THE FIRST AGENT STEPS IN AGENT-BASED TRANSPORT PLANNING

P J Fourie and J W Joubert

Optimisation Group, Department of Industrial and Systems Engineering, University of Pretoria

ABSTRACT

South Africa's socio-political history and rapidly changing urban dynamics, coupled with an eclectic mix of first and third world transport infrastructure presents a moving target to even the best traditional transport planning methodologies. A pressing need has emerged for a planning technology that can capture these dynamics and explore its evolution over time. In this regard, agent-based microsimulation has become a viable tool for analysing and predicting transport system performance, and promises to address the current and future needs of transport planners in a variety of scenarios. Our initial implementation of Multi-Agent Transport Simulation Toolkit (MATSim), a large-scale agent-based microsimulation framework proposed by Balmer et. al. (2006), models the passenger vehicle traffic of Gauteng for 2001. This paper describes the basic principles of operation of MATSim, and identifies data sources relevant to current and future states of the South African implementation.

INTRODUCTION

South Africa's socio-political history and rapidly changing urban dynamics, coupled with an eclectic mix of first and third world transport infrastructure presents a moving target to even the best traditional transport planning methodologies. A pressing need has emerged for a planning technology that can grow and adapt to change, and be capable of capturing emergent phenomena resulting from the complex interaction of the transport system's constituent entities.

Our research group has identified the demand modelling framework proposed by Balmer et. al. (2006) as a viable technology, potentially capable of modelling the increasingly complex dynamics of the South African transport system. Called the Multi-Agent Transport Simulation Toolkit (MATSim) (www.matsim.org), it has been classified as a *mesoscopic* transport demand model, capable of capturing the temporal dynamics and interaction within a transport system and the decision-making of its constituency, while discarding the detail behaviour of vehicles in order to rapidly simulate large-scale scenarios.

As the name suggests, MATSim is agent-based, therefore its picture of overall transport system performance emerges from the behaviour and interaction of its constituent entities. Generally, agent-based traffic simulations are limited to very small scenarios, due to the large computational overhead of simulating the detail behaviour of vehicles. In the case of MATSim, however, the actual mobility simulation is approximated as a stochastic queue-based simulation. As a result, it is possible to simulate very large scenarios very quickly, making MATSim viable as a transport demand planning tool.

In MATSim, agent participation in the transport system is driven by the need to realise a day plan of activities. There exists a reluctance to abandon traditional transport planning methods in favour of such an Activity-Based Approach (ABA), due to the perception that the ABA relies on far more complicated techniques and requires commensurately complex and unusual data for input and calibration.

In South Africa, some work has been done on the activity-based analysis of transport demand, but, to our knowledge, no study has progressed beyond pilot-stage application. In his review, Behrens (2000) surveys local developments in the field, and cites the perceived cost of attaining richer data through highly sophisticated surveys as the main hurdle to the application of activity-based analysis in the South African context.

Proceedings of the 28th Southern African Transport Conference (SATC 2009) 6 – 9 July 2009
ISBN Number: 978-1-920017-39-2 Pretoria, South Africa
Produced by: Document Transformation Technologies cc Conference organised by: Conference Planners

However, it has been shown how MATSim can eliminate this data hurdle by using traditional data sources to deliver a set of results far richer in detail than what is attainable through traditional means (Balmer, 2007; Balmer and Rieser, 2004). Consequently, we were convinced that it should be possible to get an initial South African implementation operational using available census and travel survey data. This initial implementation would then form the basis for a process of continuous improvement towards the ultimate goal of a fully integrated multi-modal simulator, capable of modelling private, public and commercial traffic.

It is our opinion that, if MATSim can deliver sensible results from traditional transport demand data, we can build a stronger case for the acquisition of richer travel survey data, as the relative gains from such an exercise will be more apparent to the relevant stakeholders. The aim of our initial implementation, therefore, is to model a South African transport demand scenario in the completely disaggregate, agent-based framework of MATSim, using only readily available, aggregate, traditional transport demand data.

In this paper, we provide a short overview of MATSim's principles of operation, and determine the minimum data requirements necessary to get a MATSim implementation off the ground. We identify the relevant data sources used in our initial implementation, where we attempt to model the private vehicle traffic in Gauteng during the morning and afternoon rush hour peaks of 2001. We also discuss the potential and shortcomings of available data in the further development and extension of our initial implementation.

