COMMERCIAL VEHICLE BEHAVIOUR: ANALYSING GPS RECORDS

Q. van Heerden1* and J.W. Joubert2
1 Department of Industrial & Systems Engineering
University of Pretoria, South Africa
quintin.vh@gmail.com

2 Department of Industrial & Systems Engineering
University of Pretoria, South Africa
johan.joubert@up.ac.za

ABSTRACT

The state-of-practice in commercial vehicle modelling is based on Gross Domestic Product (GDP) and aggregate economic sectors. There is a lack of understanding of the explicit activity chain characteristics of commercial vehicles. The aim of this research is to contribute to the body of knowledge on commercial vehicle activity chain characteristics by distinguishing, both temporally and spatially, between intra- and inter-provincial activity chains. Activity chains are extracted from the GPS logs of 41 711 commercial vehicles and analysed for different days of the week. The land-locked province of Gauteng, South Africa, is used as the study area since it contributes about 35% to the country's GDP. It can be considered the centre point for all major freight movements, imports, exports, and local distributions. Inter-provincial vehicles have not been considered in past research and are now the focus point in this paper. Inter-provincial vehicles are analysed as to where they enter and exit the study area and the number of activities they perform inside the area. Past research referred to inter-provincial traffic as through-traffic, but the contribution in this paper shows that the majority of inter-provincial chains enter and exit the study area through the same gateway, countering the idea of travelling through.

* Corresponding Author
1 INTRODUCTION

Proper transport infrastructure is required for both private and commercial vehicles as they compete and participate in various activities while using the same road network. Both the maintenance and development of road infrastructure and road networks are dependent on strategic decision making and transport planning, for which the government is largely responsible.

Vehicle movement models are crucial to assist in strategic decision making and transport planning. Simulation models create a safe environment in which to test different “what-if” scenarios within the transportation domain. What if the number of vehicles on the road should double in less time than expected? What if a new road is built as an alternative route?

In an attempt to model traffic behaviour, much emphasis has been placed on the modelling of private vehicle movement. While passenger and private vehicle models have been modelled with great success, these models are often merely inflated by a factor to reflect commercial traffic as background noise. These models reflect commercial vehicle behaviour to be similar to private vehicle behaviour. Commercial vehicle behaviour, though, differs quite distinctly from that of private vehicles with regard to vehicle mix, motivations, and the number of activities during a trip [1]. It is imperative to model commercial vehicle movement independently from private vehicle models.

Various commercial vehicle studies exist [1; 2; 3; 4; 5; 6] in which different approaches are used, of which the most common is transposing commodity flows into origin-destination (OD) matrices, and a three-step zonal approach, much like the Four Step Model (FSM).

The current state-of-practice in commercial vehicle modelling is outdated, still aggregate, and based on Gross Domestic Product (GDP) and economic sectors [2]. It is also based on private vehicle models, mostly the traditional FSM.

Fourie [7] compares an agent-based model to an EMME/2 model using the same data and scenario and concludes that the agent-based approach results in much richer time dependent results. He also shows that one can predict travel times more accurately with agent-based modelling. To use an agent-based approach requires disaggregate data. Commercial vehicle activity chains need to be studied explicitly and understood at a disaggregate level.

In an attempt to understand the activity- and chain-characteristics of commercial vehicles, Joubert and Axhausen [8] analyse the temporal and spatial characteristics of disaggregated commercial vehicle activities. The study includes the extraction of detailed activity-chains from GPS logs of more than 31 000 commercial vehicles over a six month period. These activity-chains are then analysed with regard to certain characteristics such as the number of activities per chain, activity durations, chain durations, and the start time of chains. The study is a great leap in the direction of understanding commercial vehicles’ movements at disaggregates level. However, there is still room to explore the differences between inter- and intra-provincial activity chains. A vehicle passing through a study area may have a different travelling behaviour to that of a vehicle that is confined to the study area.

Joubert et al. [9] build on the work of Joubert and Axhausen [8] to test the modelling approach. They use an agent based approach to implement reconstructed commercial vehicle activity chains in conjunction with private vehicles. They show that detailed movements of commercial vehicles can be modelled accurately without having to model complex logistical functions. The model is tested using the Multi-Agent Transport Simulation (MATSim) toolkit and the results show that an activity based approach in the modelling of freight vehicles, and the impact thereof on private vehicle movement, can be done
accurately. They state that the next step would be to generate through-traffic activity chains and compare those against actual traffic counts.

