger journeys completion rate during TOR1.



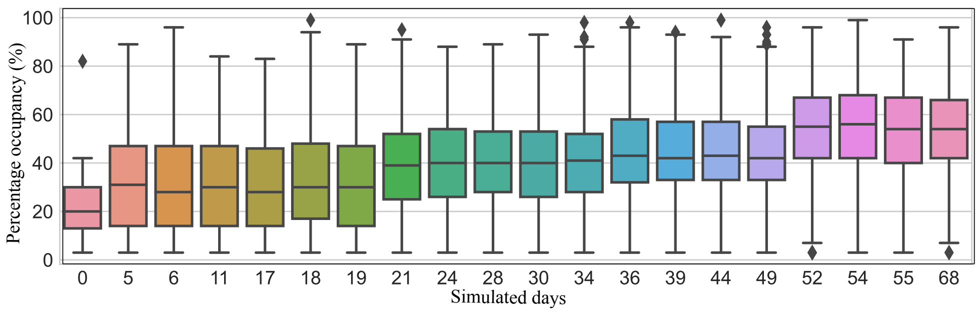
(c) Daily variation of passenger journeys completion rate during TOR2.

Figure 8.1: Comparing the variation of passenger journeys completion rates for three test experiments: (a) controlled simulation experiment (CER); (b) test experiment one (TOR1); and test experiment two (TOR2).

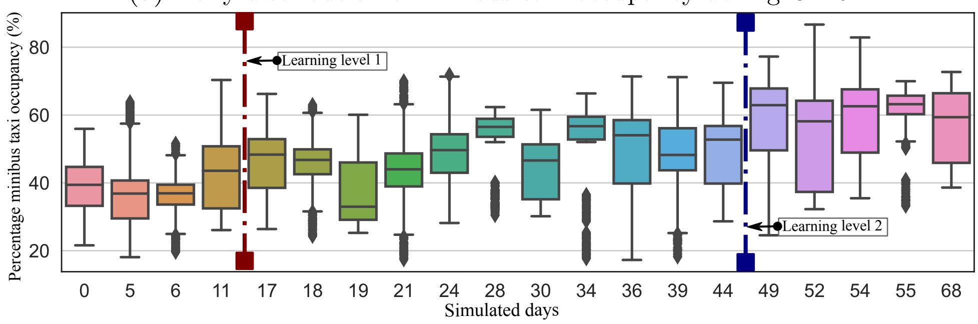
simulated paratransit system. This is because of the overall improvements observed in all the efficiency metrics values of the two test experiments (TOR1 and TOR2).

The mean *first leg distance* was reduced by 33% and 49% for CER-to-TOR1, and CER-to-TOR2, respectively (see Table 8.1). The mean *last leg distance* was reduced by 31% and 46% for CER-to-TOR1, and CER-to-TOR2, respectively. This means that passengers in TOR2 boarded minibus taxis at places closer to their journey origins, and they were dropped off at places closer to their final destination. Hence, their overall *accessibility* to the paratransit system improved.

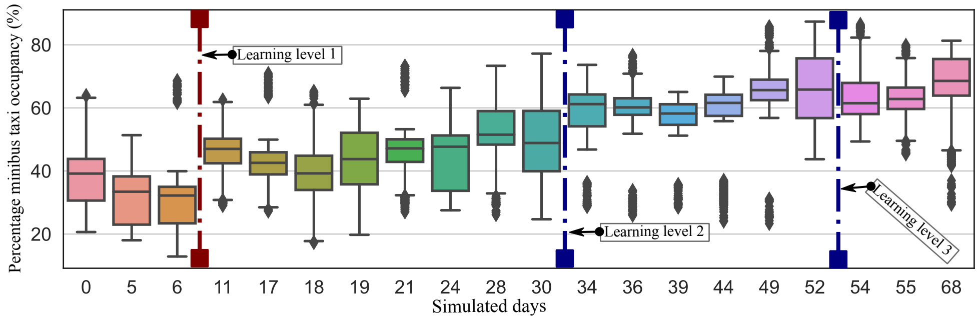
For a single journey, the mean *passenger waiting time* and the mean *count of stops* where passengers waited before getting a taxi reduced by 34% and 55%, respectively, for CER-to-TOR1 and CER-to-TOR2 (see Table 8.1). This signifies the improved ease with which the trained intelligent passenger agents waited for, found, and boarded minibus taxis. During the trip, the *hold-back per kilometre* improved by 22% and 39% for CER-to-TOR1 and CER-to-TOR2, respectively, whereas, the *journey legs count* reduced by 5% and 27% for CER-to-TOR1, and CER-to-TOR2, respectively. The reduction in "hold-back time per kilometre", and "journey legs count" means that the journeys' total travel time reduced. The percentage minibus taxi



(a) Daily distribution of minibus taxi occupancy during CER.



(b) Daily distribution of minibus taxi occupancy during TOR1.



(c) Daily distribution of minibus taxi occupancy during TOR2.

Figure 8.2: Comparing the distributions of daily minibus taxi occupancy for three test experiments: (a) controlled simulation experiment (CER); (b) test experiment one (TOR1); and test experiment two (TOR2).

occupancy increased by 14% and 21% for CER-to-TOR1 and CER-to-TOR2, respectively. Generally, the improvements in all the metrics discussed in this paragraph contribute to the general improvement in the overall *mobility* and *profitability* in the simulated paratransit system, hence the resulting system *efficiency* improvement at macro-level.

From the discussion above, and the results presented (e.g., in Figures 8.1, and Figures 8.2, and Table 8.1), we can draw three main conclusions about the simulated paratransit system. First, the paratransit system efficiency in TOR1 and TOR2 improved compared to CER. Secondly, the efficiency improvement observed could be as a result of improved agents' intelligence at micro-level in addition to improved situational awareness. Third, the simulated paratransit system underwent self-organisation as the individual agents learned and adopted new behaviour through micro-level interactions.

Our second hypothesis thus has been proved correct, i.e., that improving the intelligence and situational awareness of agents at a micro-level in a paratransit system gives rise to improved efficiency at a macro-level through self-organisation and adaptation. Accordingly, we have answered the research question RQ5.

Table 8.1: Comparing the three experiments' statistical summary of primary macro-level paratransit efficiency metrics values and the associated categories. The $\Delta\mu$ column shows the percentage change in mean values between experiments.

	Mean μ			%ge $\Delta\mu$	
	CER	TOR1	TOR2	CER-TOR1	CE-TOR2
1. Accessibility					
$\mapsto$ First leg distance d_{l1} (km)	1.4	0.9	0.7	33%	49%
$\mapsto$ Last leg distance d_{ln} (km)	1.7	1.2	0.9	31%	46%
2. Mobility					
$\mapsto$ Waiting time t_w (h)	1.2	0.8	0.5	34%	55%
$\mapsto$ Stops waited at sw_{count}	4	3	3	21%	24%
$\mapsto$ Hold-back time t_h (h)	0.9	0.8	0.6	11%	33%
$\mapsto$ Hold-back per km t_h/km (h/km)	0.18	0.14	0.11	22%	39%
$\mapsto$ Legs count l_{count}	6	5	4	05%	27%
3. Profitability					
$\mapsto$ Occupancy $\mathcal{O}$ (%)	42%	48%	51%	-14%	-21%

8.5 Summary

In this chapter, we have discussed the general aspects of minibus taxi operations, efficiency, complex behaviour, and efficiency improvement through distributed intelligence.

