

Implementing MATSim Transit Simulation in a South African context

Dennis Wevell
28108206

Project Leader: Prof J.W. Joubert

Submitted in partial fulfilment of the requirements for the degree of

BACHELORS OF INDUSTRIAL ENGINEERING

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT
AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA

11 October 2011

Executive Summary

Transportation management in modern society is increasingly dependant on reliable transportation simulation to aid in the decision making process. Due to improvements in transportation technology and urbanization trends, transportation networks in urban areas are growing increasing complex and tools supporting the decision making process are struggling to remain accurate and useful. Traditional transportation simulation that assigns fixed values to certain routes in a static process model are severely limited in terms of representing different types of traffic and do not offer sufficient results. Popular new technologies include agent-based simulation that delivers a more accurate level of results but at the cost of being cumbersome and require high computational resources.

MATSim is an open-source, agent-based transportation simulation that has proved effective at simulating private car traffic in South African conditions. The MATSim toolkit has recently been expanded to include the simulation of public transportation networks along with the original traffic simulation providing an extended functionality. The MATSim toolkit will now be able to provide the information required to take a systems view of South African transit systems. This assists in the successful development of integrated transportation systems that currently doesn't exist in third-world countries like South Africa.

This project describes in detail the processes required to simulate transit networks under South African conditions using the MATSim toolkit. The functionality of the MATSim toolkit is explained and the data requirements to run a simulation are identified. The project also uses the case study on the Nelson Mandela Bay Metropolitan Municipality to illustrate the process of building a multimodal transportation simulation in South Africa and shows the applicability of the MATSim toolkit in South African transit planning.

Contents

Executive Summary	i
List of Acronyms	v
1 Introduction and Background	1
1.1 The need for transit simulation	1
1.2 Available tools for simulation	2
1.2.1 Traditional Models	2
1.2.2 Agent-Based Models	3
1.3 The minibus taxi network	4
1.4 Understanding a transit passenger	5
1.5 Research Design and Methodology	6
1.6 Outline of project	7
2 Literature Study	8
2.1 The MATSim process	8
2.2 Building the simulation	11
3 Obtaining the required data	13
3.1 The Multimodal Network	13
3.1.1 Data Breakdown	14
3.2 The Population	15
3.3 The Transit Schedule	15
3.3.1 Data Breakdown	16
3.3.2 Data gathering	17
3.4 The Transit Vehicles	19
4 Creating a MATSim Transit Simulation	20
4.1 The example public transport simulation	20
4.2 Possible adaptations to the example transit simulation	23
5 Running the example transit simulation	27
5.1 The standard example transit simulation	27
5.2 Implementing a basic bus network	27
5.3 Including paratransit for the simulation of a full transit system	29

6	Future development	37
6.1	Zone-based paratransit fare algorithm	37
6.2	The Nelson Mandela Bay case study	39
6.2.1	Working with the data	40
7	Conclusion	42
7.1	Summary	43
	Bibliography	44
	Appendix	46
A	Recommended ERD for storing South African specific GTFS data	46
B	Example of a MATSim config file	48
C	Fare comparisons between bus and taxi networks	50
D	Specifications for creating GTFS files	52
E	Instructions for converting GTFS data to MATSim	55
F	Transit schedule data conversion process	59

List of Figures

1.1	FSM: 1. Trip Generation	2
1.2	FSM: 2. Flow Distribution	3
1.3	FSM: 3. Modal Split	3
1.4	FSM: 4. Network Assignment	3
2.1	MATSim Process cycle	10
2.2	Extension points offered by the MATSim Controller	11
3.1	The population distribution in Nelson Mandela Bay	15
3.2	Interactive stop coordinate locator	19
4.1	Example transit simulation network	21
5.1	Utility score achieved per iteration, standard model	28
5.2	Total distance travelled per iteration, standard model	29
5.3	Utility score achieved per iteration, adapted model	30
5.4	Total travel distance per iteration, adapted model	31
5.5	Basic bus implementation route	32
5.6	Basic paratransit implementation route	32
5.7	Utility score achieved per iteration, full model	33
5.8	Total travel distance per iteration, full model	34
5.9	Public transit leg histogram (it. 50) of the full transit simulation	35
5.10	Public transit leg histogram (it. 50) of the adapted transit simulation	36
6.1	Zone breakdown for the application of a zone-based cost equa- tion	38
6.2	Road map of the NMBMM	39
6.3	Transfer data conversion process	41

List of acronyms

TAZ	Traffic Analysis Zones
ABS	Agent-Based Simulation
MATSim	Multi Agent Transport Simulation
FSM	Four Step Model
NMBMM	Nelson Mandela Bay Metropolitan Municipality
IPDM	Integrated Planning Development Modelling
GTFS	General Transit Feed Specifications
OSM	Open Street Map
TDF	Transit Data Feeder
ERD	Entity Relationship Diagram
NHTS	National Household Travel Survey
CITP	Comprehensive Integrated Transport Plan
GPS	Global Positioning System

Chapter 1

Introduction and Background

In a modern society transportation shapes the way we live. People plan their daily lives around their set activities and daily schedules. Since these activities tend to be in different locations, we rely on transportation to ensure our schedules are adhered to. The state of available transportation thus effectively determines the quality of our lives and has a profound effect on our behaviour on a daily basis.

1.1 The need for transit simulation

In South Africa cities public transport only service around 50% of the passenger market. The other 50% either do not have access to public transport or doesn't see public transport as a viable alternative to private transport (Fletterman, 2008). The 50% of the passenger market that makes use of public transport consists mainly of low-income workers who do not have the luxury of a private vehicle. To further complicate matters, due to past political influences, low-income labourers in South African cities tend to live mainly on the outskirts of the urban areas in poorly developed low-cost housing, "townships", squatter camps or the surrounding rural areas. This creates challenges in the planning of public transportation networks that cannot be met by employing the methods used by developed nations.

In order to meet these challenges within budget constraints imposed in a third-world country, transportation planners cannot merely guess at the best way to develop the transportation systems. Accurate transit simulations are required taking all relevant factors into account. These include:

Traffic congestion The effect on travelling time caused by congestion.

Fare structures The effective cost of transportation.

Accessibility of transit The distance between the transit pickup points and the homes of the prospective passengers.

Social stigmas Private car users tend to prefer driving above a cheaper public transport alternative.

1.2 Available tools for simulation

1.2.1 Traditional Models

Historically in South African transportation modelling has been dominated by the Four Step Model (FSM) (Diedericks and Joubert, 2006). In order to simulate a traffic system the FSM uses trips rather than activities. The basic operation of the FSM can be explained simply as follows: (Adapted from Meyer and Miller (2001)

- 1. Trip Generation** The study area is divided into Traffic Analysis Zones (TAZ). The number of trips each area will produce or attract is calculated and assigned as illustrated in Figure 1.1.

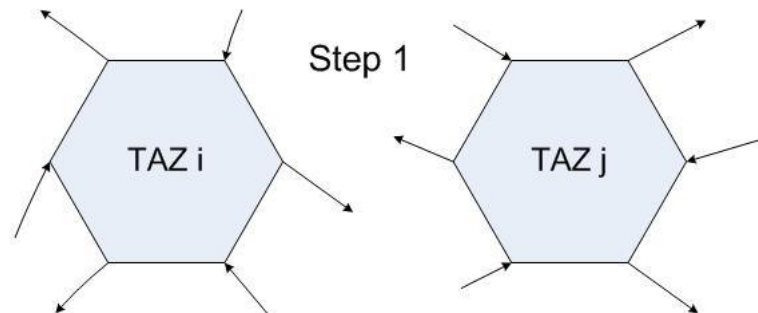


Figure 1.1: FSM: 1. Trip Generation

- 2. Flow Distribution** A fraction of trips are assigned to a traffic flow from origin TAZ i (T_i) to destination TAZ j (T_j) As illustrated in Figure 1.2. This flow is name T_{ij} .
- 3. Modal Split** Traffic flow between TAZ i and TAZ j is split into separate flows for each mode in operation between the origin-destination pair as illustrated in Figure 1.3.
- 4. Network Assignment** Each modal traffic is assigned a route through the network as illustrated in Figure 1.4. In an iterative process the flows are re-routed to satisfy capacity constraints to “balance” the network until re-routing does not improve the travel time of any trip.

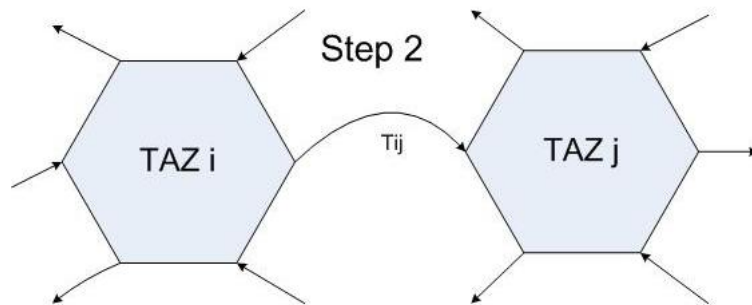


Figure 1.2: FSM: 2. Flow Distribution

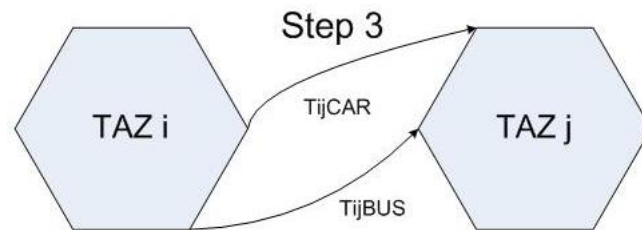


Figure 1.3: FSM: 3. Modal Split

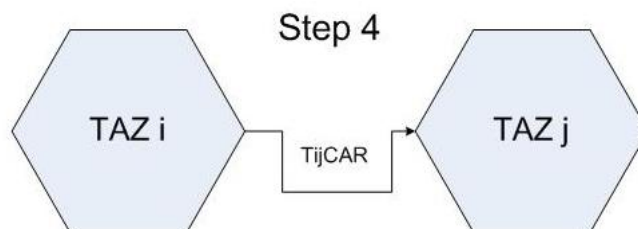


Figure 1.4: FSM: 4. Network Assignment

The problem with the traditional FSM is that traffic systems are getting more and more complex to a point where mathematical formulation of the model is no longer feasible (Rieser, 2010). This problem is enhanced when dynamic (time dependant) results are required adding additional complexity to the model. In these cases simulation is seen as the simpler way.

1.2.2 Agent-Based Models

Agent-Based Simulation (ABS) provides a modern approach transport modelling. Instead of calculating the trips employed, the activities that create a need for these trips drive the simulation (McNally, 2000). This is achieved

by creating several functional entities that in their entirety realistically simulates and predicts traffic (Rieser, 2010). This form of simulation examines the trade-off made by individuals in an attempt to punctually keep to their schedules.

The process allows individuals to gain utility by performing the activities mentioned and lose utility for negative effects like performing these activities unpunctually, the time spend travelling, the cost of travelling ect. (Fourie, 2009). This enables individuals (or agents) to evaluate the success of their day by measuring their total utility. Agents can then be allowed to adapt their behaviour in order to gain utility leading to a realistic traffic simulation.

The main problem with ABS is their scalability. Even though the simulations are based on simple concepts the complexity arises from the sheer number of agents being simulated (Rieser, 2010). This complexity creates problems in areas such as memory consumption and computing time. As a possible solution to this problem the ABS is kept as simple as possible while still remain accurate to an acceptable level.

In conclusion ABS not only enables effective transit simulation but has also proven to be more accurate at private car simulation in South African conditions as shown by Fourie (2010)

1.3 The minibus taxi network

Due to past political influences and socio-economic imbalances in South Africa the majority of low-income commuters are situated on the outskirts of the major cities in so-called townships. This provides great challenges on the transit networks in South Africa, due to the distribution of demand for public transport. As a result, a form of paratransit has developed to meet this demand in the form of minibus taxis. Minibus taxis are estimated to make up about 70% of the total transit service in South Africa. This dynamic transit form is largely unregulated, de-centralized and driven by a free market principle (Fourie, 2009).

Paratransit vehicles operate on semi-fixed routes with no fixed time tables. A number of fixed stop locations or taxi ranks are set up by the governing authorities, these ranks are serviced, but other informal stop locations are also put in use. Fares are calculated using a zone system surrounding the formal and informal rank set-up. Taxi routes tend to travel between informal and formal ranks even though the exact route is not fixed according to an interview with R. Marx, Transport Manager at Algoa Bus Company. Fares are calculated depending on the number of routes use in the trip. For example if a passenger travels any distance on the route between Rank A and Rank B the passenger will be charged the fixed fare for that route. However if the passenger wishes to travel beyond Rank B to any location between Rank B and Rank C, the fixed fare for the route between Rank B

and Rank C will be added to the total fare of the passenger.

Passenger can embark and disembark anywhere, in contrast with regular transit services. Potential passengers can flag down a passing taxi, instantly turning the passenger's location into a stop. This complex interaction is impossible to translate into analytical terms and to be fed into a traditional demand planning model. An ABS is therefore the only paradigm capable of modelling paratransit (Fourie, 2009).

1.4 Understanding a transit passenger

In order to understand a transit system, the behaviour of commuters should be studied. By understanding the needs of the commuter population and combining the user needs with the services delivered by the transit operators in a certain area an understanding is developed of the workings of the transit system in that area.

The main needs of transit passengers according to the Comprehensive Integrated Transport Plan (CITP), by SSI Engineers and Environmental Consultants (2011), are as follows:

Travel Time The times it takes for the transit user to reach the destination plays an important role in the choice of transit the user would make. This travel time is also combined with the simplicity of a certain mode of transit. If a passenger has to travel a greater distance it is shown that the user would prefer a mode of transit that travels directly to the user's destination over a mode of transport that requires a changeover, even if the latter is faster than the former.

Accessibility A transit user would tend to employ the transit provider that runs closer to the user's location and destination. The user would tend to chose the provider that requires less walking to reach and that can deliver the user as close as possible to the user's destination. The accessibility of the transit provider also effects the total travel time of the user.

Reliability Since transit users are mostly low-income workers an user cannot afford to be left waiting dueto non-performance of the transit provider. The reliability of the transit provider is therefore also decisive in the choice of the mode of transport.

Comfort and safety A transit user will tend to use the mode of transit that is the most comfortable and the safest. This also relates to the perceived danger and not necessary the real danger.