MATSIM OPERATION

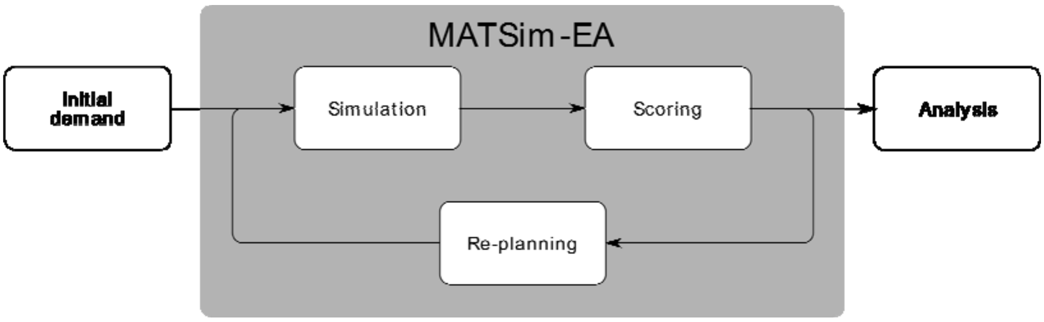

Figure 1 Schematic illustration of MATSim's transport demand modeling process. The shaded area indicates the evolutionary "engine" of MATSim, which simulates system learning and adaptation. Adapted from Rieser (2007).

Figure 1 shows a simplified view of MATSim's operation. In principle, MATSim executes, evaluates and improves agent day plans. Each agent day plan is a simple schedule of activities, their locations in the study area and the preferred mode(s) of transport to link those activities. Initial plans are fed into the mobility simulation, executed simultaneously and then evaluated to arrive at a score for each plan. The re-planning step provides the feedback that allows for the evolution of commuter plans. "Evolution" is an apt term in this context: MATSim's iterative process of mobility simulation, scoring and re-planning is based on the concept of a so-called Evolutionary Algorithm (EA). EAs solve optimisation problems by generating and improving "populations" of candidate solutions using mechanisms inspired by biological evolution such as mutation, recombination and selection (Engelbrecht, 2002).

In MATSim's context, "population" refers not to the commuter population; the simulation process ends with exactly the same "people" it started out with. Instead, it refers to the collective "memory" of the commuters; a collection of plans associated with each commuter that changes with increasing iterations. From an EA perspective, it is this "population" of plans that evolves over time.

Initialisation and execution

Each individual starts out with only one plan, specified in the process of initial demand generation. During the first iteration, each plan is routed through the network using a Dijkstra algorithm that attempts to minimise travel time (Balmer, 2007). All commuter plans are then executed simultaneously in the mobility simulator. During the simulation, events are registered for each individual, such as travel time, activity departure and arrival times.

Scoring

In the next step, a score value is calculated for each commuter plan, based on how well that plan performed in the mobility simulation. From an EA perspective, the score is a measure of fitness for each plan. MATSim uses a simple utility-based approach to calculate a plan score (Balmer, 2007) :

$$U_{total} = \sum_{i=1}^{n} U_{perf,i} + \sum_{i=1}^{n} U_{late,i} + \sum_{i=1}^{n} U_{travel,i} \quad (1)$$

where U_{total} is the total utility for the executed plan, n is the number of activities, $U_{perf,i}$ is the positive utility for performing activity i, $U_{late,i}$ is the negative utility for arriving late at activity i and $U_{travel,i}$ is the negative utility for travelling to activity i. Each one of these utility values are time-dependent functions of arbitrary complexity (the interested reader is referred to Balmer (2007) for a detailed description of these utility functions). Regardless of the constituent function complexity, it is evident from (1) that greatest utility and, consequently, plan fitness, derives from more time spent performing activities, while avoiding travel and arriving late at activity locations.

Replanning

Once executed plans have been scored, the next step in MATSim's iterative demand simulation process executes. The re-planning step refers to a sequence of configurable algorithms, and is analogous to the mutation and selection processes found in EAs. In his thesis, Balmer (2007) lists a number of algorithms used to select from and expand the "population" of plans.

These algorithms ensure that the "population" of plans gradually adapts to the transport environment. From an EA perspective, the re-planning step is analogous to the mechanisms of mutation and selection — activity times get adjusted, routes are optimised based on recent system performance and poorly performing plans get discarded. This process of adaptation is reflected in the improvement of average plan scores and travel times with increasing iterations, until a form of the Nash-equilibrium is reached, whereby changes in commuter plans do not produce any further improvement in total utility.

The interested reader is referred to the MATSim project homepage at http://www.matsim.org for a detailed description of MATSim's operation.