Chapter 9

Conclusion and recommendations

Through theoretical modelling, a field study and simulation-based experimental approaches, this study aimed to improve the efficiency of minibus taxis transportation in organically-evolved paratransit systems. The theoretical modelling work involved modelling paratransit systems as complex adaptive systems (CAS) and developing an agent-based model (ABM) for minibus taxi operations in an organically-evolved paratransit setting. The field study involved in-depth investigation of minibus taxi operations in Kampala's paratransit system, and collection and analysis of minibus taxi movement data that was used to validate the agent-based model. The experimental approaches involved three separate simulation experiments, simulating the minibus taxi transportation dynamics with varying levels of agents' intelligence and situational awareness. The results from the experiments showed that improving the micro-level agents' intelligence and situational awareness improved the overall efficiency of the minibus taxi paratransit system.

9.1 Insight into research findings and conclusions

Inefficiency has been identified as the major problem affecting organically-evolved paratransit systems in Kampala and other developing cities of the Global South. Minibus taxis contribute above 70% of the paratransit trips. We studied their efficiency from two perspectives, namely, the passengers' and drivers' perspectives. When interacting with the paratransit system, passengers execute "journeys", whereas minibus taxis execute "trips". In this dissertation, we identified and measured key efficiency metrics associated with passenger journeys (e.g., first leg distance, last leg distance and waiting time); and those associated with minibus taxi trips (i.e., hold-back time and occupancy). The results, from both the quantitative field study and the controlled simulation experiment, indicate that the journeys and trips were inefficiently executed. The mean values for the journeys' first leg distance, last leg distance, and waiting time were 1.4 km, 1.7 km, and 1.2 hours, respectively. The mean values for the minibus taxi trips' hold-back time and occupancy were 0.9 hours and 42%, respectively. Most existing paratransit studies are general, qualitative, and focus mainly on paratransit regulations and reforms. The results in this dissertation, however, provide new quantitative insight into paratransit efficiency and the associated efficiency metrics.

It was established that the minibus taxi operations within the paratransit system were complex and ineffective. The operational aspect of the minibus taxi system (studied during the field research) broadly encompassed four main facets, namely, minibus taxi management, routes, passenger search strategies, and movement characteristics. We made four major findings respectively-related to the four operational facets mentioned earlier. First, the minibus taxis paratransit system in Kampala *organically emerged* without prior planning. The vehicles' ownership is *fragmented* across many competing entities, with *no centralised management*. There are no streamlined booking and scheduling mechanisms, in addition to weak and *loosely enforced regulations*. In other words, they are *semi-autonomous*. Second, the minibus taxi routes and their associated stops are not clearly established and labelled. They often vary according to demand, traffic conditions, competition, and drivers' preference. The results from routes' profiles analysis give us reason to believe that the paratransit *routes and stops evolve*. Third, when searching for passengers, minibus taxis use three main strategies. The strategies include: *random passenger search*, where the driver starts a trip with few passengers anticipating to find more passengers en-route; *random back-off* or *holding-back*, where the driver interrupts the trip for a random period to allow for passenger demand replenishment on the route before proceeding with the trip; *trip abandonment*, where the trips deemed unprofitable by the drivers are either abandoned, or the trip routes are changed to new destinations. Fourth, it was further discovered that, during searching for, picking up and transporting passengers, minibus taxis *adopt* movement patterns where many short inter-stop distances (steps) are interspersed with long steps. This pattern often follows a heavy-tailed power-law distribution similar to the "Lévy walk" pattern defined by $f(x) \sim l^{-\alpha}$ where l is the step length, and α (referred to as the Lévy exponent) is in the range $1 < \alpha < 3$.

We draw two significant conclusions from the minibus taxis operations described in the previous paragraph. First, the minibus taxi paratransit system operations exhibit characteristics that are closely related to those of a complex adaptive system (CAS). We base this finding on Abbott and Hadžikadić's definition of a complex adaptive system (Abbott and Hadžikadić, 2017). Second, based on quantitative results discussed earlier, the operational characteristics, which include passenger search strategies, and movement characteristics are ineffective and thus, they contribute to the overall minibus taxi paratransit system inefficiency. However, despite the identified inefficiencies, the minibus taxi paratransit system adapts to fulfil the mobility needs of the urban commuters in the developing cities. For instance, it is widely known in the literature that developing cities in the Global South suffer from urban sprawl because of poor planning. However, the paratransit system organically transforms to serve the ever-growing fragmented settlements. We, therefore, believe that the paratransit system undergoes the process of self-organisation and adaptation.

The general system-wide (macro-level) behaviours in an organically-evolved paratransit system are shaped by local-level (micro-level) autonomous interactions between its entities, giving rise to a stable state at macro-level. Based on the previous finding that paratransit systems exhibited characteristics (such as many independent and autonomous components, no centralised management, self-organisation and adaptation) related to complex adaptive systems, we developed an agent-based model (ABM) to simulate transportation dynamics of a minibus taxi paratransit system. A controlled experiment (CER) of ABM simulation was setup, tuned, ran for 68 days, and validated using the data we collected from the field study.

Improving the intelligence and situational awareness of autonomous passenger and minibus taxi driver agents at the micro-level of the simulated paratransit system results in improved efficiency at the macro-level of the system. Two test simulation experiments were setup, i.e., test experiment one (TOR1) and test experiment two (TOR2). In each experiment, the passenger and minibus taxi agents were trained to make more intelligent decisions, and to improve their situational awareness. The training was done using supervised learning methods, i.e., random forests and convolutional neural networks. Results from the two test experiments showed an improvement in the macro-level efficiency metrics values. For example, the mean *passenger waiting time* reduced by 34% and 55%, for CER-to-TOR1 and CER-to-TOR2, respectively. Correspondingly, the percentage minibus taxi *occupancy* increased by 14% and 21% for CER-to-TOR1 and CER-to-TOR2, respectively. From the aforementioned results, we made three main conclusions about the simulated paratransit system. First, the minibus taxi paratransit system efficiency in TOR1 and TOR2 improved compared to CER. Second, the efficiency improvement observed was as a result of improved agents' intelligence at micro-level in addition to improved situational awareness. Third, the simulated paratransit system underwent self-organisation as the individual agents learned and adopted new behaviour through micro-level interactions.

9.2 Recommendations

Future work should consider including other modes of transport in the agent-based model. The other transport modes available in the paratransit ecosystem in the developing African cities include: the three-wheeled rickshaws and motorcycle taxis (boda bodas). These modes of transport play a big role in fulfilling the first and last legs (or first and last miles of commute) of the paratransit-related journeys. It would be interesting to model and investigate how intelligent micro-level decision making by rickshaws and boda bodas would affect the overall efficiency of the paratransit system at macro-level.