Cost of travel An additional parameter need that wasn't specified in the CITP study is the question of affordability. As mentioned the average transit user is a low-income worker and is greatly affected by changes in transportation pricing. The effect of changes in transit fares are illustrated using the data from the Algoa Bus Company in the table below. The table shows the total daily number of trips by passengers on the vehicles of Algoa Bus Company in the Nelson Mandela Metropolitan Municipality in South Africa. At the start of February paratransit fares increased causing the dramatic increase in bus passenger numbers as indicated by the data below.

Day	Month	Number
30	01	68198
06	02	75341
13	02	79357
20	02	79398
27	02	80396
06	03	81633
13	03	80836
20	03	80928
27	03	80909

1.5 Research Design and Methodology

This project attempts to solve these problems by the following means:

1. An investigation of the particular challenges facing South African commuters and the needs that drive their choice of transportation.
2. An investigation of the requirements for a practical implementation of an ABS under South African conditions such as: paratransit and the norm of not using public transport if private transport is available.
3. A discussion of a possible ABS) toolkit (The Multi-Agent Transport Simulation toolkit) and the methods required to implement this toolkit.
4. An in-depth analysis of the methods involved in building and adapting a MATSim transit simulation.
5. An initial simulation of an example network to illustrate the possibility of including paratransit simulation.
6. A study of the data gathering and conversion processes involved in the simulation of a South African city using the Nelson Mandela Bay Metropolitan Municipality as a case study.

7. An analysis of the fare structures involved in paratransit networks to enable greater accuracy in future paratransit simulations.

1.6 Outline of project

The following chapter provides greater detail from the literature on the relevancy of the MATSim ABS as well as the workings of a MATSim model and requirements of building a MATSim simulation. Chapter 3 investigates the relevant processes in data gathering and data conversion in order to enable a simulation. Chapter 4 discusses the process of adapting the MATSim model to provide results that are relevant under South African conditions using an example transit network simulation. Chapter 5 evaluates the results obtained from running the adaptations of the example network and investigates the feasibility of including bus and paratransit services into a simulation. Chapter 6 investigates possible future developments in South African transit simulation.

Chapter 2

Literature Study

The Multi Agent Transport Simulation (MATSim) toolkit MATSim Development Team (2011) is a popular ABS toolkit that provides the tools required to perform large-scale transport simulations (Balmer et al., 2008). MATSim is an open-source project and its source-code is freely available under the GNU Public License.

MATSim is currently used to simulate the private vehicle traffic for the Gauteng province in South Africa (Fourie, 2009). In his study Fourie compares the performance of MATSim versus the traditional FSM. By comparing simulated traffic counts to actual traffic counts MATSim delivered a better prediction of traffic volumes over the course of a day.

2.1 The MATSim process

The MATSim toolkit simulates a traffic system by incorporating the many decisions that agents have to make on a daily basis. Balmer (2007) identifies these decisions as follows:

- **Mode choice:** Should I take the bus or the train?
- **Route choice:** What is the best route to take?
- **Location choice:** Should I go shopping near my home or at the mall?
- **Activity type choice:** Should I go shopping today
- **Activity chain choice:** Should I go shopping before or after work?
- **Activity starting time choice:** When should I do sports today?
- **Activity duration choice:** Should I swim for an hour or two today?
- **Group composition choice:** Should we carpool?

The result of these decision is the agent's "plan" An example of such a day plan, encoded in a MATSim specific XML-format is shown below.

```
<person id="135" employed="yes">
<plan selected="yes">
<act type="h" link="1112" x="792.0" y="1020.0" end_time="07:02:15" />
<leg mode="pt">
</leg>
<act type="w" link="1323" x="1597.0" y="3105.0" end_time="17:40:05" />
<leg mode="pt">
</leg>
<act type="h" link="1112" x="792.0" y="1020.0" />
</plan>
</person>
```

All agents then attempt to execute their plans simultaneously within the limitation of the system. Rieser (2010) identifies some of the limitation an agent might encounter:

- **Speed limits:** Agents cannot drive faster than the limit of its vehicle or the road limit
- **Flow capacity of road:** A vehicle cannot drive through a crossing or section of road at the same time as 30 other vehicles
- **Storage capacity of road:** A vehicle cannot occupy a section of road if the road is already occupied by other vehicles.
- **Position of other agents:** In most cases it is not desired to simulate traffic accidents, so vehicles cannot crash into each other.
- **Right of way:** No matter if a crossing has traffic lights or not, an agent cannot just drive on, right of way is applicable.
- **Vehicle capacity:** An agent cannot board a transit vehicle that is already full (exceptions exists with minibus taxis that often exceed their legal capacity)
- **Opening times:** An agent cannot do shopping if the shop is closed.

The agents attempt to learn from their experience and adapt their plans to best interact with the limitations. In order to achieve this agents require feedback regarding the success of their plan if the form of utility gained.

MATSim incorporates these principles in its simulation by using the following steps indicated in Figure 2.1 as discussed by Balmer (2007).

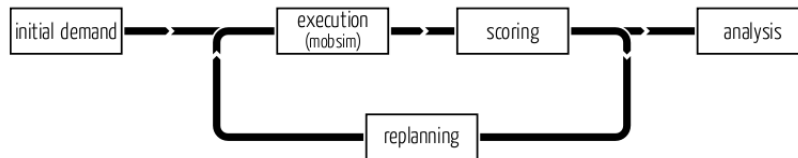


Figure 2.1: MATSim Process cycle

Initial Demand A synthetic population of agents are created. Each agent can hold certain attributes such as gender, age, employment status, car availability. Each agent lives in a virtual world generated using the socio-economic attributes of the study area van der Merwe (2011) and attempts to execute a plan to achieve the activities assigned to that agent.

Execution (mobsim) The plans of all the agents in the simulation are executed simultaneously using a mobility simulation (mobsim). Congestion is caused where multiple agents attempt to perform the same action at the same time. The following mobility simulations are available:

- QueueSimulation
- QSim
- DEQSim
- JDEQSim

However currently only “QSim” is able to accommodate public transport simulation.

Scoring After the execution is completed, each agent evaluates its plan using a utility equation (van der Merwe, 2011) where the positive utility for completing activity i is combined with the negative utility for being late for activity i and the negative utility for travelling to i for all n trips in the day.

Replanning Over a number of iterations agents learn about their environment and are able to adapt their plans to achieve a greater utility. These changes can include departing at different times, using different modes of transport or using different routes.

Over time it becomes increasingly harder for agents to improve their plans and the simulation starts to stabilize. This stable traffic state can be used to simulate and predict traffic counts and agent behaviour.

2.2 Building the simulation

In order to determine the requirements for building a MATSim model a closer look is required at the MATSim process. An ABS consist of many different building blocks (called modules) as shown in figure. In MATSim a so-called “Controller” has the job to combine all the different modules to ensure the correct data is handled at the right time (Rieser, 2010). The Controller however is much more than merely an iteration manager, it offers events to the relevant extensions to show the specific state of the Controller flow. The Controller therefore monitors the simulation and ensures that the correct modules and extensions run at the right times. The events involved in the Controller and the point of execution is shown in Figure 2.2.

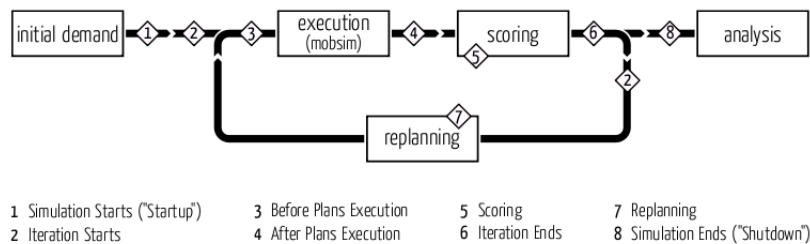


Figure 2.2: Extension points offered by the MATSim Controller

In order to adapt the simulation without writing a new Controller a Configuration file is used to set the rules and parameters of the simulation. The following indicates the type of the parameters that can be set in the Configuration file as illustrated in using a few examples provided by MATSim Development Team (2011):

- Whether public transport is enabled
- What mobility simulation to use
- The start and end time of the simulation
- The format that the Events file should be written in.
- The replanning strategy, the percentage of agents that are allowed to change their plans
- Whether agents are allowed to change their mode of transport

- The paths to the additional input required for public transport simulation

The minimum requirements for the simulation of a public transport system in MATSim as specified by MATSim Development Team (2011) will therefore be:

A Configuration, “config”, file That specifies the parameters of the simulation.

A Multimodal Network A network of links and nodes representing the road and rail infrastructure of the area in question.

A Population A synthetic population of agents each with attributes and a plan.

A Transit Schedule The transit schedules of all the transit providers in the area in question. With stop locations, departure times and routes specified.

A Transit Vehicles file A description of all types of transit vehicles operating in the area including vehicles specifications such as speed, capacity and length. And a list of all vehicles of each type operating in the area in question.

All the requirements should be saved in individual MATSim specific XML-format files and run in MATSim using the standard MATSim Controller.

Chapter 3

Obtaining the required data

The reliability of a simulation is limited by the quality of the data used to perform the simulation. Therefore in order to ensure the effective simulation of a transit system, complete cooperation with regards to data sharing is required of the public transport agencies operational in the area of the simulation model.

3.1 The Multimodal Network

The multimodal network describes the physical transportation infrastructure available to individuals (Rieser, 2010). The network description must include multimodal aspects. Different types of vehicles allowed on each network link need to be classified according to each link. To enable this on existing private vehicle network an additional attribute needs to be included. This attribute will list the vehicle types allowed on each link. It is important to note that more than one vehicle type can be allowed on a specific link, for example cars and buses.

Each link has attributes, describing the physical aspects of the link (for example the length or the number of lanes) as well as traffic related aspects (like the flow-capacity or free speed) (Rieser, 2010). Flow-capacity and physical space is measured in private car equivalents. Any other vehicle can be assigned attributes in term of private car equivalents, causing them to occupy more space on a link.

Obtaining network data is a big challenge when initiating new scenarios. A popular approach is to use data from open sources like Open Street Map (OSM) (Open Street Map Development Team, 2011), which has been taking huge steps to increase the amount of available data in recent years (Rieser, 2010). Initially simulations run using data from OSM, extracted to a road network, has shown that this is not only possible, but actually at least as useful and manageable than using private or commercial data sources. As OSM data includes information regarding transit lines (at least subways,

street cars and trains) in many regions, transit data can also be extracted from OSM. By obtaining the data for the road network and the transit network from the same source simplifies the process of matching transit lines to roads automatically (Rieser, 2010).

3.1.1 Data Breakdown

The multimodal network consists of a series of links and nodes. Simply put, nodes are points where roads intersect or where the properties of the road changes requiring a new link. Each node requires at least the following information:

- A unique id
- The “x” coordinate of the node
- The “y” coordinate of the node

The links are the road sections connecting the nodes. Links are one directional therefore a two-way road requires two links. The following information is commonly allocated to each link:

- A unique id
- The start node of the link
- The end node of the link
- The length in meters of the link
- The capacity in vehicles per hour of the link
- The freespeed of the link in meters per second
- The types of vehicles allowed on the link (modes)
- The number of lanes in the link

When simulating public transport, additional transfer links have to be included at the end of a transit route (Rieser, 2010). In its simplest form this transfer link will connect a certain node with itself, enabling transit vehicles to move from one trip to the next return trip or a different trip. The transfer link also serves the purpose of providing a stop located at the nodes, with a reference link.

3.2 The Population

A synthetic population of agents is generated using the National Household Travel Survey (NHTS). These agents are assigned virtual activities including work and being at home based on the socio-economic attributes of the study area (van der Merwe, 2011). Every agent holds several attributes and a plan that includes at minimum it's activity start times and mode choices throughout the day.

Due to computational restrictions and performance reasons the sample population is scaled down to a fraction of the full population. When scaling down the population, the network also has to be scaled down to prevent inconsistent traffic flows. This is done by reducing the storage and flow capacity of each link down by the same factor as the population (MATSim Development Team, 2011).

In order to simulate a realistic population, the population distribution over each segment of the target area is used in generating a synthetic population. The population distribution of the Nelson Mandela Bay Metropolitan Municipality (NMBMM) is illustrated visually as an example in Figure 3.1 as provided by SSI Engineers and Environmental Consultants (2011).

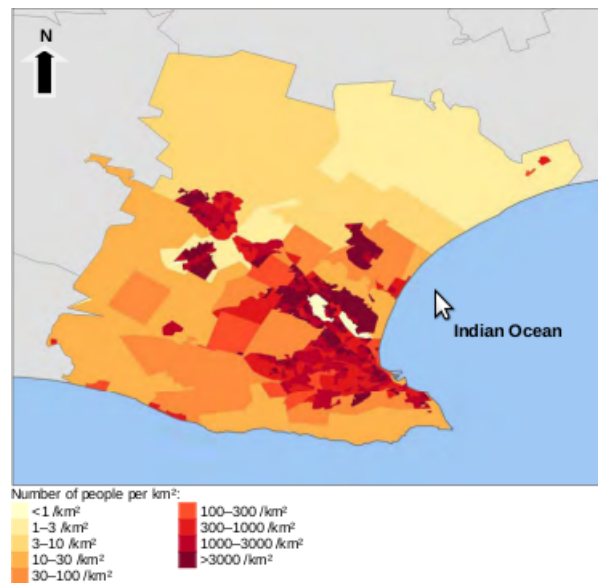


Figure 3.1: The population distribution in Nelson Mandela Bay

3.3 The Transit Schedule

Central to the data structure of simulating transit is the schedules of the transit companies involved (Rieser, 2010). It contains information regarding

one or more transit lines, the detailed routes along these lines, the stop locations used along the route and the departure times of transit vehicles on the route.

3.3.1 Data Breakdown

According to Rieser (2010) the data required for transit schedules consists of the following information:

Stop facilities A description of the locations where passengers can board and alight transit vehicles. Stop facilities are given a unique identifier, a coordinate used for estimating walk distances. Every stop facility is linked to a network link to specify how the stop can be reached by transit vehicles and a transit line indicating the transit service using that stop. The following information will be allocated to each stop at minimum:

- A unique id
- The “x” coordinate of the stop facility
- The “y” coordinate of the stop facility
- The link that the stop is attached to

Some of the additional information that can be allocated to a stop facility is indicated as follows:

- The name of the stop facility
- Whether a transit vehicle at the stop would block traffic

Transit route A description of one specific route of a transit line, including its time profile. The time profile describes the time the transit vehicle plans to take to travel from one stop to another. This information is used to determine the departure and arrival times at each stop location. Each route is described twofold: once as a list of transit route stops, and once as a network route. The latter is to enable the driver agent to travel along the correct links and the former is required by passengers for route planning. The following information will be allocated to each route:

- A unique id
- The transit line serviced
- The mode of transport used
- A list of the stops on the route
- The expected time from the start to arrival at each stop
- The expected time from the start to departure at each stop

- A list of the links used to reach the stops on the route

Departures Different departures specify at which times transit vehicles will depart from the start of the transit route.