2 ACTIVITY CHAIN EXTRACTION

Digicore Fleet Management, a vehicle tracking service provider in South Africa, provided a dataset for this research. The data contains the detailed GPS logs of 41 711 commercial vehicles for the period 1 January 2009 to 30 June 2009. These vehicles represent about 1.8% of the total light delivery vehicle and heavy vehicle population in South Africa [10]. Since these vehicles are only from Digicore's tracking service subscribers, a possible selection bias is acknowledged, yet this data is invaluable for research purposes.

The same extraction process as in Joubert and Axhausen [8] is used to extract activity chains from the GPS-logs. Ignition-on and -off signals were used to determine when a vehicle is switched on or off, which also signalled the end or start of an activity.

A 300 minute threshold of activity duration, as determined by Joubert and Axhausen [8], is used to distinguish between minor and major activities. Major activities, typically depot stops, exceed 300 minutes; and minor activities, typically refuelling or drop-off and collection activities are less than 300 minutes in duration.

Whereas Joubert and Axhausen [8] only considered chains which contained at least one minor activity, in this research all chains, consisting of only two major activities as well as containing any number of minor activities, are considered. Also, a density-based clustering approach, to cluster activities within a specified radius and a specified minimum number of points, has been omitted.

In previous research, assumptions were made that the activity chain characteristics are similar on different days of the week. For each day of the week, an initial analysis was done to determine the number of commercial vehicles with at least one activity in Gauteng. From all the calendar days, we considered two subsets separately. Firstly we extracted official public holidays. Secondly we extracted abnormal days: those days, inclusive of public holidays, that are considered abnormal by SANRAL (South African National Roads Agency Limited) in terms of traffic patterns. Examples of such abnormal days include days between weekends and public holidays when people are expected to create long weekends, as well as major school holidays. Figure 1 depicts a box plot of the number of vehicles on each day of a typical week as well as the abnormal days and public holidays.

![Figure 1: Number of vehicles with at least one activity in Gauteng](image-url)
Weekdays tend to show a similar behaviour, whereas weekends have much lower activity. Outlier 1 has a vehicle count of almost half of the average counts on Wednesdays. This abnormal activity can be attributed to the fact that it was 22/04/2009, the national election day. Companies allowed workers to visit polling stations to vote and subsequently less vehicles were on the roads. The following day saw an increase in freight movement and more vehicles than the average for a Thursday were on the road, shown by outlier 2.

All subsequent analyses were done on a day-of-the-week basis. It is important to understand which activity chain characteristic differences, if any, exist between inter- and intra-provincial activity chains.

3 ACTIVITY CHAIN CHARACTERISTICS

In South Africa, Gauteng can be considered as the centre point for all major freight movements, imports, exports, and local distributions. Gauteng contributes almost 35.1% towards the country's Gross Domestic Product (GDP) [11]. This is the reason why it is referred to as the "economic heart" of South Africa and whereas the road network that runs through it, is referred to as its "economic arteries" [12]. Since Gauteng is the economic heart of South Africa, with various gateways that link it to the rest of South Africa, and even Africa, Gauteng has both inter- and intra-provincial traffic.

Joubert and Axhausen [8] determine that 60% is a suitable threshold to distinguish between inter- and inter-provincial vehicles. If more than 60% of a vehicle's activities are inside the study area, it is considered to be an intra-provincial vehicle. If less than 60% of a vehicle's activities are inside the study area, it is considered to be an inter-provincial vehicle. The 60% threshold was used to split the vehicles into inter- and intra-provincial vehicles with the split being 77:23.

3.1 Chain length

The chain length was calculated as the number of activities per chain. For this metric, all activities in a vehicle's chain were considered and not just activities inside the study area. Summary statistics for inter-provincial vehicles are shown in Table 1.