Further modelling work should be done towards developing a "flexible bus rapid transit (BRT)" design framework that incorporates the unique and diverse mobility needs in developing cities of the Global South. We propose such framework be called the "BRT-Flexi". The conceptual BRT-Flexi should combine the benefits of BRT with the flexibility, adaptability, demand responsiveness, and near ubiquitous characteristics of paratransit. BRT-Flexi may have high-capacity buses running along pre-selected corridors during peak hours and seamlessly integrated with low-capacity paratransit modes at different connection centres along the corridors. We believe this will further improve transportation efficiency in the developing cities of the Global South.

Finally, more work should be done on integrating "smart mobility" and ICT applications in paratransit systems. The applications can support the various aspects of the paratransit system, such as journey planning, booking, scheduling, fare collection, and payments. We believe that integrating smart mobility systems at various aspects of the paratransit system may further improve its efficiency.

9.3 General concluding statement

The world has moved on from working harder to working smarter or intelligently. Today we have the technology, computing power and storage capability to develop and train advanced models to support distributed intelligence. Indeed, the efficiency of paratransit systems in the developing cities of the Global South can be fundamentally transformed by integrating intelligent transport systems (ITS) in the transportation service ecosystem. The developing cities of the Global South have unique and diverse characteristics. These characteristics are dictated by the unique culture as well as social and economic challenges. Paratransit in such a setting works, but it is inefficient. Rather than enforce the Global North style of transport systems that are capital intensive, and costly to maintain, the efficiency of paratransit in developing cities of the Global South can be improved.

Bibliography

Abbott, R. and Hadžikadić, M. (2017). Complex Adaptive Systems, Systems Thinking, and Agent-Based Modeling. In Hadžikadić, M. and Avdaković, S., editors, Advanced Technologies, Systems, and Applications, pages 1–8, Cham. Springer International Publishing.

Aggrey, N. (2017). The thriving taxi business in Kampala.

Ai, Q., Wang, X., Golbandi, N., Bendersky, M., and Najork, M. (2018). Learning Groupwise Scoring Functions Using Deep Neural Networks. CoRR, abs/1811.04415.

Ajay Mahaputra Kumar, Vivien Foster, and Fanny Barrett (2008). Stuck in traffic : urban transport in Africa. Technical report, Washington, DC: World Bank.

Alstott, J., Bullmore, E., and Plenz, D. (2014). powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. PLoS ONE, 9(1):e85777.

Awumbila, M. (2017). Drivers of Migration and Urbanization in Africa: Key Trends and Issues. Technical report, United Nations: Population Division.

Azevêdo, Lima, C., Deshmukh, N., Balakumar, Marimuthu, Oh, S., Marczuk, K., Soh, H., Kakali, Ba_ak, Toledo, T., and Peh, L. (2016). SimMobility Short-term : An Integrated Microscopic Mobility 1 Simulator 2 3 4. In Proceedigns of 6th Annual Conference of the Transportation Research Board.

Bähre, E. (2014). A trickle-up economy: mutuality, freedom and violence in cape town's taxi associations. Africa: Journal of the International African Institute, 84(4):576–594.

Behrens, R., McCormick, D., and Mfinanga, D. (2015a). Paratransit in African cities: Operations, regulation and reform. Routledge, London, 1 edition.

Behrens, R., McCormick, D., and Mfinanga, D. (2015b). Paratransit in African cities: Operations, regulation and reform.

Bell, M., Galatioto, F., Namdeo, A., Franco, P., Tiwary, A., Goodman, P., Rhys-Tyler, G., and Namdeo, D. (2012). SECURE (SElf Conserving URban Environments) project: The route to delivering transport sustainability. In 1st International Conference on Urban Sustainability and resilience, pages 12–22.

Bendor, J., Diermeier, D., and Ting, M. (2003). A Behavioral Model of Turnout. American Political Science Review, 97(2):261–280.

Benhamou, S. (2004). How to reliably estimate the tortuosity of an animal's path:: straightness, sinuosity, or fractal dimension? Journal of Theoretical Biology, 229(2):209–220.

Bennet, A. and Bennet, D. (2004). Chapter 3 - The Intelligent Complex Adaptive System. In Bennet, A. and Bennet, D., editors, Organizational Survival in the New World, KMCI Press, pages 25 – 36. Butterworth-Heinemann, Boston.

Berger, T., Birner, R., Díaz, J., McCarthy, N., and Wittmer, H. (2007). Capturing the complexity of water uses and water users within a multi-agent framework. In Craswell, E., Bonnell, M., Bossio, D., Demuth, S., and Van De Giesen, N., editors, Integrated Assessment of Water Resources and Global Change: A North-South Analysis, pages 129–148. Springer Netherlands, Dordrecht.

Berger, T. and Troost, C. (2014). Agent-based Modelling of Climate Adaptation and Mitigation Options in Agriculture. Journal of Agricultural Economics, 65(2):323–348.

Bertrand, S., Bertrand, A., Guevara-Carrasco, R., and Gerlotto, F. (2007). Scale-Invariant Movements of Fishermen: The Same Foraging Strategy as Natural Predators. Ecological Applications, 17(2):331–337.

Blume, L. (2015). Agent-Based Models for Policy Analysis. In Wallace, R., Geller, A., and Ogawa, V. A., editors, Assessing the use of agent-based models for tobacco regulation. The National Academies Press, Washington, DC.

Blyth, D., Alcaraz, J., Binet, S., and Chekanov, S. V. (2019). ProIO: An event-based I/O stream format for protobuf messages. Computer Physics Communications, 241:98 – 112.

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America, 99 Suppl 3(90003):7280–7.

Booysen, M. J., Andersen, S. J., and Zeeman, A. S. (2013). Informal public transport in Sub-Saharan Africa as a vessel for novel Intelligent Transport Systems. In 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), pages 767–772.

Bosse, T., Merk, R.-J., and Treur, J. (2012). Modelling Temporal Aspects of Situation Awareness. In Huang, T., Zeng, Z., Li, C., and Leung, C. S., editors, Neural Information Processing, pages 473–483, Berlin, Heidelberg. Springer Berlin Heidelberg.

Bradbury, A. and Howe, J. (2002). Boda boda: Uganda's rural and urban low-capacity transport services. In Godard, X. and Fatounzoun, I., editors, Urban Mobility for All. Proceedings of the 10th International CODATU conference, pages 235–240, Lisse (Netherlands). Balkema Publishers.

Brady, M., Sahrbacher, C., Kellermann, K., and Happe, K. (2012). An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services. Landscape Ecology, 27(9):1363–1381.

Breiman, L. (2001). Random Forests. In Machine Learning, volume 45, pages 5–32. Kluwer Academic Publishers, Netherlands.

Brockmann, D., Hufnagel, L., and Geisel, T. (2006). The scaling laws of human travel. Nature, 439(7075):462–465.

Brotcorne, L., De Wolf, D., Gendreau, M., and Labbé, M. (2002). A Dynamic User Equilibrium Model for Traffic Assignment in Urban Areas. In Gendreau, M. and Marcotte, P., editors, Transportation and Network Analysis: Current Trends: Miscellanea in honor of Michael Florian, pages 49–69. Springer US, Boston, MA.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Grobler, A. G., Jaques, Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pages 108–122.

Burgin, M. and Dodig-Crnkovic, G. (2009). A Systematic Approach to Artificial Agents.