3.3.2 Data gathering

One of the main problems with obtaining the transit schedules from different companies, is the lack of a common data-specification. Since two companies or organizations do not necessarily use the same data-structure it can prove difficult for one company to interpret the data output generated by another. Transit companies in general provide the same service, therefore the data from both companies should have similar content and a common data-structure is therefore possible in theory.

A popular common data specification used worldwide is the General Transit Feed Specifications (GTFS). GTFS was initially called the Google Transit Feed Specification. It was developed by Google to enable companies to upload their transit schedules on Google Maps. The value of such a data specification was then realised and the specification was renamed to be used broadly as a data-sharing specification. When using GTFS companies are able to share their transit schedules in a format that can be easily interpreted by other organizations. Companies using the GTFS format can upload their data to the GTFS Data Exchange (www.gtfs-data-exchange.com) for global data sharing and host their transit schedules directly on Google Maps (maps.google.com) providing a user-friendly way for commuters to access the company's transit schedules.

For the purpose of simulation, the transit data of all agencies operating in a area of interest needs to be consolidated. The schedules of South African transit companies are not yet converted to GTFS or any other specification and are mostly found in the form of Excel spreadsheets amongst other formats. To incorporate this data into a simulation, it has to be converted to a common format and stored for later use. To achieve this the University of Pretoria is currently hosting Transit Data Feeder (TDF), an user interface designed for the conversion of transit schedules in any format to the GTFS format. The GTFS format can store a array of extra information that are optional additions. This is also allowed for in TDF but only the mandatory data requirements will be discussed in this project.

In order to upload data in TDF, the transit agency in question needs to be created. The transit data is imported into the following mandatory entities:

1. **Stops** A stop entity requires a stop name and coordinates as mandatory input. TDF automatically provides the stop with a stop id. The coordinates of the stops can be found using an interactive map provide in TDF as shown in Figure 3.2.

2. **Routes** A route entity requires a short name, long name, the route type (car, train or bus ect.) and an additional mandatory set requirement in TDF: a definition to whether bikes are allowed on the route. A route id is assigned automatically by TDF
3. **Service Calendar** The service calendar enables the data to take into account any differences on certain days, for example fewer trips over the weekends or additional trips on Fridays.
4. **Trips** A trip entity is linked to a certain route, a certain service calendar and assigned a trip id by TDF as mandatory requirements. TDF also requires a mandatory requirement: whether bikes are allowed on the trip. The following sub-classes are linked to trips:
 - **Stop times** Stop times include the selected stop name, information regarding whether the transit vehicle will automatically stop at the stop indicated, the arrival and departure times at each stop and the sequence of the stops.

To minimize the work required for importing all the trips, a “Based on” attribute is added to the trip entity. Using the “Based on” entity one can create similar trips with different start times without having to re-enter all the stop times.

- **Frequencies** Frequencies perform a similar function as the “Based on” attribute. One trip can be set to run at a number of frequencies throughout the day. On each frequency the original trip will be repeated with a different start time
5. **Fares** The fare entity in TDF requires a price and a payment method as minimum input. The fares schedule is not mandatory.
 6. **Upcoming Holidays** The holiday or exception entity stores certain calendar days that a certain service will be suspended or activated. For example on a public holiday all normal service will be suspended, but all trips linked to the holiday service will be activated.

The main shortcoming of TDF is its inability to store the shape files for each trip. Shape files store a point-by-point GPS coordinate route for each trip in the schedule. Shape files are not a mandatory requirement in sharing GTFS data, but does provide useful information regarding the routes followed by transit vehicles.

Due to the excess information that can be stored in TDF the Entity Relationship Diagram (ERD) for TDF is highly complex. A recommended minimum requirement ERD stipulating the required data for GTFS sharing and conversion to MATSim specifically in a South African context is included in Appendix A.

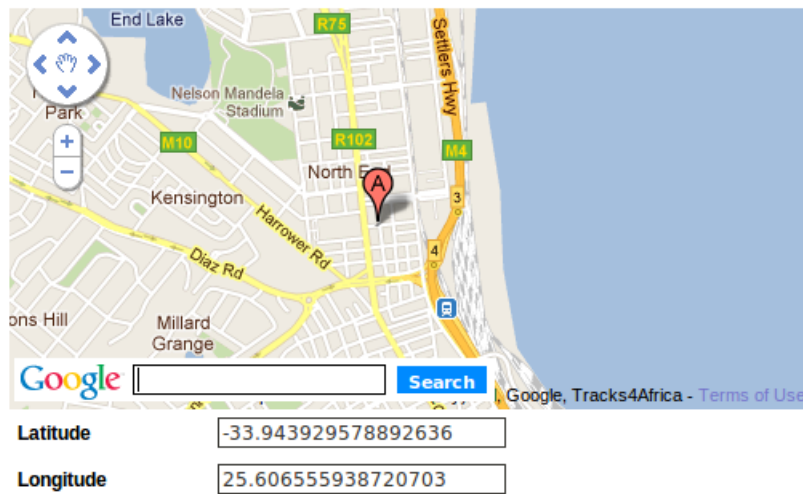


Figure 3.2: Interactive stop coordinate locator

3.4 The Transit Vehicles

All the types of transit vehicles in use in the simulation as well as a list of every specific transit vehicle in use needs to be specified to enable the simulation. When describing the types of vehicle the following information is required as minimum for each vehicle type:

- A unique vehicle type id
- The capacity standing of the vehicle
- The capacity seated of the vehicle
- The length of the vehicle

Chapter 4

Creating a MATSim Transit Simulation

In order to illustrate the process of running a MATSim transit simulation, the example simulation provided by MATSim Development Team (2011) will be investigated. This simulation includes regular private vehicle traffic as well as the inclusion of public transport in the form of a train. The project will then aim to generate an understanding of the simulation by adapting the simulation to also include other transit services such as bus and basic paratransit and by changing the specifications and adapting the configuration to suit South African conditions.

Basic route-based paratransit is possible with only a small amount of changes, namely the introduction of the paratransit routes network into the simulation and a custom time-function for the transit router. Rieser (2010) recommends including a paratransit stop on each link of the paratransit route to simulated the irregular stopping pattern of a paratransit vehicle. In order to simulate paratransit not restricted to fixed routes, much higher effort will be required (Rieser, 2010).

4.1 The example public transport simulation

The example simulation is based on the multimodal network shown in Figure 4.1. Two rail links are included, with a “modes” attribute set to “pt”. These links are the links that connect points (1), (2) and (3). All other links are considered to be for private vehicle use only, the “modes” attribute will therefore be set to “car”. The example simulation at default runs for 50 iterations before completion.

Each node is assigned a id, x-coordinate and y-coordinate as shown.

<nodes>

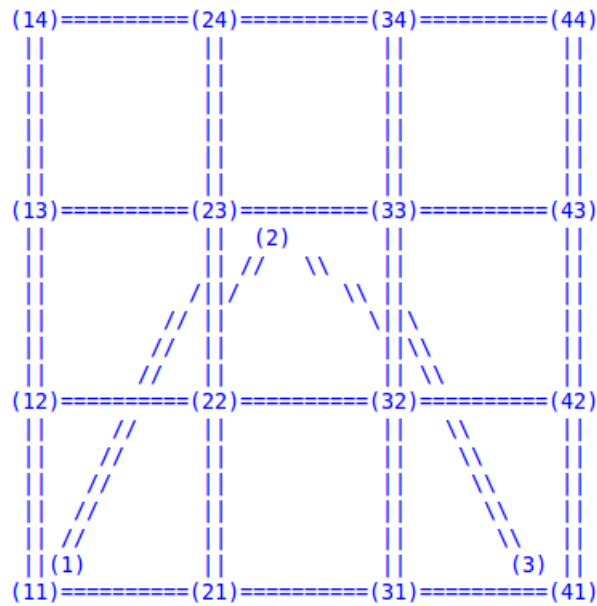


Figure 4.1: Example transit simulation network

```

<node id="1" x="1050" y="1050" />
<node id="2" x="2050" y="2950" />
<node id="3" x="3950" y="1050" />
ect...
<nodes>

```

The links of the example network are assigned either “car” or “pt” modes as discussed above. A section of the links in the example network is shown as provided by MATSim Development Team (2011). Notice the transfer link “11” that connects node 1 to node 1.

```

<links>
  <link id="11" from="1" to="1" length="100.00" capacity="2000"
freespeed="100" modes="train" permlanes="1" />
  <link id="12" from="1" to="2" length="2400.00" capacity="2000"
freespeed="12" modes="train" permlanes="1" />
  <link id="4344" from="43" to="44" length="1200.00" capacity="2000"
freespeed="12" modes="car" permlanes="1" />
  <link id="4443" from="44" to="43" length="1200.00" capacity="2000"
freespeed="12" modes="car" permlanes="1" />
</links>

```

The example network currently utilizes two vehicles of one type of transit vehicle. The specifications of the vehicle type “small train” utilized in this simulation is shown below.

```
<vehicleType id="1">
  <description>Small Train</description>
  <capacity>
    <seats persons="50"/>
    <standingRoom persons="30"/>
  </capacity>
  <length meter="50.0"/>
</vehicleType>
<vehicle id="tr_1" type="1"/>
<vehicle id="tr_2" type="1"/>
```

The example network use three stop locations in transit schedules, with a stop location in each direction added in at node 2. The stop locations used in the example network are shown below.

```
<transitStops>
  <stopFacility id="1" x="1050" y="1050" linkRefId= "11"/>
  <stopFacility id="2a" x="2050" y="2940" linkRefId= "12"/>
  <stopFacility id="2b" x="2050" y="2960" linkRefId= "32"/>
  <stopFacility id="3" x="3950" y="1050" linkRefId= "33"/>
</transitStops>
```

In the example network two routes run along the same transit line. The first route starting at node 1 and ending at node 3 and the second route representing the return trip of the first route by starting at node 3 and ending at node 1. The information regarding route one and the first two departure times are shown below.

```
<transitLine id="Blue Line">
  <transitRoute id="1to3">
    <transportMode>train</transportMode>
    <routeProfile>
      <stop refId="1" departureOffset="00:00:00"/>
      <stop refId="2a" arrivalOffset="00:03:20">
```

```

departureOffset="00:04:00"/>
    <stop refId="3" arrivalOffset="00:09:00" />
</routeProfile>
<route>
    <link refId="11"/>
    <link refId="12"/>
    <link refId="23"/>
    <link refId="33"/>
</route>
<departures>
    <departure id="01" departureTime="06:00:00" vehicleRefId="tr_1" />
    <departure id="02" departureTime="06:15:00" vehicleRefId="tr_2" />

```

The example simulation includes a config file with the minimum required parameters for simulating a public transport system. The complete minimum requirement config file is shown in Appendix A.

4.2 Possible adaptations to the example transit simulation

In order to adapt the example simulation to suit real-life South African conditions, the following changes can be made to each file of the input:

Config file changes In order to prevent private vehicle users to change to public transport and vice versa the “modes” parameter in the “changeLegModes” module in the config file can be changed from “car,pt” to “pt”. This indicates to MATSim that the only mode change allowed is public transport users changing between other modes of public transport, therefore the interchanging between private and public transportation is effectively disabled.

```

<module name="changeLegModes">
<param name="modes" value="pt" />
</module>

```

When many different forms of public transport are employed it is also recommended to increase the maximum number of plans an agent may store. Due to the increased number of transportation options this is required to improve the agent’s chance of selecting a good plan. In this investigation the “maxAgentPlanMemorySize” will be increased from 5 to 10.

```
<param name="maxAgentPlanMemorySize" value="10" />
```

Multimodal network changes In order to include a basic bus network in the simulation, the following changes are required to the network:

1. The links along the chosen bus routes must be adapted to include “bus” as a accepted mode of transport.

```
<link id="1121" from="11" to="21" length="1200.00" capacity="2000"
freespeed="12" modes="car,bus" permlanes="1" />
\end{enumerate}
```

2. If bus stops are located at a node in the network a transfer link needs to be added for that node. This is not required if the stop is located on a link.

```
<link id="1111" from="11" to="11" length="100.00" capacity="2000"
freespeed="100" modes="bus" permlanes="1" />
```

For the implementation of a paratransit network it is recommended to initially implement a fixed-route paratransit network. The same procedure will therefore be applied to implement a paratransit network as for a bus network with regards to the network.

Population changes Currently the population exists of 900 agents each with a complete daily plan. The synthetic population provided can be enlarged by adding more agents and be expanded to include information regarding whether each specific agent has access to a car

Expanding the transit schedule In order to add a bus or paratransit schedule to the transit schedule used in the simulation, the following changes are required to the transit schedule:

1. The required stop facilities should be added to transitStops. It should be noted that, with the exception of stops at the start and end of the route, stops are directional. Therefore two stops have to be added to each stop location except at the ends of the bus or paratransit route. Stops located at a node should use the transit link added to the network file as the “linkRefId”

```
<stopFacility id="11" x="1000" y="1000" linkRefId= "1111"/>
<stopFacility id="41" x="4000" y="1000" linkRefId= "4141"/>
```

2. A transit line should be created for each additional transit network.

```
<transitLine id="Green Line">
```

3. All routes required and the mode of transit of the transit line should be defined.

```
<transitRoute id="bus1">  
  <transportMode>bus</transportMode>
```

4. A stop route plan, including travel times to each stop and a link route plan needs to be added for each route defined.

```
<routeProfile>  
  <stop refId="11" departureOffset="00:00:00"/>  
  <stop refId="14" arrivalOffset="00:03:20" departureOffset="00:04:00"/>  
  <stop refId="44" arrivalOffset="00:07:20" departureOffset="00:08:00"/>  
  <stop refId="41" arrivalOffset="00:11:20" departureOffset="00:12:00"/>  
  <stop refId="11" arrivalOffset="00:15:20" />  
</routeProfile>  
<route>  
  <link refId="1111"/>  
  <link refId="1112"/>  
  <link refId="1213"/>  
  ect...
```

5. Additional information that can be included optionally is a stop name and a “isBlocking” attribute. This attribute is used to determine if the transit vehicle blocks the road when at a stop facility. For the sake of this investigation it will be assumed that paratransit vehicles do not block the road and buses do.

```
<stopFacility id="41" x="4000" y="1000" linkRefId= "4141"  
  name="2nd Busstop" isBlocking="true"/>  
<stopFacility id="01" x="1000" y="1100" linkRefId= "1112"  
  name="1st Street" isBlocking="false"/>
```

Adding additional transit vehicles Additional transit vehicle types have to be included for each additional mode of transport added.

```
<vehicleType id="2">
<description>Small Bus</description>
<capacity>
<seats persons="20"/>
<standingRoom persons="10"/>
</capacity>
<length meter="13.0"/>
</vehicleType>
<vehicleType id="3">
<description>Minibus Taxi</description>
<capacity>
<seats persons="10"/>
<standingRoom persons="0"/>
</capacity>
<length meter="4.0"/>
</vehicleType>
```

The additional vehicle types should then be included in the simulation as vehicles.

```
<vehicle id="tr_1" type="1"/>
<vehicle id="tr_2" type="1"/>
<vehicle id="bs_1" type="2"/>
<vehicle id="bs_2" type="2"/>
<vehicle id="mt_1" type="3"/>
<vehicle id="mt_2" type="3"/>
```

Chapter 5

Running the example transit simulation

This chapter analysis the workings of the MATSim transit simulation toolkit by comparing the results obtained from running the standard example transit simulation with results obtained from simulations adapted using some of the methods indicated in the previous chapter.