Table 1: Summary statistics of the number of activities per chain for inter-provincial vehicles

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Mean</th>
<th>Std dev</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>95th</th>
<th>99th</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>13.74</td>
<td>27.99</td>
<td>3</td>
<td>9</td>
<td>17</td>
<td>27</td>
<td>38</td>
<td>101</td>
<td>4811</td>
</tr>
<tr>
<td>Tuesday</td>
<td>12.22</td>
<td>23.14</td>
<td>3</td>
<td>9</td>
<td>16</td>
<td>25</td>
<td>32</td>
<td>68</td>
<td>4486</td>
</tr>
<tr>
<td>Wednesday</td>
<td>12.14</td>
<td>19.39</td>
<td>3</td>
<td>9</td>
<td>16</td>
<td>25</td>
<td>33</td>
<td>64</td>
<td>2575</td>
</tr>
<tr>
<td>Thursday</td>
<td>11.95</td>
<td>19.17</td>
<td>3</td>
<td>9</td>
<td>16</td>
<td>25</td>
<td>32</td>
<td>60</td>
<td>3768</td>
</tr>
<tr>
<td>Friday</td>
<td>11.53</td>
<td>21.11</td>
<td>3</td>
<td>9</td>
<td>16</td>
<td>24</td>
<td>30</td>
<td>54</td>
<td>2815</td>
</tr>
<tr>
<td>Saturday</td>
<td>8.66</td>
<td>26.63</td>
<td>1</td>
<td>5</td>
<td>11</td>
<td>19</td>
<td>26</td>
<td>51</td>
<td>4903</td>
</tr>
<tr>
<td>Sunday</td>
<td>8.91</td>
<td>26.81</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>19</td>
<td>29</td>
<td>79</td>
<td>2936</td>
</tr>
<tr>
<td>Abnormal Day</td>
<td>11.05</td>
<td>23.55</td>
<td>2</td>
<td>7</td>
<td>15</td>
<td>23</td>
<td>31</td>
<td>66</td>
<td>4663</td>
</tr>
<tr>
<td>Public Holiday</td>
<td>10.42</td>
<td>31.62</td>
<td>1</td>
<td>6</td>
<td>13</td>
<td>22</td>
<td>30</td>
<td>73</td>
<td>3159</td>
</tr>
</tbody>
</table>

For both inter- and intra-provincial vehicles, the chain length decreases as the week progresses from Monday to Friday. Weekends tend to have shorter chains than the rest of
the week or abnormal days and public holidays. Some businesses are closed on weekends or only open for a part of the day, thereby limiting the actual activity that could take place.

The chain lengths of inter- and intra-provincial vehicles were found to be very similar. At first glance this could be rather counter intuitive, since it is expected that inter-provincial vehicles, which are associated with long haul freight movements, should have fewer activities per chain.

3.2 Chain start and end times

Figure 2a and 2b depict the density of chain start and end times of intra-provincial vehicles for a typical weekday and weekend. The density distributions of the start times on the different days of the week do not differ much. Intra- and inter-provincial vehicles' chain start and end time characteristics were also found to be very similar. All weekdays' patterns are also similar. It is important to note that although the patterns are similar, the actual number of vehicles differ on the different weekdays.

On a Sunday, the chain start time density is slightly lower than on a weekday. This phenomenon similarly appears on a public holiday. Another similar feature is that on weekdays, weekends and public holidays, the peak chain start times are between 06:00 and 08:00, namely peak hour traffic. Commercial vehicles move slower than other vehicles and thereby contribute to peak hour traffic congestion. For chain end times, there is a peak in the late afternoon on weekdays, but on weekends the distribution is rather uniform towards the end of the day, with a slight peak between 18:00 and 20:00, as can be seen in figure 2b. Whereas Joubert and Axhausen [8] found that 60% to 87% of chains have ended before the afternoon peak starts at 16:00, in this research only 35% to 52% of activity chains have ended before the afternoon peak on the various days of the week. This means that the rest of the commercial vehicles are still on the road during peak hour traffic and contribute to peak hour traffic congestion. A difference in results between 2008 and 2009 is acknowledged, but no assignable causes could as yet be identified and this can be investigated in future research.

3.3 Chain duration

For both inter- and intra-provincial vehicles, the mean chain duration drops slightly towards the end of the week. An interesting observation is that 94.7% of all vehicles have chain durations of 24 hours or less. Some of the chains continue from one day over into the next
day and for modelling purposes, in future research, these chains may need to be divided into separate chains as in Joubert et al. [9].

3.4 Activity start time

Since activity chains consist of activities, it would also be beneficial to know when the activities take place during the day. Figures 3a and 3b depict the minor activity start times of both inter- and intra-provincial vehicles for a weekday. The inter- and intra-provincial vehicles were plotted on the same graph to determine whether differences exist between their activity start time distributions.

