

IMPROVING NETWORK RELIABILITY VIA INCIDENT MANAGEMENT

Alan Nicholson, University of Canterbury, Christchurch, New Zealand;
Alan.Nicholson@canterbury.ac.nz

Susan McMillan, NZ Transport Agency, Auckland, New Zealand;
Susan.McMillan@nzta.govt.nz

Glen Koorey, University of Canterbury, Christchurch, New Zealand;
Glen.Koorey@canterbury.ac.nz

ABSTRACT

Network reliability has become a major issue, due to increasing traffic congestion and the adverse impact on road users, especially those who have embraced the just-in-time philosophy. Traffic incidents (e.g. accidents) are a major contributor to a lack of reliability, and there is thus an increasing interest in improved incident management. This paper describes a study involving linking a microsimulation model (S-Paramics) with the SCATS traffic signal control software, to assess the benefits of adjusting traffic signal timings to mitigate the effects of incidents. While a case study indicates that the benefits of using SCATS as an incident management tool are not large, it is concluded that linking traffic signal control and microsimulation software appears to be a promising approach for developing incident management plans to improve network reliability.

Keywords: network reliability, incident management, microsimulation, adaptive signal control.

INTRODUCTION

The economies in developed countries depend heavily on their transport systems, and that dependence has increased with the adoption of 'just-in-time' production methods, the success of which depends upon the transport service being of a high quality. Surveys of transportation system users (e.g. Parkhurst et al., 1992) have shown that reliability is one of the most important determinants of transport service quality.

Nicholson and Du (1997) show that unreliability can be considered to arise from two distinctly different sources;

- demand (or flow) variations, such as day-to-day variations in the number of vehicle trips;
- supply (or capacity) variations, such as capacity reductions due to lane blockages as a result of vehicle breakdowns.

In reality, travel time variation can arise from either or both sources, and it is not always an easy matter to identify the separate effects of flow and capacity variations. For instance, if an accident occurs during the early part of a peak period and results in a road being partly blocked, it may well be difficult to separate the effect of the capacity reduction and the increasing traffic flow. In major urban areas, where the networks are typically dense and congested, the social and economic impacts of variations can be substantial, even when the variations are short, while in rural areas, where the networks are typically sparse and less congested, the impacts can be substantial if the variations are long (Nicholson, 2007).

There have in recent years been increasing efforts to mitigate the impacts of disruption of road transport networks and to develop methods for assessing the impacts (including changes in reliability) within economic appraisals of road network improvement proposals. A study for the UK Department of Transport (SACTRA, 1999) concluded that ignoring the effect of travel time variability led to the economic benefits of trunk road projects being underestimated by between 5% and 50%. While a subsequent UK study (Eddington, 2006) also stressed the importance of accounting for the reliability of travel time, de Jong et al. (2009) have commented that while major transport infrastructure projects are commonly assessed using cost-benefit analysis, "changes in the reliability of travel time are not incorporated in standard appraisals ... in The Netherlands or in other countries". In fact, New Zealand (NZ) has allowed changes in travel time reliability to be taken into account in the economic appraisal of transport infrastructure projects for over five years (Transfund NZ, 2004), but is perhaps the only country to do so.

As noted by Nicholson et al. (2003), there are several different definitions and measures of reliability. Travel time reliability, as measured by the standard deviation of travel time, is an attractive measure, because it lends itself to incorporation in procedures for the economic appraisal of transport projects. The NZ economic appraisal method (Transfund NZ, 2004; NZTA, 2010) uses this measure.

Goodwin (1992) suggested that network reliability be improved by upgrading or constructing links, to achieve or maintain a "quality margin" by planning and designing the system so that it has some spare capacity. It is interesting to note that upgrading and construction of links in the Auckland (NZ) motorway system is currently planned or under way, with an improvement in reliability frequently being used as a justification (NZTA, 2009). A transport system can be considered a system of queues, and queueing theory indicates (Wolff, 1989) that the travel time variance for a queueing system starts to increase rapidly at a flow-to-capacity ratio that is much lower than the flow-to-capacity ratio at which the mean travel time starts to increase rapidly. It is common traffic engineering practice to design for a 'practical capacity' that is 80% of the theoretical capacity (i.e. a 20% 'safety margin'), so that mean travel times are not excessive; designing to avoid large travel time variance will entail designing for a much lower 'practical capacity' (i.e. a much larger 'safety margin'). Preserving that larger 'safety margin' is very likely to be difficult, unless tolls are used to manage the travel demand.