MATSIM DATA REQUIREMENTS

In this section, we examine the minimum data requirements necessary to get a MATSim implementation off the ground. In each case, we identify the relevant data source used in our initial implementation, where we attempted to model the private vehicle traffic of Gauteng, for our base year of 2001, when the last national census was conducted. In each case we also discuss the potential and shortcomings of existing data to extend our initial implementation as we proceed towards our goal of a fully integrated, multi-modal transport
demand model.

Initial demand generation

The process of initial demand generation consists of two steps; population synthesis and day plan generation. Population synthesis requires geospatially coded census data describing the study population, which can take the form of actual person records, or be aggregated for each geographic sample unit. In general, census data is released in aggregated form to protect confidentiality. In such a case, the synthetic population needs to be constructed from aggregate data in such a way that a census of the synthetic population will return the original census. Beckman et al. (1996) describe the statistical techniques necessary to perform such a disaggregation for the purpose of activity-based microsimulation. In MATSim, the description of a synthetic individual can include pertinent attributes such as sex, age, car ownership and income, that can influence the decision-making and behaviour of that agent.

Day plan generation entails the assignment and scheduling of activities. Travel survey data are either used as input data or to calibrate and validate an activity assignment procedure. Traditionally, travel survey data are processed to take the form of an Origin-Destination (O-D) matrix, to be used in transport demand models based on the so-called Four Step Model (FSM). In MATSim, such an O-D matrix can be used to probabilistically assign activity locations to synthetic individuals based on their home location and other demographic information (Balmer and Rieser, 2004). This process of extrapolating from traditional transport data is convenient when detailed activity-based travel survey data are unavailable, and is the process used for this study.

Population synthesis

For our initial implementation we focussed on the morning and afternoon rush-hour traffic due to private vehicle drivers going to- and returning from work. We therefore needed to isolate this section of the population from available census data, and interpret these data in order to synthesise an agent population, and assign them likely home locations within the study area.

The geospatial resolution of census data has an important influence in the quality of any transport demand model. In the case of South African census data, the smallest discernible geographic unit is a so-called Enumeration Area Primary Sampling Unit (EAPSU). In preparation for the 2001 census, Statistics South Africa (Stats SA) produced a mapping of South Africa whereby the entire country was demarcated into a total of 80,787 non-overlapping EAPSUs. Each EAPSU is a small land unit of manageable population and area allocated to a single person to enumerate during the census count (Stats SA, 2001).

We first examined the 10% sample of person records from the 2001 Census as a possible source to generate our synthetic population, as the census questionnaire recorded the travel mode taken to work, making it possible to isolate private vehicle drivers. Unfortunately, for reasons of confidentiality, the smallest identifiable geographic unit in these person records is the municipality, and therefore is far too coarsely grained to serve as input to our model.

Stats SA provides another anonymous source of information, a collection of so-called Sub-Place (SP) tables, which ensures individual anonymity through aggregation, with a far higher degree of spatial resolution than the municipality level. An SP refers to a grouping of EAPSUs, and is the first level in a hierarchy of such groupings that was defined by Stats SA to allow for the systematic aggregation of census data. SP tables provide aggregate figures on parameters of interest from the census questionnaire, such as the number of males/females in each SP, or the SP's income distribution. For this study, we used the number of people who reported to drive to work in each SP to arrive at a study population size of 924,680 individuals. We assumed these individuals to be responsible for the bulk of private vehicle traffic in our study area, and proceeded to generate a 10% synthetic population for each SP as follows: if x respondents reported to drive to work within that SP, we generate $\frac{x}{10}$ synthetic individuals and equiprobably assign each a home EAPSU from the list that make up that SP. Then we assign each synthetic individual a random coordinate within the bounds of its home EAPSU to serve as its home location. The resulting home location coordinates of our synthetic population is shown in Figure 2.

SP tables can also be used to generate a synthetic population of public transport users, and it is our intention to extend our model to include public transport by the same method described above. Note that we assume a sharp distinction between public and private transport, at least for the census year. As the South African transport environment becomes increasingly multi-modal, especially in Gauteng, we will have to modify this assumption to include park-and-ride behaviour.

Work location assignment

For our initial implementation, we interpret data from the 2003 National Household Travel Survey (NHTS) to generate an origin-destination (O-D) matrix for the work trip. We then use this O-D matrix to assign work locations to our synthetic population.