5.1 The standard example transit simulation

The standard example transit simulation was executed and will be used as a baseline for the adapted simulations. The resulting utility score over each iteration is illustrated in Figure 5.1 and the total distance travelled by agents for each iteration is illustrated in Figure 5.2.

5.2 Implementing a basic bus network

A basic bus network was implemented into the simulation by adding the following changes to the simulation:

- The interchanging between private and public transport was disabled in the config file.
- The maximum amount of plans stored by an agent was increased to 10 in the config file.
- A transfer link was implemented at node 11 and node 41.
- The relevant links as illustrated in green in Figure 5.5 where updated to include “bus” as an accepted mode in the multimodal network file.
- Stops where added at node 11 and node 41 in transit schedule file, the “isBlocking” attribute was set to “true”.

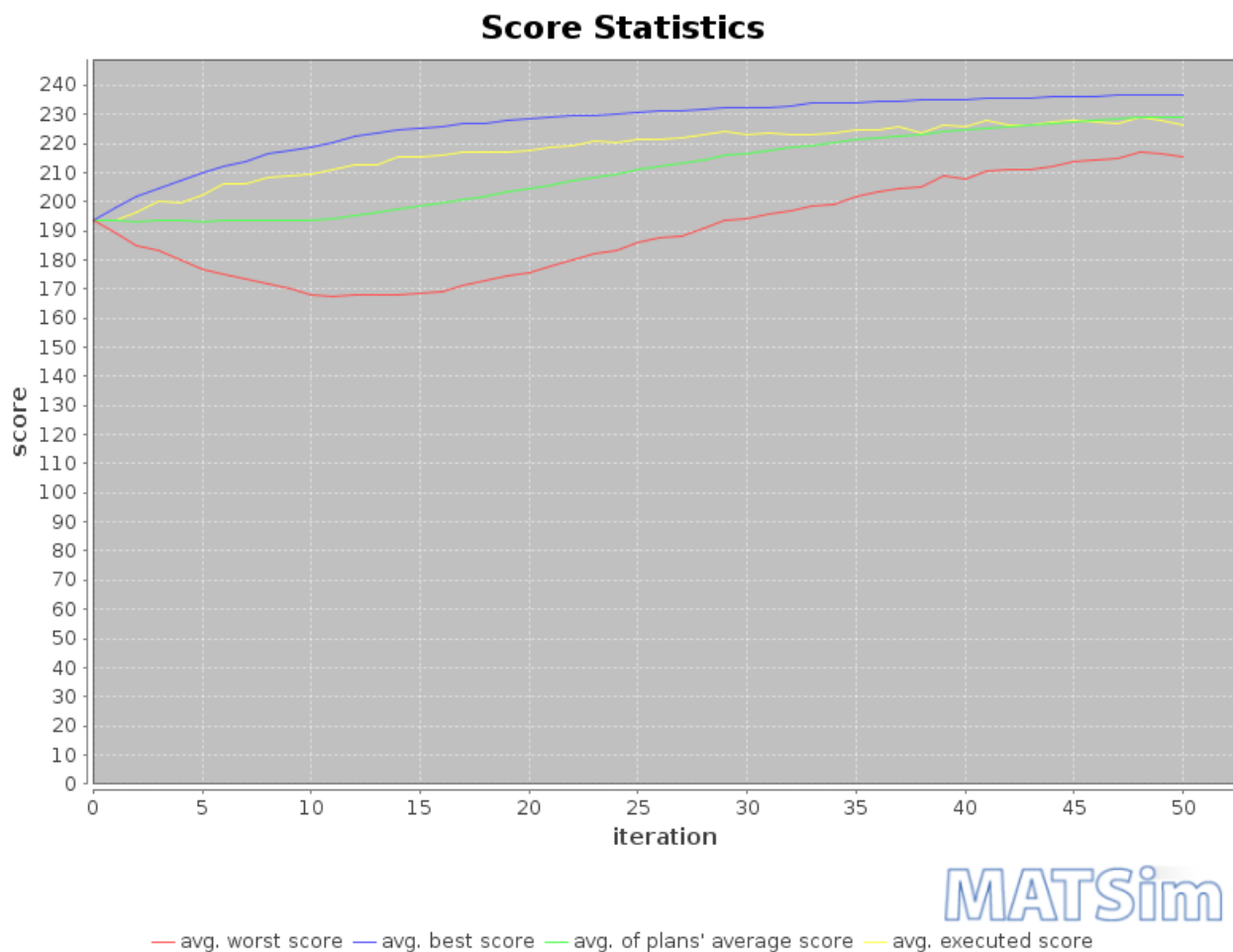


Figure 5.1: Utility score achieved per iteration, standard model

- A new transit line “Green Line” was created with two trips, bus1 and a return trip bus2. Transit mode was set to “bus”.
- Planned travelling time between the two stops was set at 5min, the route indicated in Figure 5.5 was specified as the route description.
- 26 departures were spread throughout the day, with additional departures over peak periods.
- The bus vehicle type was included in the transit vehicles file with a seated capacity of 20, standing capacity of 10 and length of 13m.
- Two vehicles of the bus type were created in the transit vehicles file.

The resulting utility score over each iteration and the total distance travelled by agents for each iteration obtained from the expanded simulation

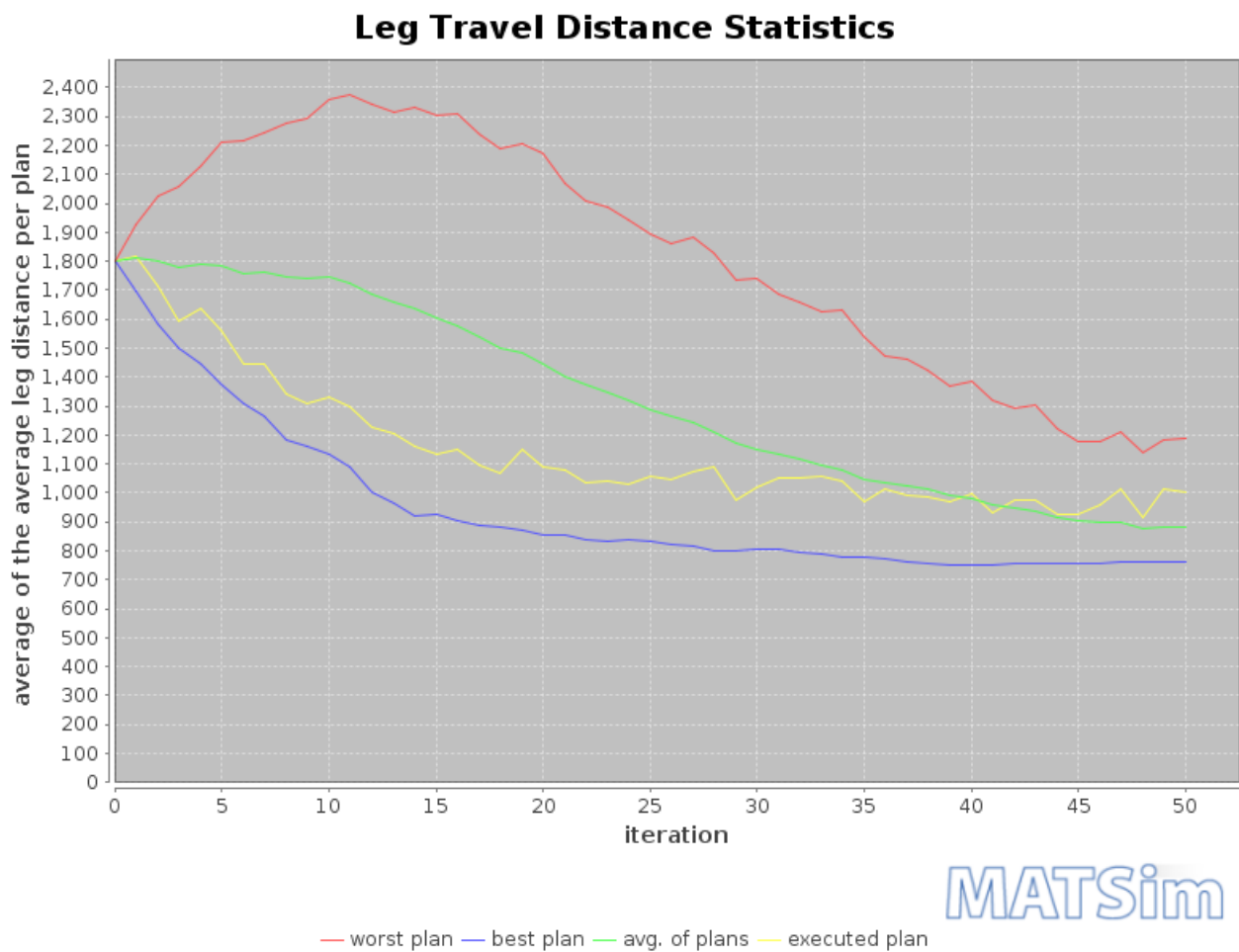


Figure 5.2: Total distance travelled per iteration, standard model

is illustrated in Figure 5.3 and 5.4 respectively.

As indicated by a study of the results, agents in the adapted simulation achieved a similar score to the standard simulation. The events file of the final iteration of the adapted simulation does show that the bus vehicles did travel along the route and passenger did use the bus network as transportation. The adapted simulation is therefore considered a success.

5.3 Including paratransit for the simulation of a full transit system

Paratransit was added to the simulation along with the bus network already included, using the same methods as for the bus network. For a basic paratransit simulation the paratransit system is treated as a bus network, with

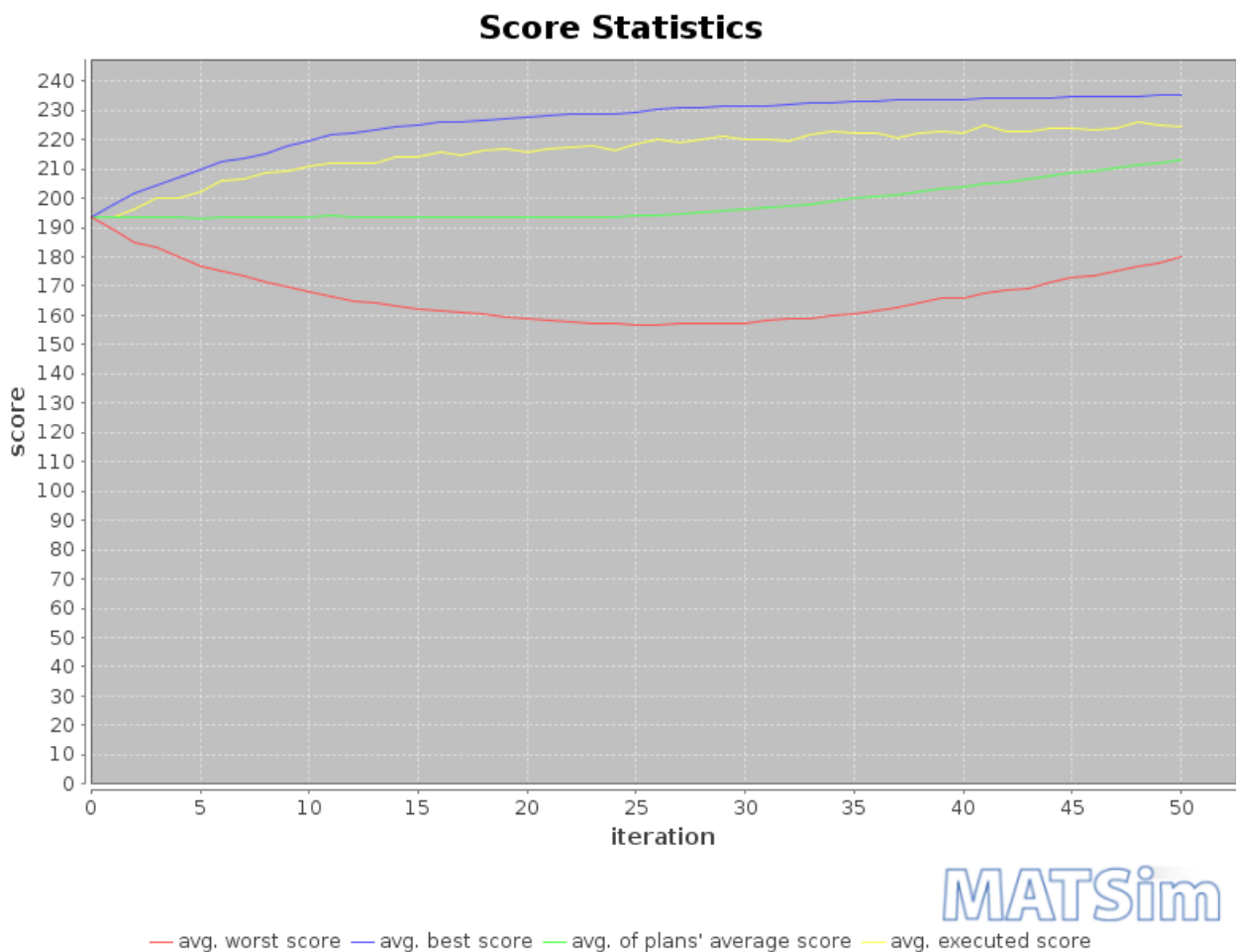


Figure 5.3: Utility score achieved per iteration, adapted model

mode set to “taxi”, with a few exceptions to mimic the behaviour of paratransit vehicles. The exceptions used to model paratransit are the following:

- Paratransit vehicles were given a seated capacity of 10, a standing capacity of 0 and a length of 4m
- Paratransit stop facilities are located on each link in the paratransit route
- “arrivalOfSet” and “departureOfSet” is equal for all stop facilities, causing the paratransit vehicle only to stop for the time required to load and offload passengers
- Paratransit departure times heavily favour peak times causing paratransit vehicles to be inactive for most of the off-peak periods

Leg Travel Distance Statistics

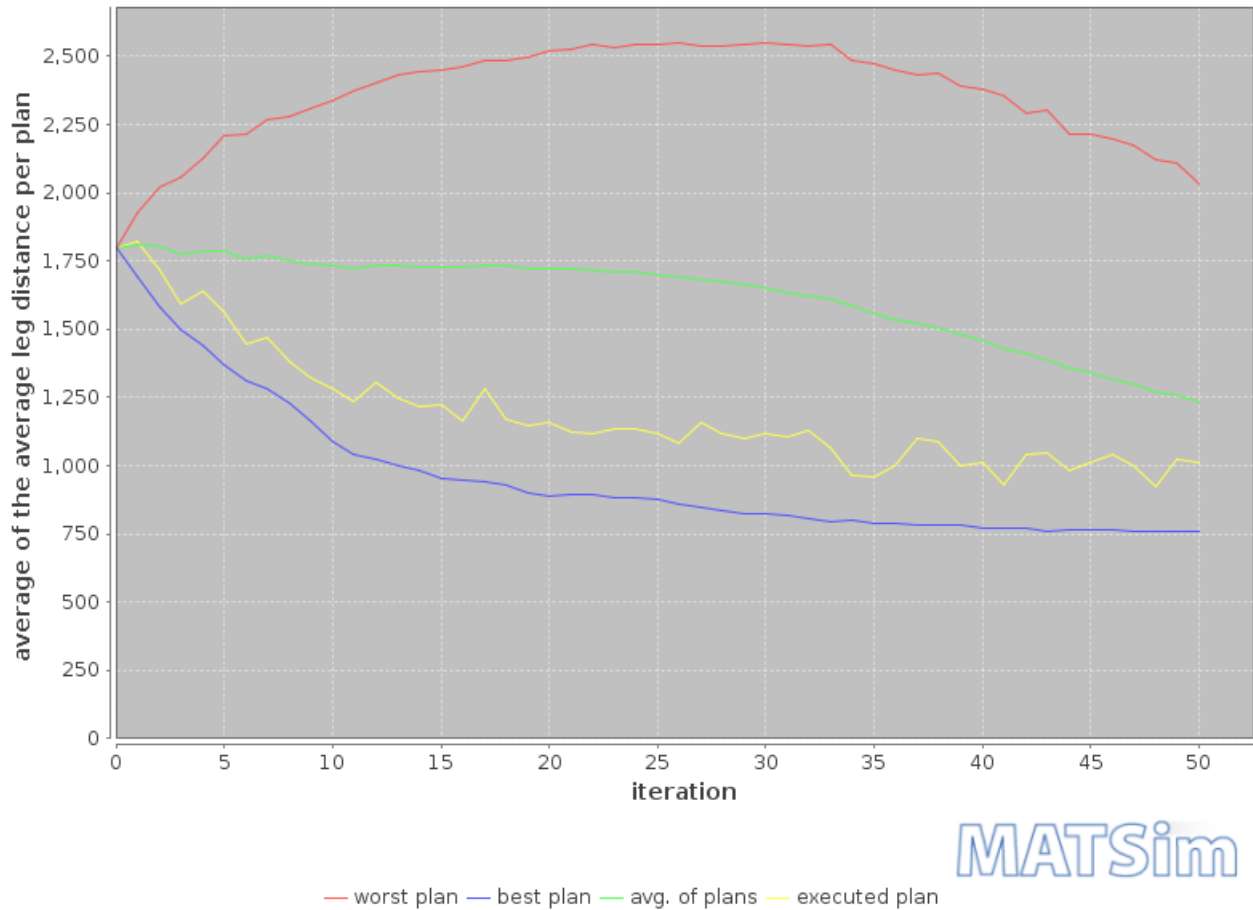


Figure 5.4: Total travel distance per iteration, adapted model

- Paratransit vehicles are set to operate at full capacity from 06:00 to 09:00 and 16:00 to 20:00
- Paratransit vehicles are set to operate at a reduced capacity from 09:00 to 16:00
- Paratransit vehicles are set not to operate after 20:00
- Planned travelling time is set at one minute per link
- The paratransit route is indicated in red in Figure 5.6

The resulting full transit simulation delivered the following results as illustrated in Figure 5.7 and Figure 5.8.

As shown, agents in the full simulation performed a lot worse than in previous simulations. The public transit leg histogram of the final iteration

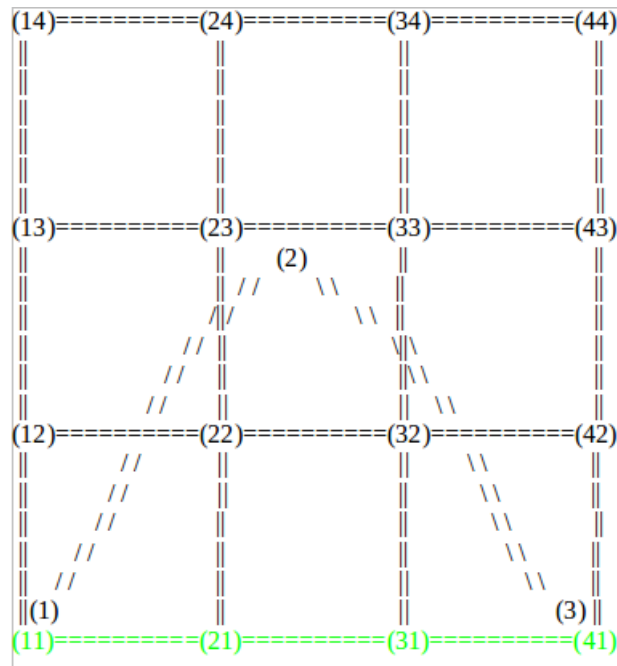


Figure 5.5: Basic bus implementation route

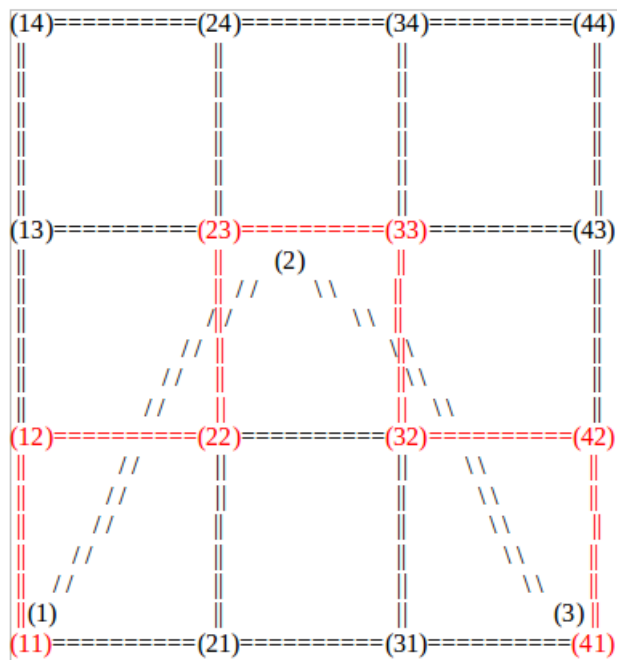


Figure 5.6: Basic paratransit implementation route

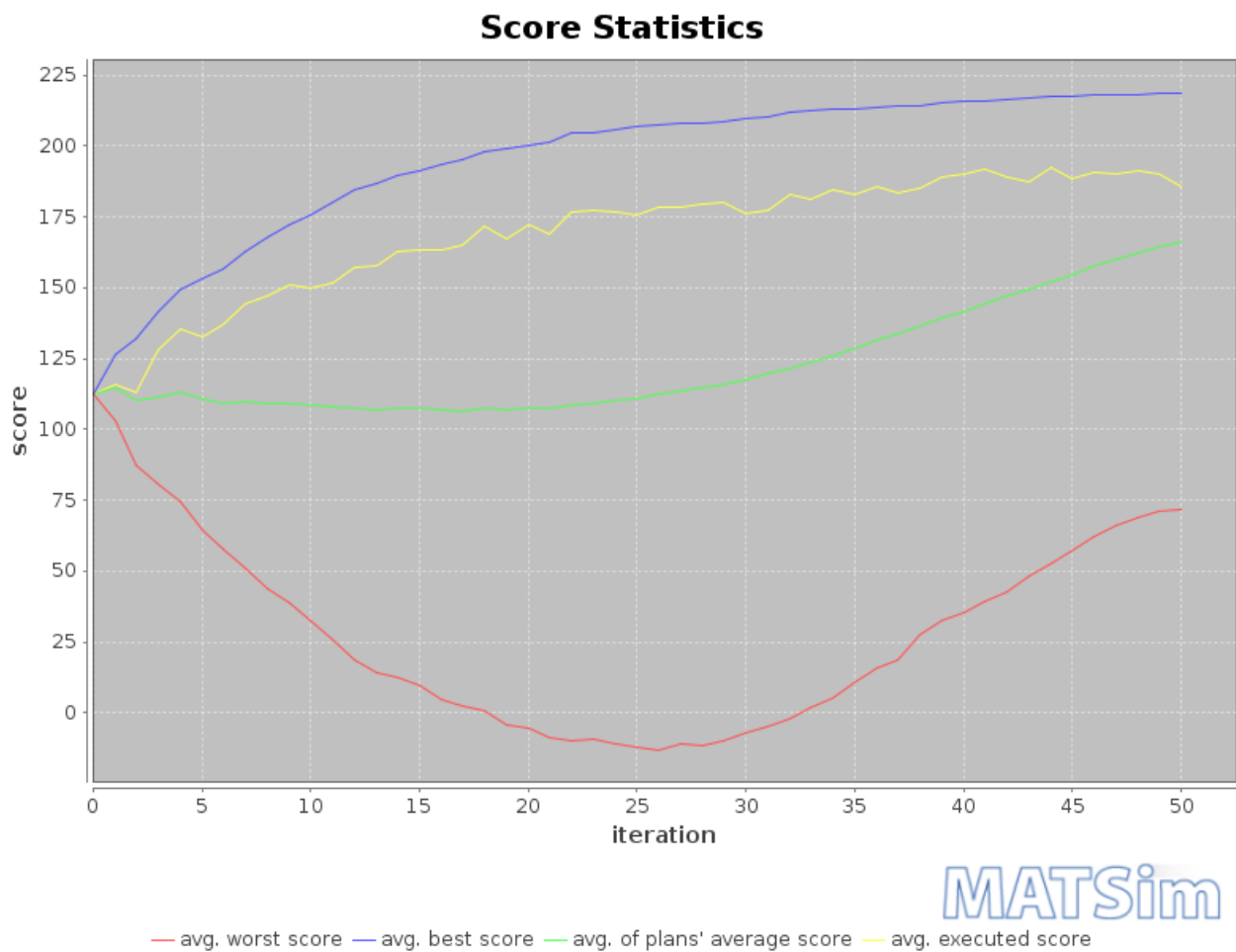


Figure 5.7: Utility score achieved per iteration, full model

shown in Figure 5.9 seems to indicate this might be due to a large amount of paratransit users being stranded during off-peak times. To illustrate this point this figure is compared to the corresponding histogram in the adapted (bus only) simulation shown in Figure 5.10. Notice how passengers enroute during non-peak hours are a lot more in the simulation that includes paratransit. The most obvious difference is the greatly increased number of passengers still en route after 20:00 at night.