This paper describes research into incident management as a means of improving reliability. After a brief discussion of the importance of incident management, the paper describes a

study involving linking a microsimulation model with traffic signal control software, to assess the benefits of adjusting traffic signal timings to mitigate the effects of incidents.

INCIDENT MANAGEMENT

Traffic congestion can be divided into two categories:

- 'recurring' congestion, which occurs when there is regularly not enough capacity in the transportation system to meet the demand (e.g. during "peak-hour" periods of the day);
- 'non-recurring' congestion, which occurs when there is a temporary and 'unexpected' reduction in capacity due to incidents (e.g. crashes, spills, weather, events, etc.).

Both types of congestion are important, with incidents (i.e. non-recurring congestion) being estimated to be responsible for about half of the congestion on US freeways (US DoT, 2000), while Schrank and Lomax (2009) have estimated that 52-58% of total motorist delay in urban areas is caused by crashes and breakdowns.

It has been suggested (FHWA, 2006) that drivers experience variations in travel time and expect therefore make allowance for some variation when planning trips. Instances of congestion and delay much greater than they have allowed for, albeit only occasional, will have a major impact on their perception of the reliability of the road network (i.e. they readily remember such events). This is consistent with recommendation of the Scottish Executive (2005) that journey time reliability be based on the proportion of trips taking more than 115% of the average journey time for that same period; this suggests that drivers expect the travel time to vary up to 115% of the expected travel time, with greater variations being perceived as evidence of unreliability.

It seems likely that variability associated with recurring congestion will be allowed for in trip planning, with the occasional instances of unexpectedly high delay being largely associated with non-recurring congestion. Hence, it seems likely that reducing non-recurring congestion will have a greater effect on the perception of reliability than reducing recurrent congestion.

Incidents causing temporary capacity reductions include vehicle-based incidents (e.g. vehicle breakdowns, accidents), other objects or obstructions on the road (e.g. debris from a landslide), roadway and utility maintenance activities, and extreme weather events (e.g. heavy rain or hailstorms). There are events that might not be expected by all road users, but which are planned events and are publicly notified (e.g. maintenance activities, sports/cultural activities). This research focuses on unplanned events.

Many incidents tend to occur during peak periods when the network is already operating at or near capacity and any disruption in the traffic flow can have a significant effect. This is not surprising, given that perturbations in traffic flow are more likely to propagate and be amplified as traffic density increases. Incidents can therefore have major impacts on transport network reliability during peak periods.

Successful incident management on road networks requires two key elements:

- quick and accurate detection of incidents (including identifying their location and nature);
- prompt and effective remedial treatment, in terms of both removing or fixing the cause of the incident (if possible) and managing vehicle flows during and after the incident.

Methods for incident detection could include, for example, telephone calls from the public, police and highway patrols, and automated incident detection systems. Incident verification is important, to maximise accuracy in the location and nature, and to minimise false alarms. Automated Incident Detection (AID) uses real time traffic data and specific algorithms to identify incidents. Williams & Guin (2007) reviewed various algorithms used since the 1970's. While early algorithms were largely based on simple speed or occupancy comparisons and classical traffic flow theory, more complex techniques such as neural networks (e.g. Srinivasan et al., 2004), fuzzy logic (e.g. Yaguang and Anke, 2006) and wavelet transformations (e.g. Samant and Adeli, 2000) have subsequently been used. Williams and Guin surveyed traffic management centres across the USA, and found widespread concern with the level of performance with existing AID algorithms, with 81% stating that even if reliable and accurate AID algorithms were widely available in the future, they would complement but not replace the other methods.