The NHTS is a survey of the detailed travel behaviour of a sample of 45,000 households throughout South Africa. The survey was developed and conducted by Stats SA for the Department of Transport in 2003. In Gauteng, 7,839 individuals participated in the survey, of which 3,038 reported to drive to work.

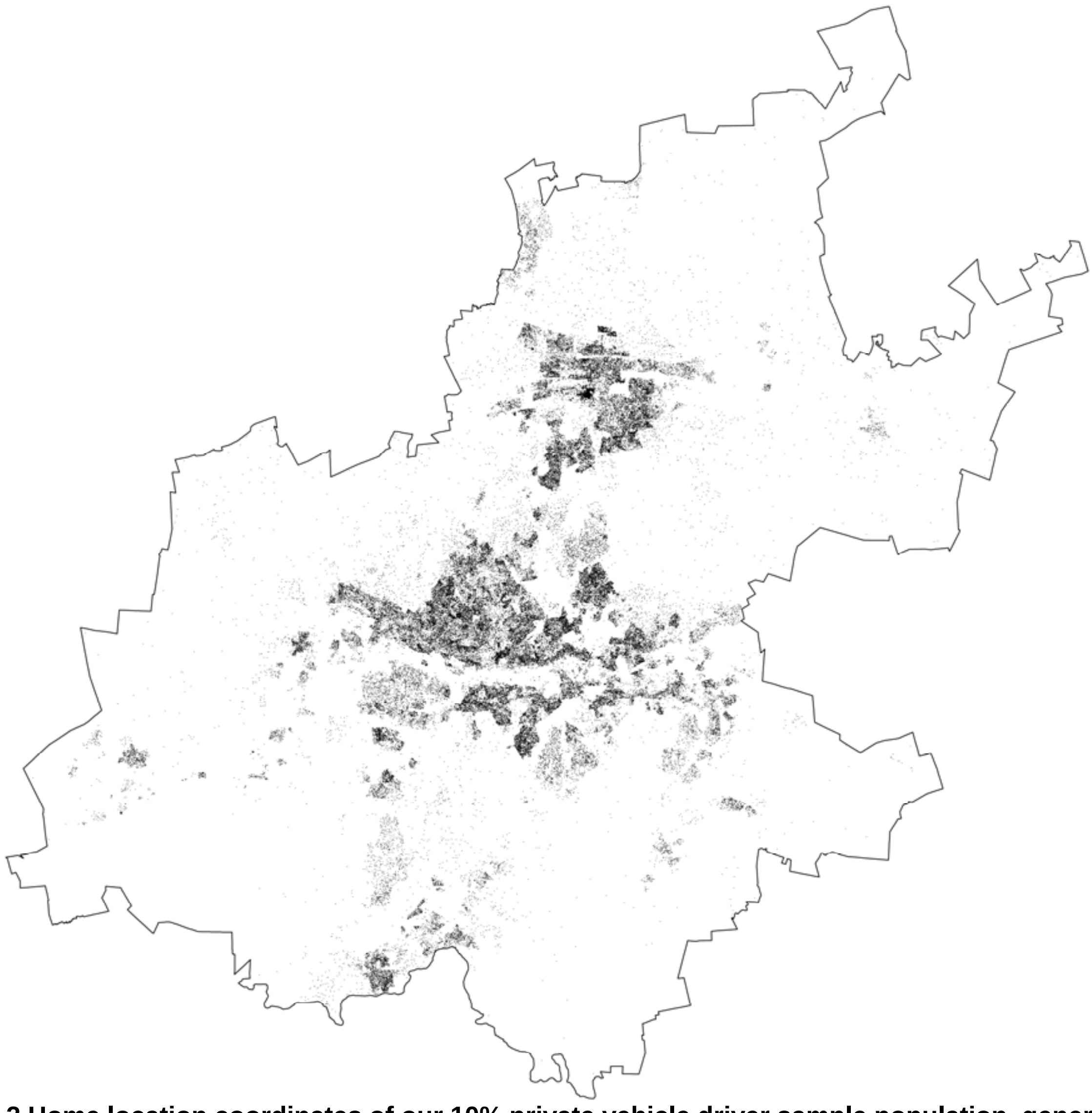

Figure 2 Home location coordinates of our 10% private vehicle driver sample population, generated from Census 2001 Sub-place tables.

In the NHTS, the smallest discernible geographic unit for work location is called a Traffic Analysis Zone (TAZ). Gauteng contains a total of 58 TAZs, each composed of a number of adjacent EAPSUs from the 2001 census.