Cao, Y., Li, Y., Zhou, J., Jin, D., Su, L., and Zeng, L. (2011). Recognizing the levy-walk nature of vehicle mobility. In IWCMC 2011 - 7th International Wireless Communications and Mobile Computing Conference, pages 1194–1199.

Cao, Z., Qin, T., Liu, T.-Y., Tsai, M.-F., and Li, H. (2007). Learning to Rank: From Pairwise Approach to Listwise Approach. In Proceedings of the 24th International Conference on Machine Learning, ICML '07, pages 129–136, New York, NY, USA. Association for Computing Machinery.

Capra, F. and Luisi, P. (2014). The Newtonian world-machine. In Capra, F. and Luisi, P. L., editors, The Systems View of Life: A Unifying Vision, pages 19–34. Cambridge University Press, Cambridge.

Cascetta, E. (2009). Random Utility Theory. In Transportation Systems Analysis: Models and Applications, pages 89–167. Springer US, Boston, MA.

Cervero, R. and Golub, A. (2007). Informal transport: A global perspective. Transport Policy, 14(6):445–457.

Ciari, F., Bock, B., and Balmer, M. (2014). Modeling Station-Based and Free-Floating Carsharing Demand: Test Case Study for Berlin. Transportation Research Record, 2416(1):37–47.

Cohen, R. and Havlin, S. (2010). Structure and Robustness of Complex Networks.

Daniel E. Agbiboa (2018). Transport, Transgression and Politics in African Cities: The Rhythm of Chaos. Routledge, London, 1 edition.

Dawid, H. and Neugart, M. (2011). Agent-based Models for Economic Policy Design. Eastern Economic Journal, 37(1):44–50.

Deadman, P. J., Schlager, E., and Gimblett, R. (2000). Simulating common pool resource management experiments with adaptive agents employing alternate communication routines. Journal of Artificial Societies and Social Simulation, 3(2).

DeAngelis, D. L. and Mooij, W. M. (2005). Individual-based modeling of ecological and evolutionary processes. Annual Review of Ecology, Evolution, and Systematics, 36:147–168.

Diaz Olvera, L., Guézéré, A., Plat, D., and Pochet, P. (2016). Earning a living, but at what price? Being a motorcycle taxi driver in a Sub-Saharan African city. Journal of Transport Geography, 55:165–174.

Diaz Olvera, L., Plat, D., and Pochet, P. (2019). Looking for the obvious: Motorcycle taxi services in Sub-Saharan African cities. Journal of Transport Geography, page 102476.

Dong, X., Chen, X., Guan, Y., Yu, Z., and Li, S. (2009). An Overview of Learning to Rank for Information Retrieval, volume 3.

Dorothy, M., Herrie, S., and David, M. (2016). The nature of paratransit operations. In Roger, B., Dorothy, M., and David, M., editors, Paratransit in African Cities: Operations, Regulation and Reform, chapter 3, pages 59–78. Routledge, New York, 1 edition.

Dorothy, N. (2018). Money in starting a taxi business.

du Preez, D., Zuidgeest, M., and Behrens, R. (2019). A quantitative clustering analysis of paratransit route typology and operating attributes in Cape Town. Journal of Transport Geography, 80(C).

Eboli, L. and Mazzulla, G. (2011). A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. Transport Policy, 18(1):172 – 181.

Eboli, L. and Mazzulla, G. (2012). Performance indicators for an objective measure of public transport service quality. European Transport \ Trasporti Europei, (51):1–4.

Edwards, A. M., Phillips, R. A., Watkins, N. W., Freeman, M. P., Murphy, E. J., Afanasyev, V., Buldyrev, S. V., da Luz, M. G. E., Raposo, E. P., Stanley, H. E., and Viswanathan, G. M. (2007). Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. Nature, 449(7165):1044–1048.

Ehebrecht, D., Heinrichs, D., and Lenz, B. (2018). Motorcycle-taxis in sub-Saharan Africa: Current knowledge, implications for the debate on "informal" transport and research needs. Journal of Transport Geography, 69:242–256.

Elsas, J. L., Carvalho, V. R., and Carbonell, J. G. (2008). Fast Learning of Document Ranking Functions with the Committee Perceptron. In Proceedings of the 2008 International Conference on Web Search and Data Mining, WSDM '08, pages 55–64, New York, NY, USA. Association for Computing Machinery.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors, 37(1):32–64.

Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. Proceedings of the National Academy of Sciences, 99(suppl 3):7243–7250.

Eubank, S., Guclu, H., Anil Kumar, V. S., Marathe, M. V., Srinivasan, A., Toroczkai, Z., and Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. Nature, 429(6988):180–184.

Evans, J., O'Brien, J., and Ch Ng, B. (2018). Towards a geography of informal transport: Mobility, infrastructure and urban sustainability from the back of a motorbike. Transactions of the Institute of British Geographers, 43(4):674–688.

Ferreira, A., Raposo, E., Viswanathan, G., and da Luz, M. (2012). The influence of the environment on Lévy random search efficiency: Fractality and memory effects. Physica A: Statistical Mechanics and its Applications, 391(11):3234–3246.

Florian, M. and Hearn, D. W. (2008). Traffic Assignment: Equilibrium Models. In Chinchuluun, A., Pardalos, P. M., Migdalas, A., and Pitsoulis, L., editors, Pareto Optimality, Game Theory And Equilibria, pages 571–592. Springer New York, New York, NY.

Franklin, S. and Graesser, A. (1997). Is It an agent, or just a program?: A taxonomy for autonomous agents. In Müller, J. P., Wooldridge, M. J., and Jennings, N. R., editors, Third International Workshop on Agent Theories, Architectures, and Languages, pages 21–35, Berlin, Heidelberg. Springer Berlin Heidelberg.

Freund, Y., Iyer, R., Schapire, R. E., and Singer, Y. (2003). An Efficient Boosting Algorithm for Combining Preferences. J. Mach. Learn. Res., 4(null):933–969.

Freund, Y. and Schapire, R. E. (1996). Experiments with a New Boosting Algorithm. In Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, ICML'96, pages 148–156, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Gauthier, A. and Weinstock, A. (2010). Africa Transforming Paratransit into BRT. Built Environment, 36(3):317–327.

Godard, X. (2013). Comparisons of urban transport sustainability: Lessons from West and North Africa. Research in Transportation Economics, 40(1):96–103.

GoMetro (2016). Discover Your City: Collect emerging transport data in emerging markets.

Goodfellow, T. (2010). 'The bastard child of nobody?': Anti-planning and the institutional crisis in contemporary Kampala.

Goodfellow, T. (2012). State effectiveness and the politics of urban development in East Africa: A puzzle o PhD thesis, London School of Economics and Political Science.

Goodfellow, T. (2017). 'Double Capture' and De-Democratisation: Interest Group Politics and Uganda's 'Transport Mafia'. The Journal of Development Studies, 53(10):1568–1583.

Gorkem, G., Huseyin, C., Ozgur, B., and Halim, C. (2014). Using Potential Accessibility Measure for Urban Public Transportation Planning: A Case Study of Denizli, Turkey. Promet - Traffic and Transportation, 26(2):129–137.

GPy (2012). GPy: A Gaussian process framework in python.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U., and DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. Ecological Modelling, 198(1):115–126.

Grisan, E., Foracchia, M., and Ruggeri, A. (2003). A novel method for the automatic evaluation of retinal vessel tortuosity. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), volume 1, pages 866–869.