While incident detection algorithms for high-speed freeways are well established, this is not the case for lower-speed urban arterials. One study of incident detection for such roads (Zhang & Taylor, 2006) involved researching lane-blocking incidents (including the entire link being blocked) between two adjacent intersections on an arterial network. Due to a lack of real incident data on the arterial network, incidents were modelled using Q-Paramics software. They concluded that more detectors than is current practice would be required to successfully detect incidents on urban arterials.

Incident treatment can be preventive (e.g. limiting the flow via tolling or controlling access) or reactive (e.g. dissemination of real time traffic information to users, using variable message signs (VMS) to inform motorists of traffic conditions and suggest detours, adjusting traffic signal timings). Adaptive traffic control systems may also adjust signal timings in response to observed changes in traffic demands due to an incident, although the default response rate is often relatively slow. While VMS can be effective, motorists who are familiar with the network will often ignore the suggested routing and select their own perceived optimal route, while motorists who are not familiar with the network will often not take detours because they are unsure of their reliability. Hidas (2001) found that traffic diversion rates for VMS routing can range from 5% to 80%, with the most important factor affecting the decision to divert being network familiarity.

It is not appropriate to use classical equilibrium methods for investigating the effect of incidents causing short disruptions to traffic flow, as the traffic flow pattern is unlikely to change sufficiently quickly to reach an equilibrium situation until some time after the end of the period of disruption. It is therefore necessary to use a simulation approach, say, to allow for the traffic flow pattern changing during the period of disruption.

It is difficult to quickly identify the most effective treatment for a particular incident; the variety of locations and flow conditions invariably mean that no two incidents are exactly the same. To date, this has made it particularly difficult for automated systems to provide more effective treatment guidance than that determined by “human experience”. Research into the most appropriate treatment options is also complicated by the difficulty in collecting suitable field data from incidents for analysis.

The study described here involved using microsimulation (S-Paramics) to investigate the effectiveness of the Sydney Coordinated Adaptive Traffic System (SCATS) in dealing with the effects of an incident on a section of the Auckland Northern Motorway (i.e. the diversion of traffic from the motorway onto nearby arterial roads). SCATS is a real time adaptive signal control system used extensively by road controlling agencies in Australia and NZ and is also used in Asia, North and South America, and Europe.

STUDY METHOD

Microsimulation models generally assume fixed traffic signal phase times, and it has been common practice to model SCATS controlled intersections in microsimulation models by using the average signal timings. For options testing, the splits, phasing, cycle length and offsets can be optimised in an external signal optimisation package for input into a microsimulation model. However, in reality, the modelled signals are actually likely to be controlled by SCATS which adapts in real time in response to current traffic patterns. It is difficult to effectively model SCATS controlled signals using fixed-time signal plans (Zhang and Taylor, 2006), and this is a particular problem when modelling incidents, as it means that the model does not allow for changes in signal timings to reflect changes in traffic flows due to incidents. It was therefore important to link the microsimulation model directly to SCATS, and this was done using the FUSE software (baseplus, 2007), which enables S-Paramics to use SCATS signal timings and allows the model to determine how SCATS will adapt to changes in traffic flows as a result of incidents.

Incidents can be modelled in S-Paramics using the ‘incident editor’, which allows for vehicles slowing down and/or stopping for the duration of the modelled incident (SIAS, 2007). The information input into S-Paramics to model an incident includes: the duration of the incident, the speed of the vehicles (zero for stopped), the delay experienced by affected vehicles (seconds), the lane(s) affected, and the incident rate (the percentage of vehicles using the specified lane that will experience the incident). If ‘feedback’ is turned on, vehicles are allowed to re-route to avoid additional delay caused by the incident.

ITS measures (e.g. VMS signs and transmitters, so that vehicles receive information about the incident when they are in a defined area) can also be modelled using S-Paramics. The incident information can include speed restrictions, lane restrictions, delay warnings, diversion routing and car-park availability advice. Driver aggression (this affects gap acceptance and overtaking behaviour), driver awareness (this affects the propensity to divert to an alternative route) and vehicle headways (for a specified area such as a ramp) can also

be modified during the incident. The ITS messages can be applied to all vehicles, or to specific vehicle types (e.g. to represent specially-equipped vehicle fleets).