Each person record in our NHTS data set collection cites home and work location as a reference to one of these 58 TAZs. It was therefore possible to construct a 58x58 O-D matrix associating the origins ("home TAZs") and destinations ("work TAZs") for these 3,038 individuals. By dividing each entry in the matrix by its row total, we get an estimated conditional probability of work TAZ given home TAZ.

We then used this normalised O-D matrix to assign a work location coordinate to each of the 92,468 private vehicle driver agents from the population synthesis step. For each agent, we locate its home TAZ from its home location EAPSU. Then we use a biased roulette-wheel technique to probabilistically assign a work TAZ, using the probability estimates from the corresponding home TAZ row in the normalised O-D matrix. Each individual is assigned a work EAPSU, chosen equiprobably from the work TAZ's constituency, and assigned a random work location coordinate within the boundaries of that EAPSU. The resulting work location coordinates of our synthetic population is shown in Figure 3.

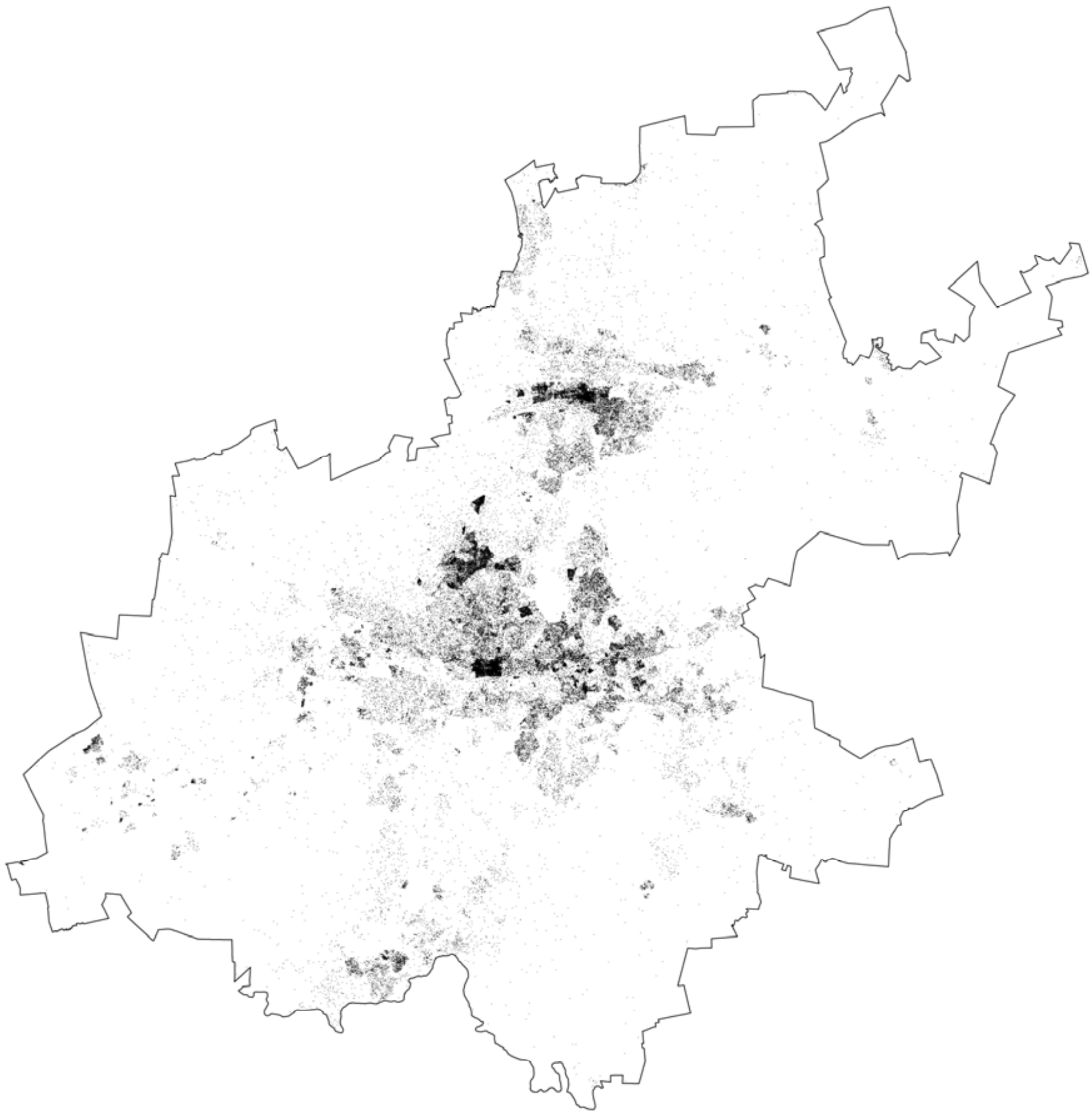

Figure 3 Work location coordinates of our 10% sample population of private vehicle drivers.