![Minor activity start times](image1)

![Major activity start times](image2)

**Figure 3: Activity start times of inter- and intra-provincial vehicles**

The minor activity starts time distributions for inter- and intra-provincial vehicles are similar for all weekdays where they tend to peak in the middle of the day. This confirmed the results of Hunt and Stefan [1] that commercial vehicle activities concentrate towards the middle of the workday and not as much towards the morning and afternoon peaks. However, since minor activities are typically drop-off and collection activities, the activity chains had to start earlier than the start times of these minor activities. This means that vehicles could have started their trips during the morning rush hour.

From the data, it was observed that a substantial percentage (77%) of the commercial vehicles travelling in the province of Gauteng, South Africa, is inter-provincial vehicles, emphasising the need to understand their unique travel behaviour.

4 INTER-PROVINCIAL TRAFFIC

The average flow of vehicles through the 8 major gateways into and out of Gauteng was determined, over the 182-day period from 1 January 2009 to 30 June 2009, for each day of the week. Figure 4 shows the temporal distribution of the number of vehicles that enter and exit Gauteng through the 8 major gateways on a Monday.
At all the gateways, the number of vehicles that entered and then left the study area on a weekday, exceeded the number of vehicles on weekends. It is however noted that abnormal days had a significantly higher activity than weekdays, specifically at gateways 1, 2, and 7, suggesting that organisations continue with, and even increase, their business during the school holidays. On public holidays, however, a much lower count was noted than on any other day of the week. This pattern on different days of the week is evident at all gates. Contrary to the majority of the gates, gate 2 does not have the morning and afternoon peaks, but instead has a constant inflow and outflow of vehicles throughout the day.

As expected, gateway 7, which is associated with the N1 towards Cape Town, has a high gate activity since Cape Town has the country's second largest port that accounts for many freight movements to and from Cape Town.
Gate 4 links Gauteng with the Limpopo province as well as the rest of Southern Africa. Being referred to as the "gateway into Africa", gate 4 was expected to have a high volume of vehicles as well. This was not the case, though. It is possible that very few of Digicore's subscribers are cross-border carriers, or conduct only a few of their activities in the Limpopo province.

Next, we wanted to determine how many activities inter-provincial vehicles perform within the study area before eventually leaving it. An analysis was done and an exponential distribution was fitted to the data. This yielded a probability density function $\lambda e^{-\lambda x}$ with a rate parameter $\lambda$ of 0.1591 and estimated standard error of 0.0004.

Whereas Joubert and Axhausen [8] found that 90% of vehicles have 4 activities or less within the study area, the findings in this research reveal that only 50% of vehicles perform 4 activities or less in the study area. Instead, 90% of vehicles performed a higher number of 14 activities or less within the area. There was thus an increase, between 2008 and 2009, in the number of activities that vehicles have within the study area before departing from it. In density based clustering, if activities in close proximity are clustered together, chains consist of fewer activities. Since density based clustering was not used in this research, it could be a valid cause leading to the longer activity chains in this research. Ongoing research is considering and evaluating this discrepancy.

The bimodality, i.e. peaks of entries in the morning and exits in the afternoon at the same gate, suggests that there is a net inflow of vehicles into Gauteng in the morning and a net outflow of vehicles from Gauteng in the afternoon. This raised the question whether the same vehicles enter and exit through the same gateway.

To investigate this, consecutive activities were paired and referred to as $(a, b)$. If both $a$ and $b$ were inside or outside Gauteng, the pair was ignored. If activity $a$ was inside Gauteng and activity $b$ was outside Gauteng, an exit was noted. The point of exit was determined by connecting point $a$ with point $b$ by means of a straight line. The point where the line intersected the border of Gauteng was noted as point $c$. The nearest gate to point $c$ was considered to be the exit gate. Similarly, an entry was registered whenever $a$ was outside and $b$ was inside Gauteng.

The nearest gate to point $c$ was considered to be the entry gate. From the activity chains it is possible to determine the activity start and end times. The time at which the entry or exit occurred was estimated by calculating the fraction $f$, which is the straight line distance between $a$ and the point of entry or exit, $c$, divided by the straight line distance between $a$ and $b$. From the activity chains, the duration $d$ was determined as the time that elapsed from the end of activity $a$, denoted as $a_{\text{end}}$, to the beginning of activity $b$, denoted as $b_{\text{begin}}$. The duration $d$ was denoted as $d = b_{\text{begin}} - a_{\text{end}}$. The estimated time of entry or exit, denoted as $e$, was then determined using both the duration and fraction measures, as $e = a_{\text{end}} + fd$.