In normal “on-street” applications, SCATS uses information from vehicle detectors, predetermined operation boundaries and historical data to determine cycle times, phase splits, phase sequences and coordination offsets. FUSE enables the actual on-street operation of SCATS-controlled intersections to be replicated in S-Paramics. It is relatively simple to code intersections in S-Paramics to be linked to SCATS through FUSE. Allowed and banned movements and vehicle detector locations are coded in S-Paramics and the signal controller settings and SCATS settings are loaded into WinTraff, which emulates the signal controllers. S-Paramics, WinTraff and SCATSim (software that replicates SCATS when connected to a traffic model) can all be run on the same computer.

Microsimulation models could be linked directly to a live traffic situation, to enable real-time assessment of incidents, modelling and performance evaluation of possible treatments (including taking no action) and selection of the best option. The key to this would be faster-than-real-time modelling speed (and possibly parallel systems testing different scenarios), so that evaluation can be done in a timely manner to give useful information.

For practical purposes, it was assumed that any incident to be modelled occurs within the model area and that its effects do not extend beyond the model area. This might be difficult to achieve unless an exceptionally large network is modelled, and a less detailed (mesoscopic) model such as SATURN could be used to identify the area that should be modelled, in order that virtually all the effects of an incident are captured. During this study, the traffic demand was consistent across all incident scenarios, and the simulation period was sufficient to ensure that all of the demand was loaded onto the network (i.e. that trips were not queued within zones at the end of the simulation run).

Microsimulation has variability in demand due to the stochastic nature of the simulation, and this means variability in travel time. Microsimulation linked to SCATS means even greater variability. Ten simulation runs were done for each incident scenario, using different “random seeds”, with the results being averaged.

There are several measures that might be used to assess the impact of an incident on network performance, including:

- the change in vehicle travel times;
- the amount of re-routing that takes place;
- the level of service (volume/capacity ratios) at key locations of the network;
- the time for the network to recover.

In NZ, the impact is assessed using the travel time reliability, as measured by the standard deviation in travel time (NZTA, 2010).

CASE STUDY

A calibrated model of a small region of Auckland’s North Shore was used for this case study

12th WCTR, July 11-15, 2010 – Lisbon, Portugal

(Koorey et al., 2008), and included a portion of Auckland's Northern Motorway as well as a parallel route using the urban arterial network along Wairau and Taharoto Roads (see Figure 1). The Auckland Harbour Bridge lies just to the south of the study area, with the CBD just south of the bridge. The Northern Motorway is part of State Highway 1 (SH 1), and is not only a key part of Auckland's traffic network, but also the main north/south route on NZ's North Island.

The network model included the Northern Motorway and the nearby arterial roads (see Figure 2). Three different scenarios were modelled:

- base scenario with no incident;
- an incident on the motorway with traffic diverting to the arterial route, with SCATS left to adapt as per normal (incident with original SCATS);
- an incident on the motorway with traffic diverting to the arterial route, with changes to SCATS to give priority along the diversion route (incident with modified SCATS).