Activity schedule generation

Now that we have the location of the home and work activity for each of our virtual commuters, we need to order and schedule the activities taking place at these locations, as well as describe the transport mode and route to take through the network from one activity to the next. Only then will we have a compliant set of day plans for our sample population that can serve as input to the MATSim demand simulation process.

In our case, we come up with the order of activities and connecting transport mode (private vehicle) and MATSim takes care of the scheduling and time-dependent routing through repeated mobility simulation and re-planning (the EA solution procedure, illustrated in Figure 1). It only requires a sample schedule for each commuter as an initial condition to the demand simulation process. These schedules will be discarded during the demand simulation process as MATSim comes up with improved timings and durations for activities, and more efficient routes through the network.

In our case, the initial schedule was constructed as follows: each commuter departs for work from their home location at a uniformly distributed random time between 05h00 and 07h00. For our initial implementation, we assumed the preferred duration of the work activity to be 9 hours, as South African labour law prescribes a work day of eight hours plus a compulsory lunch hour, so once an agent arrives at work, it will stay there for nine hours or until 18h00 (the arbitrarily chosen closing time for the work activity), depending on which time comes first. Then, the agent returns to their home coordinate.

Beyond the initial implementation

The bulk of the remainder of NHTS records for Gauteng record the travel behaviour for individuals who have to rely on one or more public transport modes to get to their primary activity locations. The NHTS recorded the combination and order of transport modes used to get to work. Therefore, the NHTS can be used to assign primary activity locations to public transport users in a similar way as we did for private vehicle drivers, as well as to validate the combination of transport modes used to travel to and from the primary activity location.

Unfortunately, modelling public transport usage depends just as much on having data describing supply as demand. We therefore require routing and scheduling information for bus, rail and the minibus para-transit modes. Rail schedules are readily available from a single authority (Metrorail), but the road-bound public transport modes are run by a plethora of operators, and will require extensive efforts to document.

A further complication is that MATSim will need major modifications in order to model the highly dynamic and interactive minibus para-transit mode. Our research group is in the process of determining the requirements of modelling para-transit with colleagues in Zürich and Berlin, as similar systems are used in various parts of the world.

In order to extend our initial implementation beyond the primary activity chain, both for private and public transport, we require detailed land-use data for the study area. These appear to be readily available from commercial sources such as www.geoterraimage.com. Coupled with building height data, a detailed description of secondary activity opportunities can be derived, in terms of the number and type of opportunities that exist at each building location.

As for commercial traffic, our research group has gained access to a massive commercial vehicle tracking dataset. Efforts are underway to analyse this data to develop an activity-based model of commercial transport.

Simulation

MATSim requires a transport network description to execute commuter plans in the simulation step. A valid network description is a weighted digraph of nodes and links that should satisfy the following two conditions:

1. Each link must specify its flow capacity, usually in terms of vehicles per hour, along with its length in meters, free speed in meters per second, and number of lanes.
2. Each node in the network should be reachable from any other node, to prevent vehicles from becoming trapped.

Network data are relatively easy to come by, due to the proliferation of the Global Positioning Sysem (GPS) as a navigation aid in private and commercial vehicles. Generally, however, the flow capacities of network links aren't explicitly specified, and either has to be assigned heuristically based on free speed and number of lanes, or has to be determined through a traffic count.

Our research group obtained a Geographic Information System (GIS) shapefile depicting the Gauteng road transport network from Business Connexion GIS, a Gauteng-based business solutions provider. The GIS dataset forms part of their business intelligence product range, and is known to be valid for 2007. Each line shape in the file denotes the centre line of a road segment, tagged with attributes, including the type of road, speed limit, measured length and topology. We interpreted the road type and speed limit fields to assign flow capacities and number of lanes to each network link.

Initial results for private vehicle traffic suggest that network misclassification is an important contributor to simulation error and, consequently, we are investigating the possibility of using image recognition techniques on aerial photography data to improve our network description.