Güting, R. H., Schneider, M., Güting, R. H., and Schneider, M. (2005). Modeling and Querying History of Movement. Moving Objects Databases, pages 99–185.

Hammod, R. A. (2015). Considerations and Best Practices in Agent-Based Modeling to Inform Policy. In Wallace, R., Geller, A., and Ogawa, V. A., editors, Assessing the use of agent-based models for tobacco regulation. The National Academies Press.

Hamstra, C. (2017). Complexity storytelling: The science of complexity within the art of communication. Emergence: Complexity and Organization, 1.

Hawking, S. (2002). On the Shoulders of Giants: The Great Works of Physics and Astronomy. Running Press.

Hayles, K. (1991). Chaos: More than a metaphor. The University of Chicago Press.

Heckbert, S., Baynes, T., and Reeson, A. (2010). Agent-based modeling in ecological economics. Annals of the New York Academy of Sciences, 1185(1):39–53.

Holland, J. H. (1992). Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applic MIT Press, Cambridge, MA, USA.

Holland, J. H. (1995). Hidden Order: How Adaptation Builds Complexity, volume 17. SAGE Publications Inc, New York.

Holte, J. (1993). Chaos The New Science. In Nobel Conference Series. University Press of America.

Hoogendoorn, M., Van Lambalgen, R. M., and Treur, J. (2011). Modeling Situation Awareness in Human-like Agents Using Mental Models. In Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two, IJCAI'11, pages 1697–1704. AAAI Press.

Hooker, C. (2011). Introduction to Philosophy of Complex Systems: A. Part A: Towards a framework for co volume 10. Elsevier B.V.

Institute of Medicine (2015). Assessing the Use of Agent-Based Models for Tobacco Regulation. The National Academies Press, Washington, DC.

ITP (2010). Pre-feasibility Studies for the Development of a Long Term Integrated Bus Rapid Transit system for Greater Kampala Metropolitan Area. Technical report, World Bank, Kampala City Council, Government of Uganda.

Jahan, S. (2016). Human Development for Everyone. Technical report, United Nations Development Programme (UNDP).

Jakob, M., Moler, Z., Komenda, A., Yin, Z., Jiang, A. X., Johnson, M. P., Pěchouček, M., and Tambe, M. (2012). AgentPolis: Towards a Platform for Fully Agent-Based Modeling of Multi-Modal Transportation (Demonstration). In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3, AAMAS '12, pages 1501–1502, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

James, A., Plank, M. J., and Edwards, A. M. (2011). Assessing Lévy walks as models of animal foraging. Journal of The Royal Society Interface, 8(62):1233–1247.

Jean, T., Priti, G., and Paul, K. (2018). Road Safety Performance Review: Uganda. Technical report, United Nations Economic Commission for Africa (UNECA), New York and Geneva.

Jennings, G. and Behrens, R. (2017). The Case for Investing in Paratransit: Strategies for regulation and re

Jiang, B., Yin, J., and Zhao, S. (2009). Characterizing the human mobility pattern in a large street network. Phys. Rev. E, 80(2):21136.

Jiang, M. (2020). Improving situational awareness with collective artificial intelligence over knowledge graphs. In Proc.SPIE, volume 11413.

JICA (2008). Dar Es Salaam Transport Policy and System Development Master Plan: Techinical Report 3. Technical report, Dar Es Salaam City Council, Dar Es Salaam.

JICA (2010). The study on greater Kampala road network and transport improvement in the republic of Uganda. Technical report, Ministry of Works and Transport (MOWT), Republic of Uganda, Kampala, Uganda.

JICA (2014). The Project on Integrated Urban Development Master Plan for the City of Nairobi in the Republic of Kenya: Final Report. Technical report, Nairobi City County (NCC), Nairobi City County (NCC).

Joachims, T. (2002). Optimizing Search Engines Using Clickthrough Data. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '02, pages 133–142, New York, NY, USA. Association for Computing Machinery.

Jokinen, K. and Wilcock, G. (2019). Multimodal open-domain conversations with robotic platforms. In Alameda-Pineda, X., Ricci, E., and Sebe, N., editors, Multimodal Behavior Analysis in the Wild, pages 9–26. Academic Press.

Jonker, N. J. and Venter, C. J. (2019). Modeling Trip-Length Distribution of Shopping Center Trips from GPS Data. Journal of Transportation Engineering, Part A: Systems, 145(1):4018079.

KCCA (2016). Multimodal Urban Transport Master Plan for Greater Kampala Metropolitan Area. Technical report, Kampala Capital City Authority (KCCA), Kampala.

Klopp, J. M. and Cavoli, C. (2019). Mapping minibuses in Maputo and Nairobi: engaging paratransit in transportation planning in African cities. Transport Reviews, 39(5):657–676.

Kumar, A. M. (2011). Understanding the emerging role of motorcycles in African cities : a political economy perspective. Technical Report 13, Washington, DC: World Bank, Washington, DC.

Kussul, N., Lavreniuk, M., Skakun, S., and Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geoscience and Remote Sensing Letters, 14(5):778–782.

Ladyman, J., Lambert, J., and Wiesner, K. (2013). What is a complex system? European Journal for Philosophy of Science, 3(1):33–67.

Laine, T. and Bruun, E. (2017). Changes to Travel Demand Modeling in the Age of MaaS – example of Brutus. In ICoMaaS.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature, 521(7553):436–444.

Lee, B. Y., Brown, S. T., Korch, G. W., Cooley, P. C., Zimmerman, R. K., Wheaton, W. D., Zimmer, S. M., Grefenstette, J. J., Bailey, R. R., Assi, T.-M., and Burke, D. S. (2010). A computer simulation of vaccine prioritization, allocation, and rationing during the 2009 H1N1 influenza pandemic. Vaccine, 28(31):4875–4879.

Levine, J., Grengs, J., Shen, Q., and Shen, Q. (2012). Does Accessibility Require Density or Speed? Journal of the American Planning Association, 78(2):157–172.

Lindberg, C. and Schneider, M. (2013). Combating infections at Maine Medical Center: Insights into complexity-informed leadership from positive deviance. Leadership, 9(2):229–253.

Litman, T. (2012). Current mobility trends: implications for sustainability. In Keep moving, towards sustainable mobility, pages 23–44. Eleven International Publishing, Amsterdam.

Liu, T.-Y. (2011). Applications of Learning to Rank. In Liu, T.-Y., editor, Learning to Rank for Information Retrieval, pages 181–191. Springer Berlin Heidelberg, Berlin, Heidelberg.

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. Journal of the Atmospheric Sciences, 20(2):130–141.

Lucas, K., Jennings, G., Zuidgeest, M., Behrens, R., Venter, C., Mitullah, W., and Thornton, B. (2019). Transport Social Exclusion in Five African Cities. Technical report, International Network for Transport and Accessibility in Low Income Communities, Leeds,UK.

Luhmann, N. and Gilgen, P. (2012). General Systems Theory. In Introduction to Systems Theory, pages 25–120. Polity Press.

Marcius, C., Takao, S., and Daniel, N. (2015). Efficiency and Effectiveness Analysis of Public Transport of Brazilian Cities. Journal of Transport Literature, 9:40 – 44.