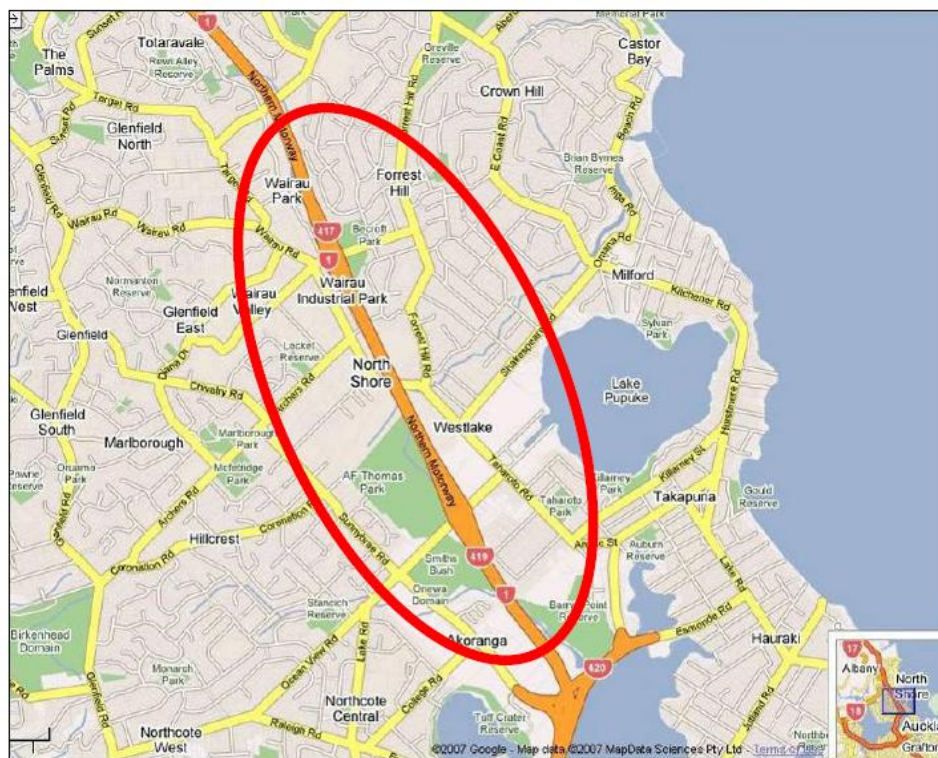


Figure 1: Case Study Area

It is likely that the timing of an incident will have a substantial effect upon the benefits of using SCATS to give priority along the diversion route. If an incident occurs late at night when traffic flows are relatively low, the benefit accruing from any incident management is likely to be very small. Conversely, for a network operating at or near capacity, it may be very difficult to achieve any meaningful improvement in network performance if an incident occurs. It was therefore decided to start with assessing the benefit of incident management using SCATS for an incident occurring during the “shoulder” of the peak period, when some spare capacity on the network still exists but network performance is sensitive to incidents.