Analysis

Simulation output needs to be validated against some real-world measures of transport system performance. In MATSim, simulated hourly traffic counts are compared with actual data in order to determine simulation accuracy. As a full second-by-second record exists of each vehicles simulated movement, it would be possible to stratify simulated counts using arbitrary parameters such as vehicle type or weight, in order to perform more detailed validation.

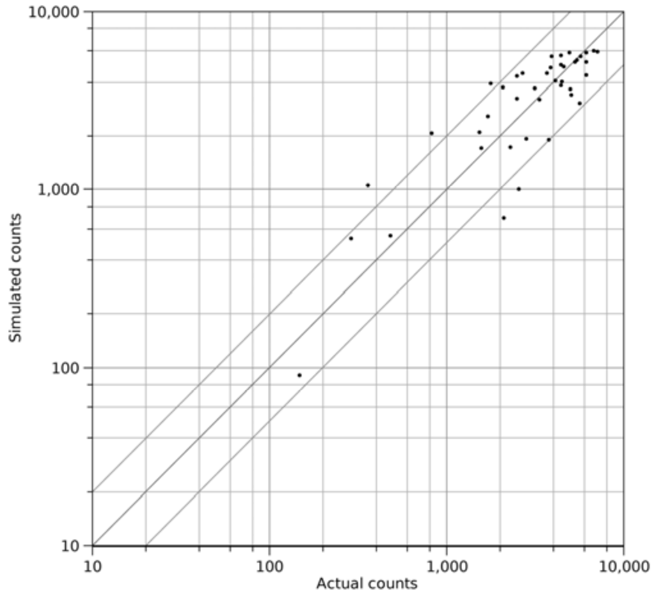

Figure 4. A log-log scatter plot summarises the correlation between simulated and actual traffic counts during the morning traffic peak for all 20 pairs of counting stations in Gauteng used in our preliminary study.

For our initial implementation we compare simulated traffic counts against the actual 2001 South African National Roads Agency Ltd. (SANRAL) counting station data from a selection of 20 pairs of network links in the Gauteng area. These stations are mostly situated near important intersections on the main arterial routes in Gauteng. We are in the process of capturing count station data for the complement of stations on record.

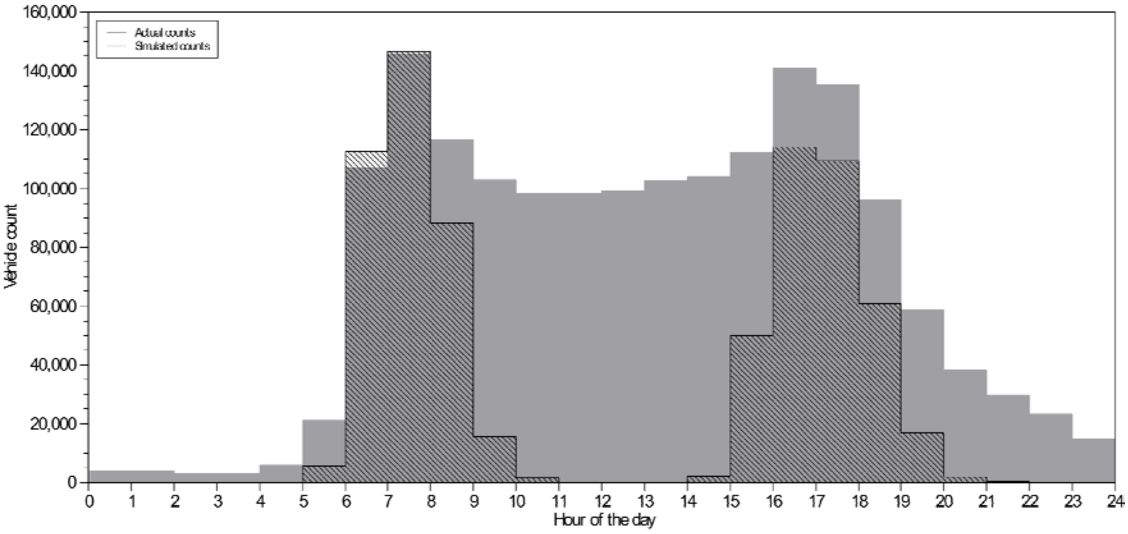

Figure 5 Total simulated vs. actual counts of all recorded count stations for the entire work day. Our initial implementation only models the home-work-home trips, assumed to be responsible for the morning and afternoon traffic peaks.