Marion, R. (1999). The Edge of Organization: Chaos and Complexity Theories of Formal Social Systems. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States .

Maroulis, S., Bakshy, E., Gomez, L., and Wilensky, U. (2014). Modeling the Transition to Public School Choice. Journal of Artificial Societies and Social Simulation, 17(2):3.

Martín H., J. A., de Lope, J., and Maravall, D. (2008). Adaptation, anticipation and rationality in natural and artificial systems: computational paradigms mimicking nature. Natural Computing, 8(4):757.

Masad, D. and Kazil, J. (2015). Mesa: An Agent-Based Modeling Framework. In Huff, K. and Bergstra, J., editors, Proceedings of the 14th Python in Science Conference, pages 51–58.

Matthews, M. L., Bryant, D. J., Webb, R. D. G., and Harbluk, J. L. (2001). Model for Situation Awareness and Driving: Application to Analysis and Research for Intelligent Transportation Systems. Transportation Research Record, 1779(1):26–32.

Mills, F. and Stufflebeam, R. (2017). Introduction to Intelligent Agents.

MoWT, M. o. W. and Transport (2009). National Transport Master Plan including a transport master plan for the Greater Kampala Metropolitan Area (NTMP/GKMA). Technical report, Ministry of Works and Transport, Republic of Uganda, Kampala.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., and Schwarz, N. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environmental Modelling & Software, 48:37–48.

Mutiso, W. and Behrens, R. (2011). ‘Boda Boda’ bicycle taxis and their role in urban transport systems: case studies of Kisumu and Nakura, Kenya. In 30th Annual Southern African Transport Conference 11-14 July 2011 “Africa on the Move”, CSIR International Convention Centre, pages 430–444, Pretoria, South Africa.

Naderpour, M., Lu, J., and Kerre, E. (2012). A Conceptual Model for Risk-Based Situation Awareness. In Wang, Y. and Li, T., editors, Foundations of Intelligent Systems, pages 297–306. Springer Berlin Heidelberg, Berlin, Heidelberg.

Ndibatya, I. and Booysen, M. (2020a). Minibus taxis in Kampala's paratransit system: Operations, economics and efficiency. Journal of Transport Geography, 88.

Ndibatya, I. and Booysen, M. (2020b). Transforming Paratransit in Africa's congested Cities: An ICT- enabled Integrated Demand Responsive Transport (iDRT) approach. In Miriam, C. and Paul, C., editors, IST-Africa 2020 Conference Proceedings, pages 1–10, Kampala, Uganda. IST-Africa Institute and IIMC.

Ndibatya, I., Booysen, M. J., and Quinn, J. (2014). An adaptive transportation prediction model for the informal public transport sector in Africa. In 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), pages 2572–2577.

Ndibatya, I., Coetzee, J., and Booysen, M. J. (2016). Mapping the informal public transport network in Kampala with smartphones: Making sense of an organically evolved chaotic system in an emerging city in sub-Saharan Africa. In Proc. 35th Southern African Transport Conference, Pretoria, pages 4–7.

Newman, M. E. J. (2011). Complex Systems: A Survey. American Journal of Physics, 79(8):800–810.

Newton, I. (1687). Philosophiae Naturalis Principia Mathematica ("Mathematical Principles of Natural Phi volume 1 of Mathematical Principles of Natural Philosophy. London, 1 edition.

Newton, I. (1713). Philosophiae Naturalis Principia Mathematica ("Mathematical Principles of Natural Phi volume 2. 1 edition.

Newton, I. (1726). Philosophiae Naturalis Principia Mathematica ("Mathematical Principles of Natural Phi volume 3. 1 edition.

Nkurunziza, A., Zuidgeest, M., Brussel, M., and Maarseveen, M. (2012). Modeling Commuter Preferences for the Proposed Bus Rapid Transit in Dar-es-Salaam. Journal of Public Transportation, 15:95–116.

North, M. J., Collier, N. T., Ozik, J., Tatara, E. R., Macal, C. M., Bragen, M., and Sydelko, P. (2013). Complex adaptive systems modeling with Repast Simphony. Complex Adaptive Systems Modeling, 1(1):3.

ODA (2020). Minibus Taxi Operators in Kampala Ready to Engage on BRT Plan.

Odell, J. (2002). Agents and complex systems. Journal of Object Technology, 1(2):35–45.

Oh, S., Byon, Y.-J., Jang, K., and Yeo, H. (2018). Short-term travel-time prediction on highway: A review on model-based approach. KSCE Journal of Civil Engineering, 22(1):298–310.

Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. Nature, 441(7092):502–505.

Ommeh, M. S., Chitere, P., Orero, R., Mitullah, W., and McCormick, D. (2015). The politics behind the phasing out of the 14-seater matatu in Kenya.

OpenStreetMap contributors (2017). Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org.

Ormerod, P. and Rosewell, B. (2009). Validation and Verification of Agent-Based Models in the Social Sciences. In Squazzoni, F., editor, Epistemological Aspects of Computer Simulation in the Social Sciences, pages 130–140, Berlin, Heidelberg. Springer Berlin Heidelberg.

Ortúzar, J. D. and Willumsen, L. G. (2011). Modelling Transport. John Wiley and Sons.

Pablo, S. F. (2015). Paratransit: a key element in a dual system. Technical report, Cooperation for urban mobility in the developing world (CODATU), Paris, France.

Pasumarthi, R. K., Bruch, S., Wang, X., Li, C., Bendersky, M., Najork, M., Pfeifer, J., Golbandi, N., Anil, R., and Wolf, S. (2019). TF-Ranking: Scalable TensorFlow Library for Learning-to-Rank. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 2970–2978.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12:2825–2830.

Pelletier, C., Webb, G., and Petitjean, F. (2019). Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. Remote Sensing, 11:523.

Perreault, C. and Brantingham, P. J. (2011). Mobility-driven cultural transmission along the forager–collector continuum. Journal of Anthropological Archaeology, 30(1):62 – 68.

Phillips, O. and Mesharch, W. K. (2018). Effectiveness of KCCA in the Traffic Management and Solid Waste Management Sectors in Kampala: The Informal Sector and the City Economy. In Joshua Mugambwa and Mesharch W. Katusiimeh, editors, Handbook of Research on Urban Governance and Management in the Developing World, chapter 17, pages 306–318. IGI Global, Hershey, PA.

Plano, C., Behrens, R., and Zuidgeest, M. (2020). Towards evening paratransit services to complement scheduled public transport in Cape Town: A driver attitudinal survey of alternative policy interventions. Transportation Research Part A: Policy and Practice, 132:273–289.

Poincare, H. (1881). Mémoire sur les courbes définies par une équation différentielle. Journal de mathématiques pures et appliqués, 7:375–422.

Poincare, H. (1908). Science et méthode. Journal de mathématiques pures et appliqués, 2:200–231.

Pojani, D. and Stead, D. (2017). The Urban Transport Crisis in Emerging Economies: An Introduction. In Pojani, D. and Stead, D., editors, The Urban Transport Crisis in Emerging Economies, pages 1–10. Springer International Publishing, Cham.

Pojani, D. and Stead, D. (2018). Policy design for sustainable urban transport in the global south. Policy Design and Practice, 1(2):90–102.