Leg Travel Distance Statistics

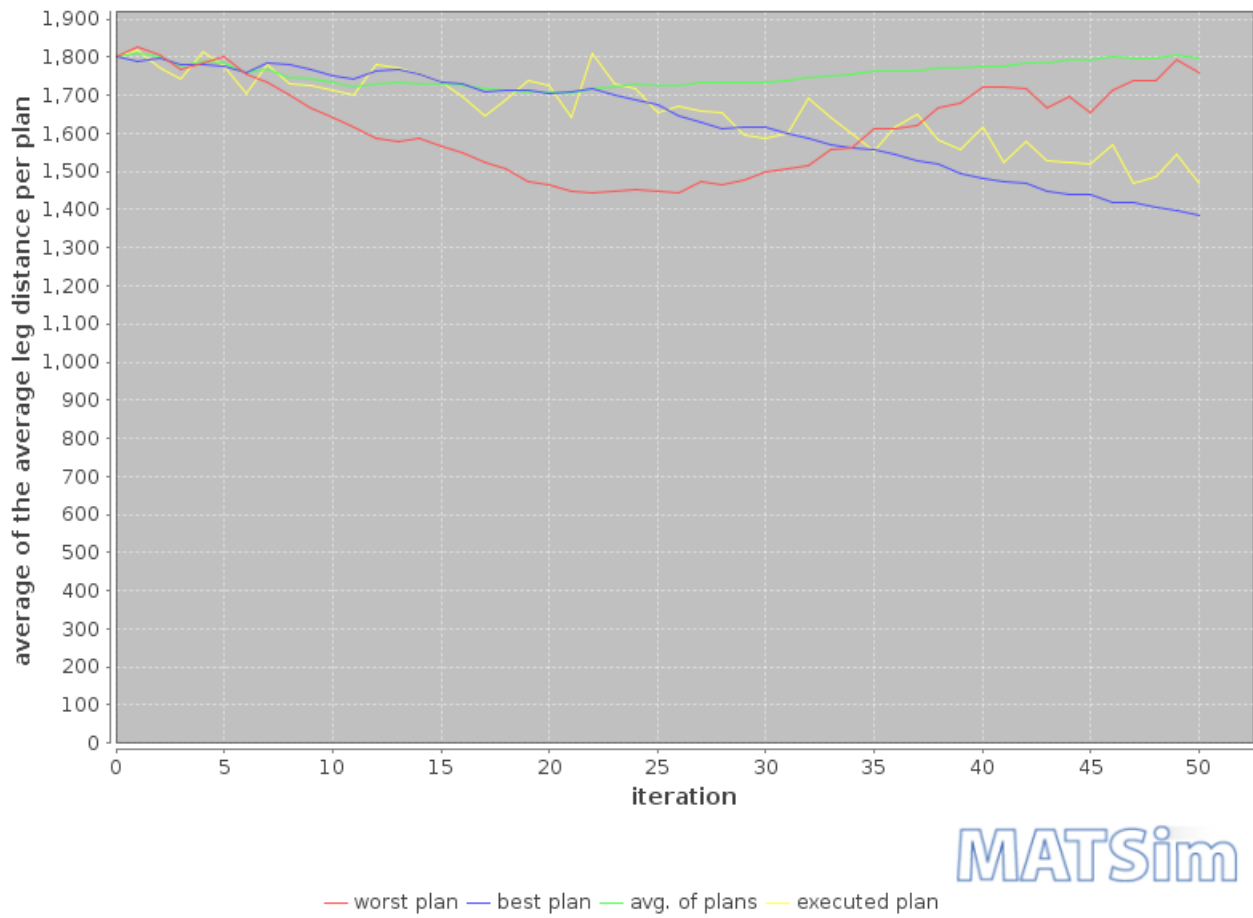


Figure 5.8: Total travel distance per iteration, full model

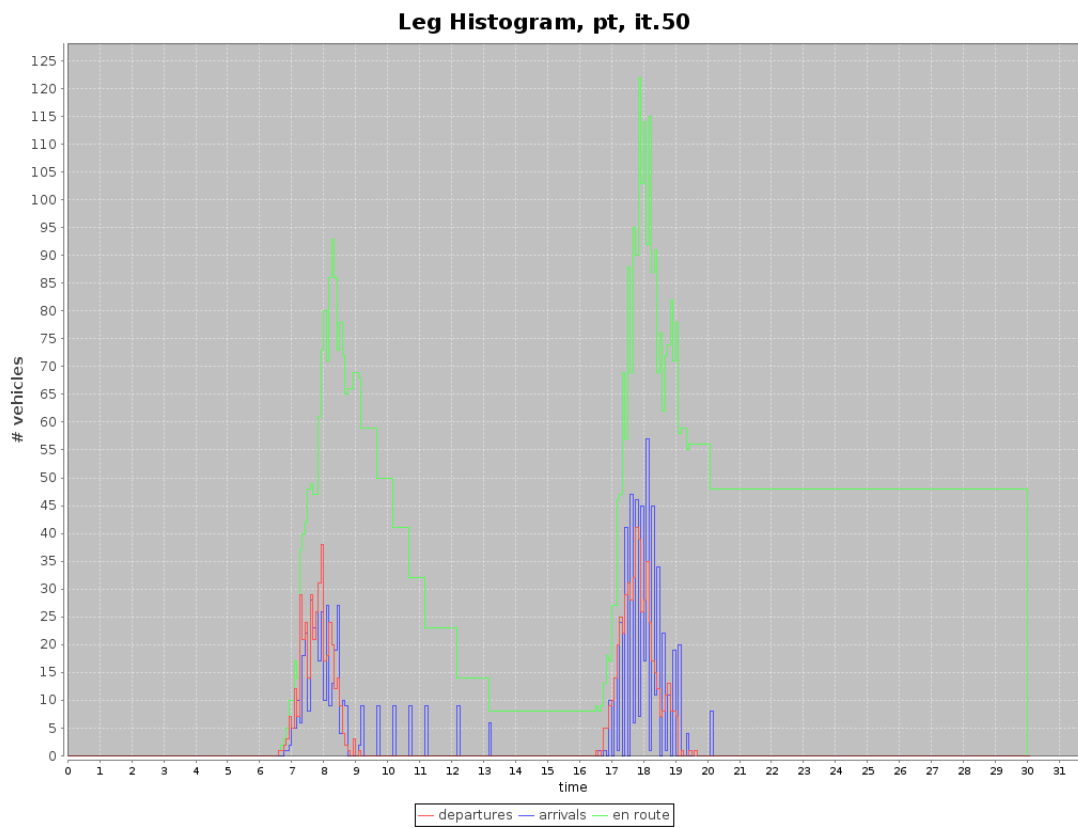


Figure 5.9: Public transit leg histogram (it. 50) of the full transit simulation

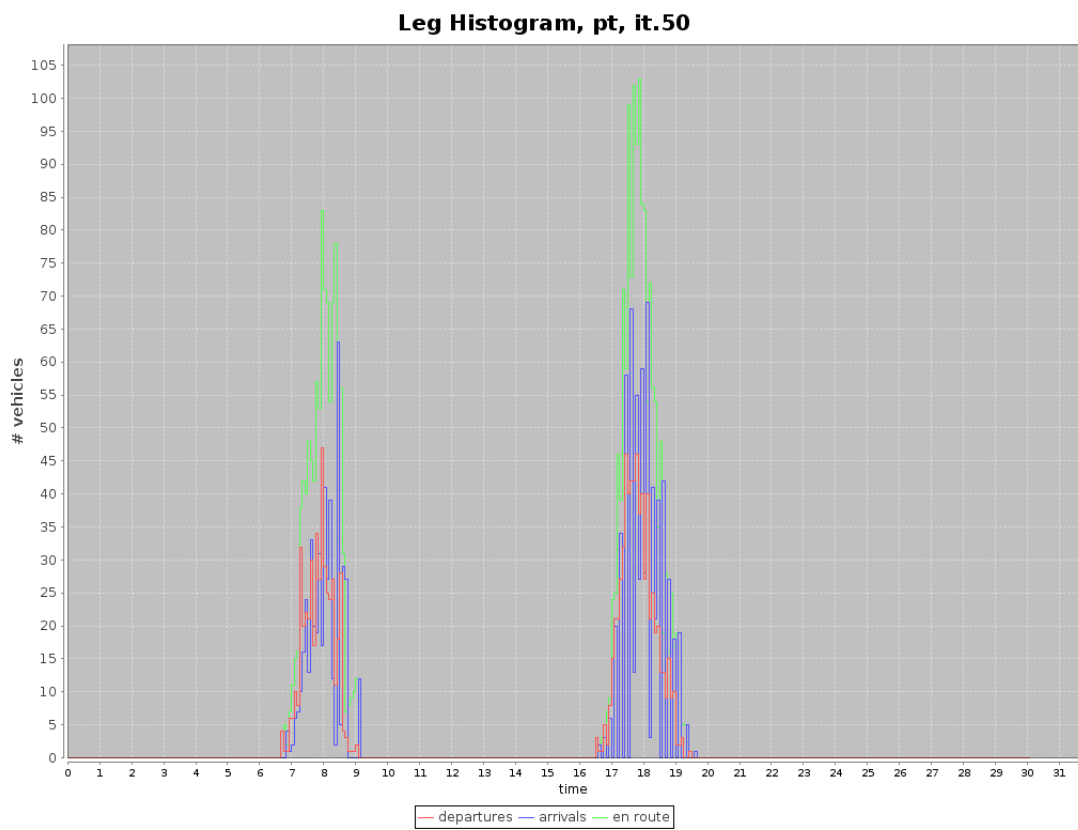


Figure 5.10: Public transit leg histogram (it. 50) of the adapted transit simulation

Chapter 6

Future development

In South African transit systems close competition exists between the paratransit and bus network according to information obtained from an interview with R. Marx, Transport Manager at Algoa Bus Company (2011). In order to simulate the behaviour of passengers in these competitive transit systems, the logic that drives the behaviour of passenger can be implemented to predict the effect of changes in the provided service.

6.1 Zone-based paratransit fare algorithm

One such change in service is a change of transit fares. In order to determine the effect of such a fare change on a specific passenger's travel cost an algorithm can be introduced to convert transit fares to travel cost per distance or per zone.

According to R. Marx (2011), paratransit fares in the NMBMM are calculated according to the number of taxi-ranks involved in the trip. A trip between two ranks is assigned the fare of travel between those two ranks and if a trip goes beyond a rank, a second fare is added onto the total fare. When comparing bus fares, usually calculated by route, to paratransit fares it was found that the total transit cost for a passenger on a single fare transit trip tend to be less than the equivalent bus trip. However the total cost for a passenger travelling a paratransit trip involving multiple fares tend to be more than the equivalent bus fare as according to R. Marx (2011). The feasibility of this fare comparison in other parts of South African is yet to be confirmed. A summary of the results of a fare comparison survey done by the Algoa Bus Company is shown in Appendix C.

In order to incorporate this fare comparison into an algorithm, the implementation of a trip-based fare structure will be attempted. By connecting the locations of taxi ranks in a specific area a series of overlapping zones are formed. Theoretically an agent travelling in one zone will prefer to use paratransit from an economical point of view. Should the specific agent

require transportation beyond the agent's zone of origin, the agent would prefer the bus over paratransit.

To illustrate this theory the location of the mayor ranks in the NMBMM was provided by R. Marx (2011) as indicated in the table below.

No	Name	Address
1	Brister House	190 Govan Mbeki Ave
2	Cleary Park	Cnr Standford and Rensburg
3	Korsten	42 Highfield Rd
4	Motherwell Shopping Centre	Cnr W.M. Maku and Ngonyama
5	Ntoli Square	Cnr Njoli and Daku
6	Uitenhage terminus	59 Durban st
7	Dwesi Complex	Cnr Mkwenkwe and Tofile
8	Green Acres	8 Ring Rd

By plotting the mayor ranks onto a map of the NMBMM the area can be divided into fare zones. It should be noted for a accurate zone configuration all of the formal and informal taxi ranks employed in the NMBMM will be required. The resulting zone breakdown is indicated in Figure 6.1. By employing the zone-based costing method the most economical option between paratransit and bus transit can be determined. If the zone id of the origin equals the zone id of the destination, paratransit will be the most economical option. Otherwise if the zone id of the origin does not equal the zone id of the destination, bus transit will be the most economical option.

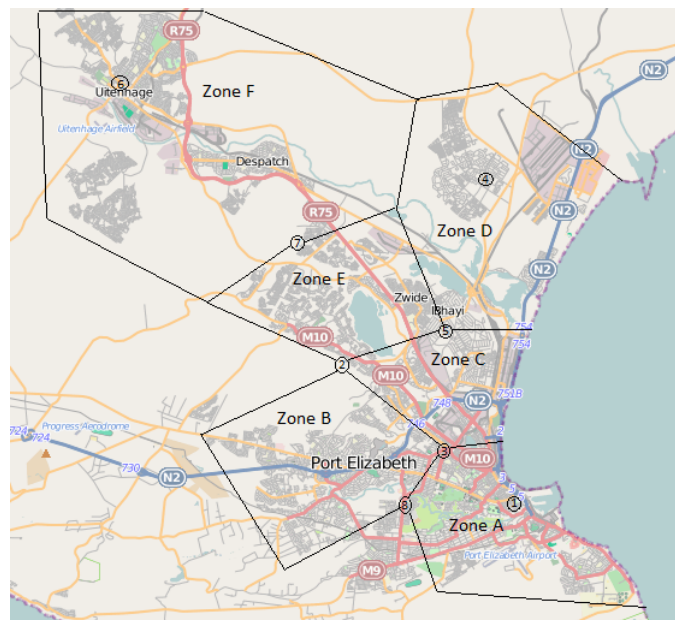


Figure 6.1: Zone breakdown for the application of a zone-based cost equation

6.2 The Nelson Mandela Bay case study

This project has investigated the requirements for a full-scale MATSim simulation. This section outlines the progress currently made towards the aim of simulating a South African city as a case study.

The NMBMM was selected as the case study for this project because of its isolation from the rest of South Africa. This area provides a transportation system with little influence from outside the area which creates a closed system transit network that simplifies the simulation of the transport system. Furthermore, a private car transport model was created by the Integrated Planning Development Modelling (IPDM) for the municipality making it the preferred choice for the case study of the implementation of a multimodal transit simulation. Figure 6.2 shows the road map of the NMBMM as obtained from Open Street Map Development Team (2011).

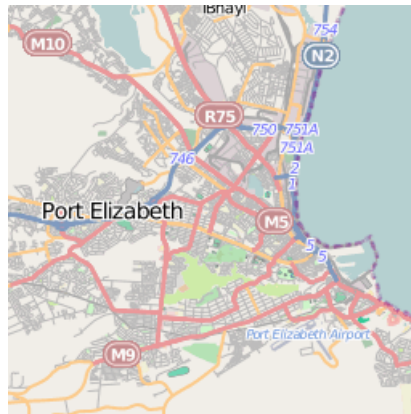


Figure 6.2: Road map of the NMBMM

In the NMBMM three modes of public transport should be included in a complete simulation of the transit system. These three networks are the rail network, the bus network and the minibus taxi network. Simulating the rail and bus network is a relatively simple matter once the required data is gathered, as the MATSim toolkit is set up to simulate these forms of transit. The minibus taxi network provides additional challenges in the modelling process.

In the NMBMM two official transit operators are present as well the semi-informal paratransit system. The formal transit providers are the Eastern Cape Metrorail, that operates a railway from Uitenhage to Port Elizabeth, and the Algoa Bus Company that operates 298 routes in the greater NMBMM area. The Eastern Cape Metrorail line in the NMBMM uses the rail infrastructure from a former freight line. The rail service therefore is inaccessible to the majority of transit users.

According to theCITP the full paratransit infrastructure consists of 49

taxi ranks and an estimated 258 basic routes. To further complicate matters the CITP estimates that only 44% of taxis operating in the NMBMM are licensed. An estimated total of 2347 mini-bus taxis are operating in the NMBMM (SSI Engineers and Environmental Consultants, 2011).

6.2.1 Working with the data

The transit schedules for the formal transit providers in the NMBMM are obtained in Excel sheets for the Metrorail service and extracted as raw data for the Algoa Bus Company. The data can then be converted to the GTFS format using TDF in the case of the Excel spreadsheets and by cleaning the raw data received from the Algoa Bus Company.

In order to generate MATSim specific files from the data the GTFS files will be combined with a multimodal network in a MATSim specific XML-format and imported into the java program `GTFS2MATSimTransitScheduleFileWriter` as developed by Sergio Arturo Ordóñez Medina. The program consists of three main phases:

- **The import phase** The GTFS files for all modes of transit are imported and combined with the network.
- **The data cleaning phase** The data is cleaned using a semi-automatic procedure that employs a user interface.
- **The data conversion phase** The required MATSim specific XML files are created.

The full process of running `GTFS2MATSimTransitScheduleFileWriter` using the Eclipse platform is shown in Appendix E.