The activity chains were split into in-out chains, chains that start outside the study area, enter, perform a few activities and eventually leave the area; and out-in chains, chains that start inside the area, exit, perform a few activities outside, and eventually return. The results captured in table 2 show fractions of in-out chains that entered at a specific gate (from), and then exited at specific gates (to). The percentage of the total exits is also shown. Fractions under 5% are omitted from the table. To ease the interpretation of the table, the cells are colour coded according to the weight of the fraction in the cell. The higher the fraction, the darker the shade of grey.
Table 2: Fraction of gate activity for in-out chains on a Monday

<table>
<thead>
<tr>
<th>From</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Percentage of total exits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.640</td>
<td>0.125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.094</td>
<td>21%</td>
</tr>
<tr>
<td>2</td>
<td>0.087</td>
<td>0.704</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.092</td>
<td>19%</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.384</td>
<td>0.413</td>
<td></td>
<td></td>
<td></td>
<td>0.082</td>
<td></td>
<td>9%</td>
</tr>
<tr>
<td>4</td>
<td>0.115</td>
<td>0.109</td>
<td></td>
<td>0.427</td>
<td>0.113</td>
<td></td>
<td>0.144</td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>5</td>
<td>0.055</td>
<td>0.062</td>
<td>0.069</td>
<td>0.055</td>
<td>0.671</td>
<td></td>
<td></td>
<td></td>
<td>9%</td>
</tr>
<tr>
<td>6</td>
<td>0.067</td>
<td>0.116</td>
<td></td>
<td></td>
<td></td>
<td>0.458</td>
<td>0.211</td>
<td>0.073</td>
<td>7%</td>
</tr>
<tr>
<td>7</td>
<td>0.077</td>
<td>0.126</td>
<td></td>
<td></td>
<td></td>
<td>0.638</td>
<td></td>
<td></td>
<td>23%</td>
</tr>
</tbody>
</table>
| 8    | 0.105| 0.091|      | 0.154| 0.097| 0.123| 0.385|      | 7%                        

Although this was repeated for out-in chains and for each day of the week, only one table is shown due to the similar trend in the data. The darker diagonal is much more evident in out-in tables than in in-out tables. Out-in chains are therefore much more likely to exit and enter at the same gate than in-out chains. The total number of in-out chains is almost double that of out-in chains on most days, which suggests that there are a number of distribution centres situated outside Gauteng.

The prominent diagonal feature in in-out pairs correlates well with Figure 2 and confirms the morning and afternoon peaks at the same gate, especially for gates 1 and 7.

For out-in chains that exit through gate 3, 79% return through gate 3, while 11.3% return through gate 2. This could be vehicles that make use of the road network outside Gauteng once they leave Gauteng, and eventually re-enter Gauteng through gate 2, or vice versa.

Overall, out-in vehicles tend to enter the study area through the same gate they exited. In-out vehicles showed the same pattern except that they also tend to exit at other gates than where they entered the study area. A possible explanation could be that the road network inside Gauteng is better than outside, or that pickups and deliveries take place throughout Gauteng, leading to the vehicles considering other routes to exit Gauteng.

5 CONCLUSION

This research added to the body of knowledge in commercial vehicle behaviour as the differences in characteristics and structure of intra- and inter-provincial vehicles' activity chains were analysed in detail. It was determined that chain characteristics should be modelled separately for weekdays, weekends, abnormal days, and public holidays.

Chain end times and the density based clustering approach were identified as areas for further research.

Special focus was placed on inter-provincial vehicles to understand where the chains start and end, where they enter and exit the study area, whether they enter and exit the study area through the same gate, all while considering it on different days of the week and for different times of the day.

The next step would be, similar to Joubert et al. [9], to generate a synthetic population consisting of both inter- and intra-provincial commercial vehicles and simulating a full combined private and commercial vehicle model and analysing and validating the results.
Such a model will be more accurate since it accounts for commercial vehicles separately from private vehicles. These models can then be used to improve decision making and a wiser expenditure of tax money.

6 REFERENCES