Prigogine, I. (1961). Introduction to thermodynamics of irreversible processes. 2 edition.

Qin, T., Liu, T.-Y., and Li, H. (2010). A general approximation framework for direct optimization of information retrieval measures. Information Retrieval, 13(4):375–397.

Raichlen, D. A., Wood, B. M., Gordon, A. D., Mabulla, A. Z. P., Marlowe, F. W., and Pontzer, H. (2014). Evidence of Levy walk foraging patterns in human hunter-gatherers. Proceedings of the National Academy of Sciences of the United States of America, 111 2:728–33.

Raisio, H. and Lundström, N. (2014). Managing Chaos: Lessons From Movies on Chaos Theory. Administration & Society, 49(2):296–315.

Ramos-Fernandez, G., Mateos, J. L., Miramontes, O., Cocho, G., Larralde, H., and Ayala-Orozco, B. (2003). Levy walk patterns in the foraging movements of spider monkeys (Ateles geoffroyi).

Ranacher, P. and Tzavella, K. (2014). How to compare movement?: A review of physical movement similarity measures in geographic information science and beyond. Cartography and geographic information science, 41(3):286–307.

Rand, W. and Rust, R. (2011). Agent-Based Modeling in Marketing: Guidelines for Rigor. International Journal of Research in Marketing, 28:181–193.

Ray, T. and Triantaphyllou, E. (1999). Procedures for the evaluation of conflicts in rankings of alternatives. Computers & Industrial Engineering, 36(1):35–44.

Reynolds, A. (2015). Liberating Lévy walk research from the shackles of optimal foraging. Physics of Life Reviews, 14:59 – 83.

Reynolds, A. M. (2018). Current status and future directions of Lévy walk research. Biology Open, 7(1).

Rhee, I., Shin, M., Hong, S., Lee, K., and Chong, S. (2008). On the Levy-Walk Nature of Human Mobility. In Proceedings - IEEE INFOCOM, pages 1597–1605.

Rhee, I., Shin, M., Hong, S., Lee, K., Kim, S. J., and Chong, S. (2011). On the Levy-Walk Nature of Human Mobility. IEEE/ACM Trans. Netw., 19(3):630–643.

Richard, I. (2005). Regulation of Public Transport Services. In Public Transport in Developing Countries, chapter 19, pages 403–443. Emerald Group Publishing Limited, 2 edition.

Riley, K. (2012). The Butterfly Effect. In International Congress on Mathematical Education, volume 69, pages 369–373, Seoul, Korea.

Ringler, P., Keles, D., and Fichtner, W. (2016). Agent-based modelling and simulation of smart electricity grids and markets – A literature review. Renewable and Sustainable Energy Reviews, 57:205 – 215.

Robert Heinze (2018). "Taxi Pirates": a comparative history of informal transport in Nairobi and Kinshasa, 1960s–2000s. In Agbiboa, D., editor, Transport, Transgression and Politics in African Cities: The Rhythm of Chaos, chapter 1, pages 31–52. Routledge, London and New York, 1 edition.

Robertson, D. (2005). Agent-Based Modeling Toolkits NetLogo, RePast, and Swarm. Academy of Management Learning and Education, The, 4:524–527.

Rouzafzoon, J. and Helo, P. (2016). Developing service supply chains by using agent based simulation. Industrial management & data systems, 116(2):255–270.

Russell, S. and Norvig, P. (2003). Intelligent Agents. In Artificial Intelligence: A Modern Approach, chapter 2. Prentice Hall Press, USA, 2rd edition.

S Kisaalita, W. and Sentongo-Kibalama, J. (2007). Delivery of urban transport in developing countries: the case for the motorcycle taxi service (boda-boda) operators of Kampala. Development Southern Africa, 24(2):345–357.

Saddier, S., Patterson, Z., Johnson, A., and Chan, M. (2016). Mapping the Jitney Network with Smartphones in Accra, Ghana: The AccraMobile Experiment. Transportation Research Record, 2581(1):113–122.

Saleh, M., Esa, Y., Onuorah, N., and Mohamed, A. A. (2017). Optimal microgrids placement in electric distribution systems using complex network framework. In 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), pages 1036–1040.

Sampaio, B. R., Neto, O. L., and Sampaio, Y. (2008). Efficiency analysis of public transport systems: Lessons for institutional planning. Transportation Research Part A: Policy and Practice, 42(3):445–454.

Sanstad, A. H. (2015). Assessing Agent-Based Models for Regulatory Applications: Lessons from Energy Analysis. In Wallace, R., Geller, A., and Ogawa, V. A., editors, Assessing the use of agent-based models for tobacco regulation. The National Academies Press.

Scafetta, N. (2011). Understanding the complexity of the Lévy-walk nature of human mobility with a multi-scale cost/benefit model. Chaos, 21(4).

Schalkwyk, D. and Sadie, Y. (2011). A troubled journey: The South African government and the taxi recapitalisation policy, 1998-2008. 43(1):79–108.

Schelling, T. C. (1971). Dynamic models of segregation. The Journal of Mathematical Sociology, 1(2):143–186.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61:85–117.

Schneider, M. and Somers, M. (2006). Organizations as complex adaptive systems: Implications of Complexity Theory for leadership research. The Leadership Quarterly, 17(4):351 – 365.

Shiell, A., Hawe, P., and Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. BMJ, 336(7656):1281–1283.

Shlesinger, M. F. (2006). How does one best search for non-replenishable targets at unknown positions? An optimized search strategy could be applied to situations as diverse as animal foraging and time-sensitive rescue missions. Nature, 443(7109):281–282.

Simon, H. (1957). Models of Man: Social and Rational. In Mathematical Essays on Rational Human Behavior in a Social Setting. Wiley.

Spooner, D., Mwanika, J., and Manga, E. (2020). Kampala Bus Rapid Transit: Understanding Kampala's P

Stewart-Wilson, G., Sewankambo, N., Muwanga, N., Kasimbazi, E., Musuya, T., Droruga, N., Mwadime, R., Eriki, P., and Muyonga-Namayengo, F. (2017). Owning Our Urban Future: The Case of Kampala City. Technical report, Uganda National Academy of Sciences :Forum on Cities, Urbanization, and Services (FOCUS), Kampala.

Susan Eve Woolf (2014). South African taxi hand signs : documenting the history and significance of taxi h
PhD thesis, University of the Witwatersrand.

Svore, K. M., Volkovs, M. N., and Burges, C. J. C. (2011). Learning to Rank with Multiple Objective Functions. In Proceedings of the 20th International Conference on World Wide Web, WWW '11, pages 367–376, New York, NY, USA. Association for Computing Machinery.

Talabis, M. R. M., McPherson, R., Miyamoto, I., Martin, J. L., and Kaye, D. (2015). Analytics Defined. In Talabis, M. R. M., McPherson, R., Miyamoto, I., Martin, J. L., and Kaye, D., editors, Information Security Analytics, pages 1–12. Syngress, Boston.

Tambulasi, R. I. C. and Kayuni, H. M. (2008). Can the state perpetuate the marginalisation of the poor? The socioeconomic effects of the state's ban on minibus 'callboys' in Malawi. Development Southern Africa, 25(2):215–226.