A summary of the complete data conversion process is illustrated in Figure 6.3 and the transit schedule process is illustrated in greater detail in Appendix F:

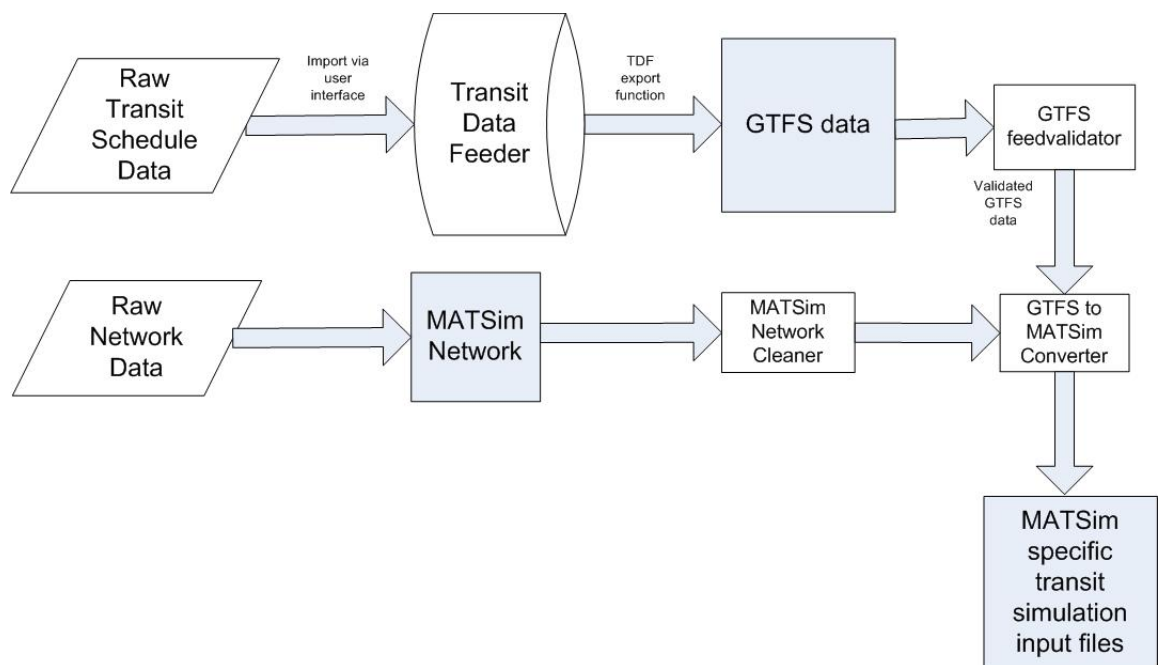


Figure 6.3: Transfer data conversion process

Chapter 7

Conclusion

Implementing mass transport simulations in South Africa provides the opportunity to improve the cooperation of transportation companies and other organizations involved in transportation. Currently integrated public transport systems in South African cities are virtually non-existent (SSI Engineers and Environmental Consultants, 2011). Transit operators in these cities compete with each other for passengers instead of combining their efforts to improve the service provided to the passenger thereby enlarging the potential passenger market. In theory by combining all modes of transit into a single network, all transit operators involved will be rewarded with more passengers and passengers will be rewarded with better, safer, faster public transport (Department of development planning and urban management, 2008). This can be achieved by obtaining a better understanding of the transport challenges and needs of passengers and by enabling effective transportation management with the use of transportation modelling.

ABS like the MATSim toolkit provides the opportunity to prove to transit operators the value of an integrated public transport plan. MATSim also enables the effective design of a integrated transit plan in the best interest of all parties effected. This report shows that integrating a full-scale MATSim transit simulation is possible for South African conditions. In order to achieve this goal cooperation is required from the relevant transportation and government agencies to provide the data and funding required to perform the necessary simulation and analysis.

If we as South Africans want to live in a modern society it is up to us to create it. “You must be the change you wish to see in the world” Mahatma Gandhi.

7.1 Summary

Although the project was unable to complete a full-scale transit simulation in the time-frame provided, important steps were taken towards achieving that goal. Chapter 3 indicated the data required to perform a full-scale simulation and indicates the data-converting processes that currently are available.

In Chapter 4 possible ways to adapt a simulation to create a more realistic representation of the target transportation network were determined by extending and adapting the standard MATSim example transit simulation.

In Chapter 5 the effects of changing certain parameters of the MATSim example transit simulation were measured and some relevant parameters for simulation under South African conditions were determined. A feasible simulation of bus and basic paratransit networks was also achieved in the example network.

Chapter 6 developed a possible fare-comparison structure that can be employed to improve simulations of the competition for passengers between bus and taxi networks in South African cities. The progress made towards simulating the NMBMM as a case study for MATSim simulation in South Africa is also indicated.

Bibliography

- Balmer, M., 2007. Travel demand modeling for multi-agent transport simulations: Algorithms and systems. Ph.D. thesis, Swiss Federal Institute of Technology (ETH) Zurich, Switzerland.
- Balmer, M., Meister, K., Rieser, M., Nagel, K., Axhausen, K., 2008. Agent-Based Simulation of travel demand: Structure and computational performance of MATSim-T. In: In 2nd TRB Conference on Innovations in Travel Modeling.
- Department of development planning and urban management, 2008. City of johannesburg integrated transport plan. Tech. rep., Development planning and urban management for the City of Johannesburg.
- Diedericks, D., Joubert, J., 2006. Towards transportation system integration in the City of Tshwane Metropolitan Municipality. Urban Transport and the Environment in the 21st Century Urban Transport XII.
- Flettermann, M., 2008. Designing multimodal public transport networks using metaheuristics. Master's thesis, University of Pretoria.
- Fourie, P. J., 2009. An initial implementation of a multi-agent transport simulation for South Africa. Master's thesis, University of Pretoria.
- Fourie, P. J., 2010. Agent-based transport simulation versus equilibrium assignment for private vehicle traffic in Gauteng. South African Transport Conference ISBN 978-1-920017-47-7.
- MATSim Development Team, 2011. Multi-Agent Transport Simulation, MATSim.
URL www.matsim.org
- McNally, M., 2000. Handbook of Transport Modelling. Vol. 1. Pergamon, Ch. 3, pp. 35–69.
- Meyer, M., Miller, E., 2001. Urban Transportation Planning - A Decision-Oriented Approach. MCGraw-Hill.

Open Street Map Development Team, 2011. Open Street Map, The Free Wiki World Map.

URL www.openstreetmap.org

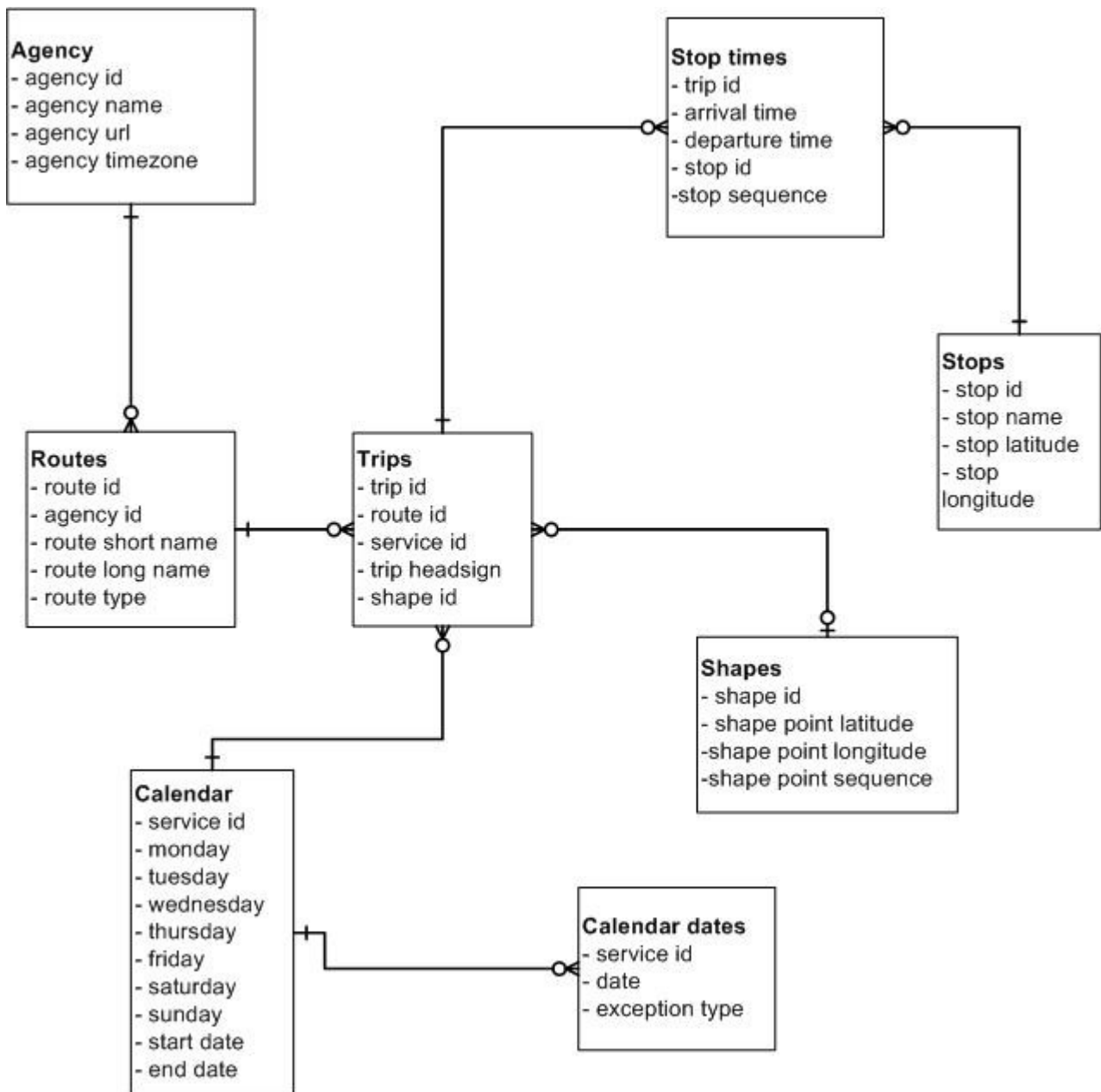
Rieser, M., 2010. Adding Transit to an Agent-Based Transportation Simulation. Ph.D. thesis, Technische Universitat Berlin.

SSI Engineers and Environmental Consultants, 2011. Comprehensive Integrated Transport Plan. Tech. rep., Infrastructure and Engineering Directorate Nelson Mandela Bay Municipality.

van der Merwe, J., 2011. Agent-Based Transport Demand Modelling for the South African Commuter Environment. Master's thesis, University of Pretoria.

Appendix A

Recommended ERD for storing South African specific GTFS data



Appendix B

Example of a MATSim config file

```

<?xml version="1.0" ?>
<!DOCTYPE config SYSTEM "http://www.matsim.org/files/dtd/config_v1.dtd">
<config>

  <module name="global">
    <param name="randomSeed" value="4711" />
    <param name="coordinateSystem" value="Atlantis" />
  </module>

  <module name="network">
    <param name="inputNetworkFile" value="examples/pt-
tutorial/multimodalnetwork.xml" />
  </module>

  <module name="plans">
    <param name="inputPlansFile" value="examples/pt-
tutorial/population.xml" />
  </module>

  <module name="scenario">
    <param name="useTransit" value="true" />
    <param name="useVehicles" value="true" />
  </module>

  <module name="controler">
    <param name="outputDirectory" value="./output/pt-tutorial" />
    <param name="firstIteration" value="0" />
    <param name="lastIteration" value="50" />
    <param name="eventsFileFormat" value="xml" />
    <param name="mobsim" value="qsim" />
  </module>

  <module name="qsim">
    <!-- "start/endTime" of MobSim (00:00:00 == take earliest activity
time/ run as long as active vehicles exist) -->
    <param name="startTime" value="00:00:00" />
    <param name="endTime" value="30:00:00" />

    <param name = "snapshotperiod"      value = "00:00:00"/> <!--
00:00:00 means NO snapshot writing -->
    <param name = "snapshotFormat"     value = "otfvis"/> <!--
googleearth, transims, otfvis -->
  </module>

  <module name="planCalcScore">
    <param name="learningRate" value="1.0" />
    <param name="BrainExpBeta" value="2.0" />

    <param name="lateArrival" value="-18" />
    <param name="earlyDeparture" value="-0" />
    <param name="performing" value="+6" />
    <param name="traveling" value="-6" />
    <param name="waiting" value="-0" />

    <param name="activityType_0"      value="h" /> <!-- home -->
    <param name="activityPriority_0"  value="1" />
    <param name="activityTypicalDuration_0" value="12:00:00" />
    <param name="activityMinimalDuration_0" value="08:00:00" />

    <param name="activityType_1"      value="w" /> <!-- work -->
    <param name="activityPriority_1"  value="1" />
    <param name="activityTypicalDuration_1" value="08:00:00" />