Figures 4 and 5 shows our initial results in two formats. Figure 4 is a detailed counts comparison for the morning peak, showing our largest deviations from perfect prediction on intermediate flow links. From the histogram in Figure 5 can be seen that the MATSim demand modelling process has changed the initial timings of activities, as described in Section 3.1, to arrive at a reasonable fit with the timings observed in reality during the morning and afternoon traffic peaks. In general, it appears that the scale of the commuter population that we generated from census data is of the correct order of magnitude, and the virtual commuters organise themselves in a similar fashion to their real-world counterparts. Their detail routing on the network differs for links of intermediate capacity, however. It appears that these deviations are due to flow rate misclassifications on feeder routes in our virtual network, and we expect to improve as we develop a more accurate network description. In a forthcoming paper we will discuss our results in more detail.

Besides traffic counts comparisons, we foresee that public transport ridership figures will become an important validation measure once our public transport implementation is operational. Although data sources still need to be identified, it is expected that ticketed modes should be well accounted for in the form of ticket sales data.

Unfortunately, the minibus taxi para-transit mode operates without any formal record of passenger volumes, and will therefore require primary data sampling in order to validate ridership. Several methods to collect such data can be conceived, but the willingness of operators to participate is questionable, due to the competitive and confidential nature of ridership information.

Traffic counts comparisons aren't the only forms of analysis open to the transport demand planner, as MATSim records the full simulation behaviour of all agents in text format. These data can be interpreted to derive detailed, time-dependent measures of transport system efficiency, such as emissions estimates, average system travel speed, or even overall flow rate vs. average travel speed for the entire system. Therefore, even when using limited traditional transport demand data, as in this initial implementation, it becomes possible to achieve far richer results than would be possible for traditional demand modelling methods.

CONCLUSION

Our experience with implementing MATSim for private vehicle traffic in Gauteng has confirmed that data availability need not represent a major hurdle in implementing a modern transport demand planning technology. Available (aggregate) data sources such as census data and the NHTS proved useful in generating a (disaggregate) model of private vehicle transport demand, and we expect to also use this data in our initial implementation of public transport.

The most important challenges to extending our current model pertain to public transport. Very little information is available on the transport provider side, which presents difficulties both to modelling and validation. However, we suspect that this knowledge gap is important to more stakeholders than just our research group. We expect that the modernisation of public transport in Gauteng will create an urgent need for updated transport demand information.

The success of expensive infrastructure investments such as Bus Rapid Transit (BRT) and the Gautrain rapid rail system will depend on how these modes integrate with existing public transport options. The only way to achieve integration of modes is by gaining knowledge on the detail operation and behaviour of each mode. Our hope is that the relevant stakeholders will come to this realisation, and make efforts to close this major knowledge gap.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the following persons and institutions:

- Prof. Dr. Kai Nagel and the MATSim development team from Transport Systems Planning and Transport Telematics, Institute for Land and Sea Transport Systems, Technische Universität Berlin, where Mr. Fourie was received for two months in the European summer to lay the groundwork for this study.
- Business Connexion GIS, for their GIS dataset of the Gauteng transport network.
- Mr. Piet Alberts from Stats SA, who arranged to provide us with Census 2001 and NHTS data.

This work is based upon research supported by the National Research Foundation (NRF) (grant FA2007051100019). The authors also acknowledge financial aid from the University of Pretoria and the NRF Knowledge Interchange and Collaboration programme (UID 67539), which made possible the research visit to Berlin during June and July 2008.

REFERENCES

Balmer, M. 2007. Travel Demand Modeling for Multi-Agent Transport Simulations: Algorithms and Systems. PhD thesis, ETH Zürich.

Balmer, M., Axhausen, K. and Nagel, K. 2006. Agent-Based Demand-Modeling Framework for Large-Scale Microsimulations. Transportation Research Record: 125–134.

Balmer, M. and Rieser, M. 2004. Generating daily activity chains from origin-destination matrices. In Transportation Research Board 84th Annual Meeting.

Beckman, R., Baggerly, K., and McKay, M. 1996. Creating synthetic baseline populations. Transportation Research Part A, 30(6):415–429.

Behrens, R. 2000. Activity-travel analysis: a review of the theoretical origins, recent developments and local application. Paper presented to the South African Transport Conference, South Africa, 17 - 20 July.

Engelbrecht, A. 2002. Computational Intelligence: An Introduction. J. Wiley & Sons.
Rieser, M. 2007. Simulating road pricing/tolls with MATSim.

Stats SA. 2001. Census 2001 metadata: Geography hierarchy and attributes. Technical Report 03-02-25 (2001), Statistics South Africa.