Toroczkai, Z. and Eubank, S. (2006). Agent-based Modeling as a Decision Making Tool: How to Halt a Smallpox Epidemic. Frontiers of Engineering, pages 21–27.

Turner, J. R. and Baker, R. M. (2019). Complexity Theory: An Overview with Potential Applications for the Social Sciences. Systems, 7(1):4.

UBOS (2014). National population and housing census 2014 - Uganda. Technical report, Uganda bureau of statistics (UBOS), Kampala, Uganda.

Uganda Parliament (1998). Traffic and Road Safety Act 1998.

Uganda Parliament (2011). The Kampala Capital City Act, 2010.

Uganda Parliament (2012). Traffic and Road Safety (Driving Tests and Special Provisions for Drivers of Public Service Vehicles and Goods Vehicles) Regulations, 2012.

Usyukov, V. (2017). Methodology for identifying activities from GPS data streams. Procedia Computer Science, 109:10 – 17.

Valizadegan, H., Jin, R., Zhang, R., and Mao, J. (2009). Learning to Rank by Optimizing NDCG Measure. In Proceedings of the 22nd International Conference on Neural Information Processing Systems, NIPS'09, pages 1883–1891, Red Hook, NY, USA. Curran Associates Inc.

Veng, Kheang, P. and Tetsuo, Y. (2016). State of the Art of Paratransit Literatures in Asian Developing Countries. Asian Transport Studies, 4(1):57–77.

Venter, C., Jennings, G., Hidalgo, D., and Pineda, A. F. V. (2018). The equity impacts of bus rapid transit: A review of the evidence and implications for sustainable transport. International Journal of Sustainable Transportation, 12(2):140–152.

Venter, C., Mahendra, A., and Hidalgo, D. (2019). From Mobility to Access for All: Expanding Urban Transportation Choices in the Global South. Technical report, World Resources Institute.

Vermeiren, K., Verachtert, E., Kasaija, P., Loopmans, M., Poesen, J., and Van Rompaey, A. (2015). Who could benefit from a bus rapid transit system in cities from developing countries? A case study from Kampala, Uganda. Journal of Transport Geography, 47:13–22.

Viswanathan, G., Raposo, E., and da Luz, M. (2008). Lévy flights and superdiffusion in the context of biological encounters and random searches. Physics of Life Reviews, 5(3):133–150.

Viswanathan, G. M., Buldyrev, S. V., Havlin, S., da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (1999). Optimizing the success of random searches. Nature, 401(6756):911–914.

Viswanathan, G. M., da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (2011). Lévy flight foraging. In The Physics of Foraging: An Introduction to Random Searches and Biological Encounters, pages 85–99. Cambridge University Press.

Weidlich, A. and Veit, D. (2008). A critical survey of agent-based wholesale electricity market models. Energy Economics, 30(4):1728 – 1759.

Wheeler, G. (2020). Bounded Rationality. In Zalta, E. N., editor, The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, fall 2020 edition.

Whitney, C. W., Lind, B. K., and Wahl, P. W. (1998). Quality Assurance and Quality Control in Longitudinal Studies. Epidemiologic Reviews, 20(1):71–80.

Williams, S., White, A., Waiganjo, P., Orwa, D., and Klopp, J. (2015). The digital matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. Journal of Transport Geography, 49:39–51.

Wood, A. (2015). Multiple Temporalities of Policy Circulation: Gradual, Repetitive and Delayed Processes of BRT Adoption in South African Cities. International Journal of Urban and Regional Research, 39:568–580.

Woolf, S. E. and Joubert, J. W. (2013). A people-centred view on paratransit in South Africa. Cities, 35:284–293.

Xiang, X., Kennedy, R., Madey, G., and Cabaniss, S. (2005). Verification and Validation of Agent-based Scientific Simulation Models. In 2005 Agent-Directed Simulation Symposium ADS'05.

Zhang, L. and Levinson, D. (2004). Agent-Based Approach to Travel Demand Modeling: Exploratory Analysis. Transportation Research Record, 1898(1):28–36.

Zimmerman, B., Lindberg, C., and Plsek, P. (2001). A Complexity Science Primer. In Edgeware: Insights from complexity science for health care leaders, volume 1, page 38. John Wiley & Sons, Ltd.

Appendix A

Algorithms

A.1 Class definitions

Algorithm 4: PASSENGERAGENT: Passenger Class definition

```
Class PassengerAgent
    Data:
        passenger_id
        state
        journey_diary
        pos                  // lat lon position
        wait_time
        load_cond
        sensor
    Functions:
        planJourney()
        searchForTaxi()
        executeJourney()
        updateJourneyDiary()
        disembark()
        boardMv()
        moveTo()
        runISM()
        runBCM()
        runACM()
        dropJourney()
        updateLocation()
        loadNextJourney()
end
```

Algorithm 5: MINIBUSTAXIAGENT: Minibus Taxi Class definition

```
Class MinibusTaxiAgent
    Data:
        vehicle_id
        state
        pos
        pob          // passengers on board
        trip
        sensor
        hbtm                // hold back time
        stop_interrupt
    Functions:
        tautForPassengers()
        searchForRoute()
        executeTrip()
        abandonTrip()
        findInitStop(sData)
        moveTo()
        runLFM()
        runRCM(sData)
        runHBM(sData)
end
```

Algorithm 6: MODEL: Model Class definition

```
Class Model
    Data:
        step_no
        pax_wait_threshold
        hold_back_threshold
        load_factor_threshold
        exit_condition
    Functions:
        generateJourneyDiaries()
        generateTripsMatrix()
        extractStopsNetwork()
        updateModelSchedule()
        saveModelInstance()
        step()
end
```

Glossary

autonomous Ability to make independent decisions. 5

autonomous agents Autonomous agents are software programs which respond to states and events in their environment independent from direct instruction by the user or owner of the agent, but acting on behalf and in the interest of the owner. 5

Global South Countries South of the equator in Africa, Asia and Latin America. 1

instantiate To instantiate is to create an instance of an object in an object-oriented programming (OOP) language. An instantiated object is given a name and created in memory or on disk using the structure described within a class declaration. 64

minibus A minibus, microbus, or minicoach is a passenger carrying motor vehicle that is designed to carry more people than a multi-purpose vehicle or minivan, but fewer people than a full-size bus. In the United Kingdom, the word "minibus" is used to describe any full-sized passenger carrying van. Minibuses have a seating capacity of between 8 and 30 seats. 1

paratransit A flexible transport service that provides individualised rides without fixed routes or timetables to supplement the fixed-route mass transit. 1

passenger A passenger (also abbreviated as pax) is a person who travels in a vehicle but bears little or no responsibility for the tasks required for that vehicle to arrive at its destination or otherwise operate the vehicle. Passengers often ride on buses, passenger trains, airliners, ships, ferryboats, and other methods of transportation. 1

public transport Public transport (also known as public transportation, public transit, or mass transit) is a shared passenger-transport service which is available for use by the general public, as distinct from modes such as taxicab, carpooling, or hired buses, which are not shared by strangers without private arrangement. Public transport modes include city buses, trolleybuses, trams (or light rail) passenger trains, rapid transit (metro/subways/undergrounds etc) ferries, and minibuses common in African cities and the Global South. 1