```

Appendix C

Fare comparisons between bus and taxi networks

BUS AND TAXI COMPARISON

PASSES	ORIGIN	DESTINATION	CASH FARE = R			TAXI FARE =R		VARIANCE TAXI/BUS
			CURRENT	NEW		CURRENT	NEW	
2XB	Uitenhage	Kirkwood	R 15.60		R 15.60		R 18.00	-R 2.40
A	Bethelsdorp	Markman	R 10.50		R 10.00			
A	Bethelsdorp	Uitenhage	R 10.50		R 10.00			
A	Despatch	Port Elizabeth	R 10.50		R 10.00		R 10.00	R 0.00
A	Motherwell	Kwanobuhle	R 10.50		R 10.00		?	
A	Motherwell	Lorraine	R 10.50		R 10.00		R 9.00	R 1.00
A	Motherwell	Summerstrand	R 10.50	-R 0.50	R 10.00		R 9.00	R 1.00
A	Uitenhage	Port Elizabeth	R 10.50		R 10.00		R 12.00	-R 2.00
B	Bethelsdorp	Perseverance	R 7.90		R 8.50		R 10.00	-R 1.50
B	Cleary Park	Markman	R 7.90	R 0.60	R 8.50			
B	Gelvandale	City	R 7.90		R 8.50		R 7.00	R 1.50
B	Joe Slovo	Lorraine	R 7.90		R 8.50		R 7.30	R 1.20
B	Kwadwesi	Uitenhage	R 7.90		R 8.50			
B	Kwadwesi	Markman/Motherwell	R 7.90		R 8.50			
B	Malabar	City	R 7.90		R 8.50		R 10.00	-R 1.50
B	Motherwell	City	R 7.90		R 8.50		R 7.50	R 1.00
B	Motherwell	Greenacres	R 7.90		R 8.50		R 8.00	R 0.50
B	Motherwell	Korstein	R 7.90		R 8.50		R 7.50	R 1.00
B	Motherwell	Uitenhage	R 7.90		R 8.50		R 11.00	-R 2.50
B	New Brighton	Summerstrand	R 7.90		R 8.50		R 7.00	R 1.50
B	Uitenhage	Motherwell	R 7.90		R 8.50		R 11.00	-R 2.50
B	Veeplaas	Summerstrand	R 7.90		R 8.50		R 7.30	R 1.20
B	City	Lorraine/Mnt Pleasant/Sherwood/Rowallan Pk/Booys	R 7.90		R 8.50			
B	Zwide	Uitenhage	R 7.90		R 8.50		R 10.30	-R 1.80
C	Booyesen Park	Cleary Park	R 6.10		R 6.50		R 6.00	R 0.50
C	Cleary Park	Deal Party	R 6.10		R 6.50			
C	Cleary Park	Struandale	R 6.10	R 0.40	R 6.50		R 7.00	-R 0.50
C	Cleary Park	Northern Areas/Korsten	R 6.10		R 6.50		R 5.00	R 1.50
C	Cleary Park	Gelvandale	R 6.10		R 6.50		R 7.00	-R 0.50
C	Cleary Park	Korsten	R 6.10		R 6.50		R 6.00	R 0.50
C	Govan Mbeki	City	R 6.10		R 6.50			
C	Korsten	City	R 6.10		R 6.50		R 7.00	-R 0.50
C	Korsten	Greenacres	R 6.10		R 6.50		R 6.00	R 0.50
C	Korsten	Western Areas	R 6.10		R 6.50		R 7.00	-R 0.50
C	Kwanobuhle	City	R 6.10		R 6.50		R 6.00	R 0.50
C	Kwanobuhle	Roesdale	R 6.10		R 6.50		R 6.00	R 0.50
C	Kwanobuhle	City + Langa	R 6.10		R 6.50		R 5.50	R 1.00
C	Kwanobuhle	City - Hospital/Fairbridge Heights	R 6.10		R 6.50		R 7.00	-R 0.50
C	Malabar	Korsten	R 6.10		R 6.50		R 5.00	R 1.50
C	Motherwell	Deal Party	R 6.10		R 6.50		R 7.00	-R 0.50
C	Motherwell	Township	R 6.10		R 6.50		R 5.00	R 1.50
C	New Brighton	City	R 6.10		R 6.50		R 6.50	R 0.00
C	New Brighton	Korsten	R 6.10		R 6.50		R 6.50	R 0.00
C	New Brighton	Markman/Motherwell	R 6.10		R 6.50		R 7.00	-R 0.50
C	New Brighton	Greenacres	R 6.10		R 6.50		R 7.00	-R 0.50
C	New Brighton	Township	R 6.10		R 6.50			
C	Rowallan Park	Korsten	R 6.10		R 6.50			
C	City	Walmer/Summerstrand/Khans/Algoa Pk	R 6.10		R 6.50			

Appendix D

Specifications for creating GTFS files

GTFS specifications adapted from
http://code.google.com/transit/spec/transit_feed_specification.html

The following text(*.txt) files are required:

- agency.txt
- calendar.txt
- calendar_dates.txt
- routes.txt
- stops.txt
- trips.txt
- stop_times.txt
- shapes.txt

1) Data required for agency:

agency_name,agency_url,agency_timezone,agency_lang,agency_phone
example: Metrorail,http://www.metrorail.co.za,Africa/Johannesburg,en,043-700-2030

2) Data required for calendar:

service_id,monday,tuesday,wednesday,thursday,friday,saturday,sunday,start_date,end_date
example: sW,1,1,1,1,1,0,0,20110404,20111231

If the specific service is active on a particular day of the week it is indicated using a "1", "0" indicates the particular service is not active on that day of the week. The start and end date indicates the period for which the schedule is valid. The calendar file therefore indicates the breakdown on the different service schedules employed for certain days, for example a schedule can be created for standard weekdays, public holidays and weekends.

3) Data required for calendar_dates:

service_id,date,exception_type
example: sW,20110422,2

Calendar dates list the days that has exceptions to the normal schedule, this mostly consists of public holidays. Exception type "1" adds a extra service to the schedule and exception type "2" removes a service from the schedule.

4) Data required for routes:

route_id,route_short_name,route_long_name,route_type
example: r1,Main,PE Metrorail,0

The different routes that are serviced are listed to be referenced in trips. The following routes types are possible:

- 0 - Tram, Streetcar, Light rail. Any light rail or street level system within a metropolitan area.
- 1 - Subway, Metro. Any underground rail system within a metropolitan area.
- 2 - Rail. Used for intercity or long-distance travel.
- 3 - Bus. Used for short- and long-distance bus routes.
- 4 - Ferry. Used for short- and long-distance boat service.
- 5 - Cable car. Used for street-level cable cars where the cable runs beneath the car.
- 6 - Gondola, Suspended cable car. Typically used for aerial cable cars where the car is suspended from the cable.
- 7 - Funicular. Any rail system designed for steep inclines.

5) Data required for stops:

stop_id,stop_name,stop_lat,stop_lon,location_type

example: 1,Port Elizabeth,-33.960280,25.624170,0

List all the stops employed in the schedule as well as the other locations that effects the movement of the transit vehicles. Location type indicates if a location is a stop on a route "0" or a bus depot or other important facility "1".

6) Data required for trips:

route_id,service_id,trip_id,direction_id,trip_headsign

example: r1,sW,0002,0,Uitenhage

Each trip in the trip file shows a specific instance of travel along a certain route. The direction is set as outbounds "0" or inbounds "1" from a reference point. The trip headsign serves as a description of the direction of the route.

7) Data required for stop_times:

trip_id,arrival_time,departure_time,stop_id,stop_sequence

example: 0002,00:00:00,05:43:00,1,1

0002,05:46:00,05:47:00,2,2

0002,05:50:00,05:51:00,3,3

0002,05:54:00,05:55:00,4,4

The stop times list the expected arrival time and departure time of each vehicle on a specific trip at each stop on that trip.

8) Data required for shape file:

shape_id,shape_pt_lat,shape_pt_lon,shape_pt_sequence

PE_out,-33.9579350,25.62210110,1

PE_out,-33.9564397,25.62011160,2

The shape file is a point to point GPS map of the route travelled for calculated per bus trip, and not as per bus route as expected.

Appendix E

Instructions for converting GTFS data to MATSim

How to convert GTFS data to a MATSim Transit Schedule by Sergio Arturo Ordonez Medina.

1. Put the set of GTFS files of each public transport system in a different folder of your file system.
2. Create a java program that constructs an object of the **GTFS2MATSimTransitSchedule** class located in */gtfs2matsimtransitschedule/src/main/java/GTFS2PTSchedule* package of the **gtfs2matsimtransitschedule** contrib project of MATSim. For this you need to specify:
 - a) An array of folders (**File** class) where your public transport system specifications are located.
 - b) An array of Strings representing the network modes correspondent to each public transport defined in b) (e. g. "car" for buses, "rail" for metro).
 - c) The MATSim network object with the nodes in latitude and longitude coordinates (WGS84).
 - d) An array of Strings with the names of the calendar services that are desired (e. g. "weekday", "daily"). Remember that MATSim only simulates one day, but the GTFS files specify routes for many calendar days or dates.
 - e) The desired output coordinates system
3. Call the method *TransitSchedule* *getTransitSchedule()*. Then, each route of each given public transit systems is going to be processed with the semi-automatic procedure presented in the figure:

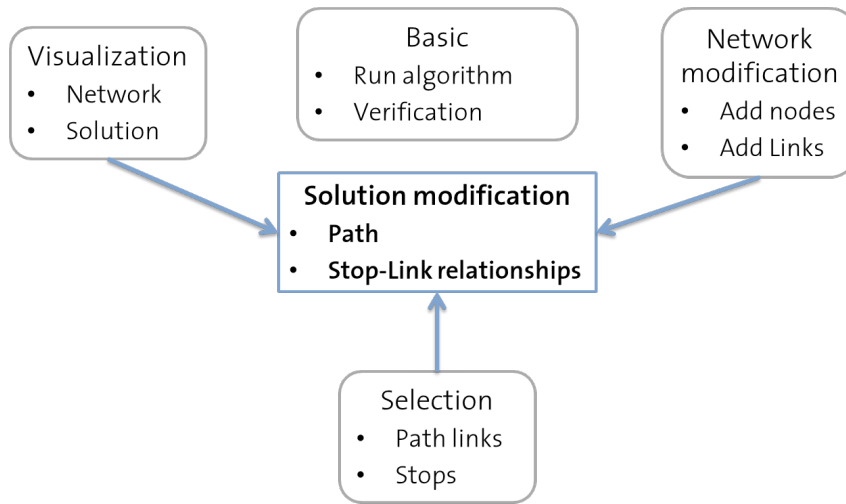
For the manual editing process one can visualize, edit and verify the solution for each route:

- a) Visualization: A navigation network is displayed, including all relevant information for working with one single route. This includes the route's profile, the given sequence of GPS points and its current solution (path and stop-link relationships). Selected elements are drawn in a different color. All is displayed in a bi-dimensional interactive way with refresh of the cursor location in the working coordinates, and panning, zoom and view-all options.
- b) Selection: Different options for selecting elements of the solution or elements from the network are provided. It is possible to select the nearest link (solution or network), nearest node (network) or nearest stop (solution) to a point indicated by the user. When a stop that has a stop-link relationship already, the corresponding link is selected as well. If a link of the solution path is selected and it does not have a subsequent link connected, a new one from the network is selected with one click; the selected link is the one with the most similar angle than the line defined by the end node of the initial link and a point indicated by the user.
- c) Solution modification: The first link of the sequence can be added selecting any link of the network. If a link of the solution path does not have a subsequent link connected, it is possible to add one according to the selection function described in (b). If there are two subsequent links in the solution that are not connected (a gap), a subsequence that connects the mentioned links is added, using the shortest path algorithm, with the current parameters. Furthermore, selecting one link of the path, it is possible to delete it, or to delete all the links before or after it. Finally, stop-link relationships can be modified selecting both elements. If the modified relationship was fixed, the user is prevented because the tool erase the solutions of the routes to which the selected stop belongs.
- d) Network modification: New nodes to the road network can be added. In addition, with any node selected, it is possible to add a new link selecting the end node.

Hints and interaction details:

- It is necessary to pass the verification process ("Is OK") for saving a route result
- The routes are saved in temporal files located in the ./data/paths/ folder relative to the program location.
- **Panning** and **zoom** functions are provided dragging the mouse and moving the mouse wheel.
- **View all** function is provided typing the "v" key
- Up and down keys allow to select the next or previous link of the path.

- “<” or “>” keys allow to select the previous or next stop



4. Finally, after the semi-automatic process, the Transit Schedule object is returned and the network object is modified (splitting, new nodes and links, and modes of the links). One can save in a XML the **TransitSchedule** object constructing a **TransitScheduleWriter** object, and the modified network with a **NetworkWriter**.

Appendix F

Transit schedule data conversion process

Raw Transit Schedules

Transit Data Feeder

Import via user interface

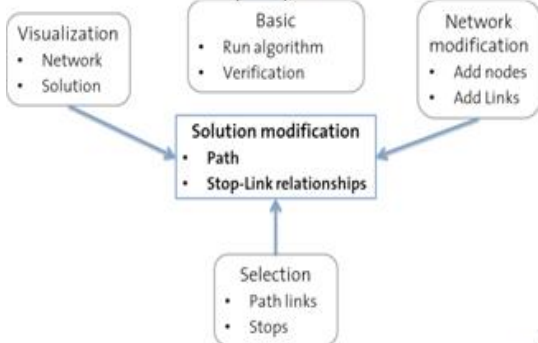
GTFS Data

TDF export function

Name	Size	Type	Date Modified
agency.txt	139 bytes	plain text d...	02 April 2011, 07:55
calendar.txt	161 bytes	plain text d...	18 April 2011, 12:00
calendar_dates.txt	167 bytes	plain text d...	18 April 2011, 12:00
routes.txt	121 bytes	plain text d...	19 May 2011, 15:00
shapes.txt	14.2 KB	plain text d...	20 May 2011, 11:13
stop_times.txt	4.1 KB	plain text d...	19 May 2011, 14:59
stops.txt	418 bytes	plain text d...	15 April 2011, 21:58
trips.txt	551 bytes	plain text d...	20 May 2011, 11:06

8 objects (19.9 KB)

Combine with MATSim network



GTFS to MATSim Conversion

MATSim input files

```
<nodes>
<node id="1" x="1050" y="1050" />
<node id="2" x="2050" y="2050" />
<node id="3" x="3050" y="1050" />
</nodes>
<links>
<link id="11" from="1" to="2" length="100.00" capacity="2000" freespeed="100" nodes="train" perlanes="1" />
<link id="12" from="1" to="3" length="100.00" capacity="2000" freespeed="100" nodes="train" perlanes="1" />
<link id="21" from="2" to="3" length="100.00" capacity="2000" freespeed="100" nodes="train" perlanes="1" />
<link id="1111" from="11" to="12" length="1200.00" capacity="2000" freespeed="12" nodes="car,bus,taxi" perlanes="1" />
<link id="1112" from="11" to="13" length="1200.00" capacity="2000" freespeed="12" nodes="car" perlanes="1" />
</links>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE SYSTEM "http://www.matsim.org/files/dbd/transitschedule_v1.dtd">
<itSchedule>
<transitStops>
<stopFacility id="1" x="1050" y="1050" linkRefId="11"/>
<stopFacility id="2a" x="2050" y="2050" linkRefId="12"/>
<stopFacility id="3" x="3050" y="1050" linkRefId="33"/>
</transitStops>
<transitLine id="Blue Line">
<transitRoute id="It03">
<transportMode>train</transportMode>
<routeProfile>
<stop refId="1" departureOffset="00:00:00"/>
<stop refId="2a" arrivalOffset="00:03:20" departureOffset="00:04:00"/>
<stop refId="3" arrivalOffset="00:09:00" />
</routeProfile>
<route>
<link refId="11"/>
<link refId="12"/>
<link refId="23"/>
<link refId="33"/>
</route>
<departures>
<departure id="01" departureTime="00:00:00" vehicleRefId="tr 1" />
<departure id="02" departureTime="06:15:00" vehicleRefId="tr 2" />
<departure id="03" departureTime="06:30:00" vehicleRefId="tr 1" />
<departure id="04" departureTime="06:40:00" vehicleRefId="tr 2" />
<departure id="05" departureTime="06:50:00" vehicleRefId="tr 1" />
<departure id="06" departureTime="07:00:00" vehicleRefId="tr 2" />
<departure id="07" departureTime="07:10:00" vehicleRefId="tr 1" />
<departure id="08" departureTime="07:20:00" vehicleRefId="tr 2" />
<departure id="09" departureTime="07:30:00" vehicleRefId="tr 1" />
</departures>
</itSchedule>
```