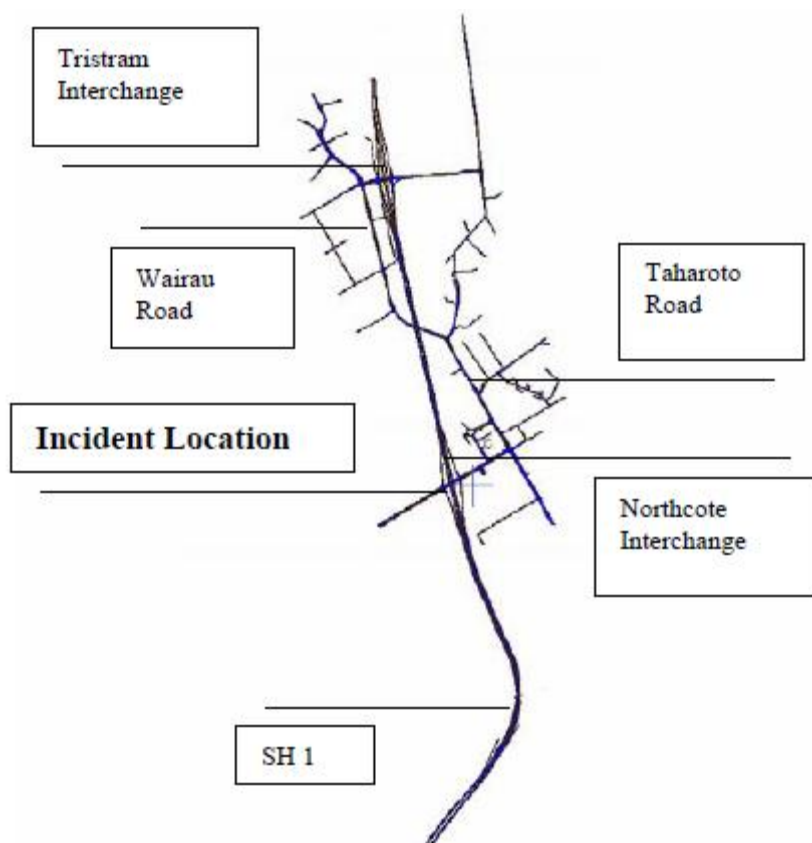


Figure 2. Network Model and Incident Location

One time period (from 3:15pm-4:30pm) has been modelled so far, with the incident being located on the northbound 3-lane section of the motorway (just north of the Northcote interchange) and involving complete closure of the kerb-side lane from 3:30pm to 4:00pm and complete closure of the centre lane from 3:30pm to 3:45pm. The incident can be considered a combination of a 'primary' incident and a 'secondary' (or consequential) incident. Such combinations are not unusual, with research indicating that approximately 14-18% of all incidents are secondary incidents (FHWA, 2009).

This incident scenario was chosen to ensure that there would be sufficient congestion on the motorway to:

- divert some motorists on the motorway away from the motorway onto the parallel arterial route;
- deter some motorists from entering the motorway (i.e. encourage them to remain on the parallel arterial route).

The pre-peak demand allowed for some spare capacity in the network during the base condition. This allowed the SCATS control on the parallel arterial route an opportunity to make changes as the flow pattern changed as a result of the incident. When the network is already fully congested, SCATS cannot adapt well to the change in flow pattern, as phase times and cycle lengths are probably already at their maximum values.

Two diversion routes were analysed. The first diversion route (the Taharoto diversion) is shown in Figure 3. This is available for drivers travelling northwards with destinations at (or beyond) the north end of the study area. These motorists have the option of taking the motorway (the black line) or continuing on the parallel SCATS-controlled arterial (the grey line) and joining the motorway at the Tristram interchange if travelling further north.

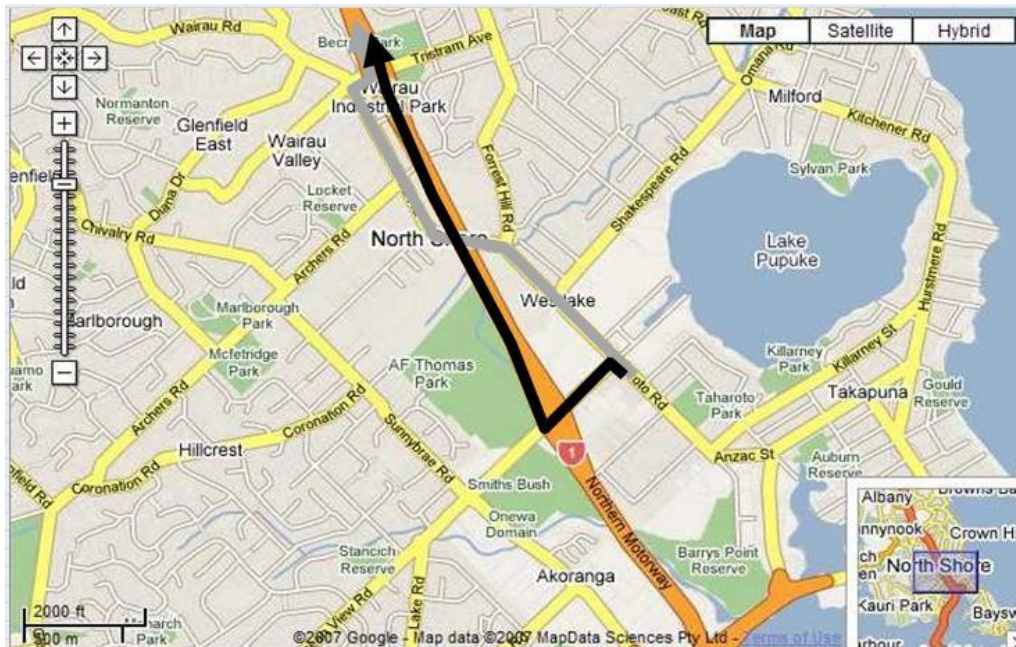


Figure 3. Taharoto diversion route.

The results of this diversion are shown in Figure 4 and demonstrate a small improvement in travel time on the diversion route when SCATS is modified for this incident condition.

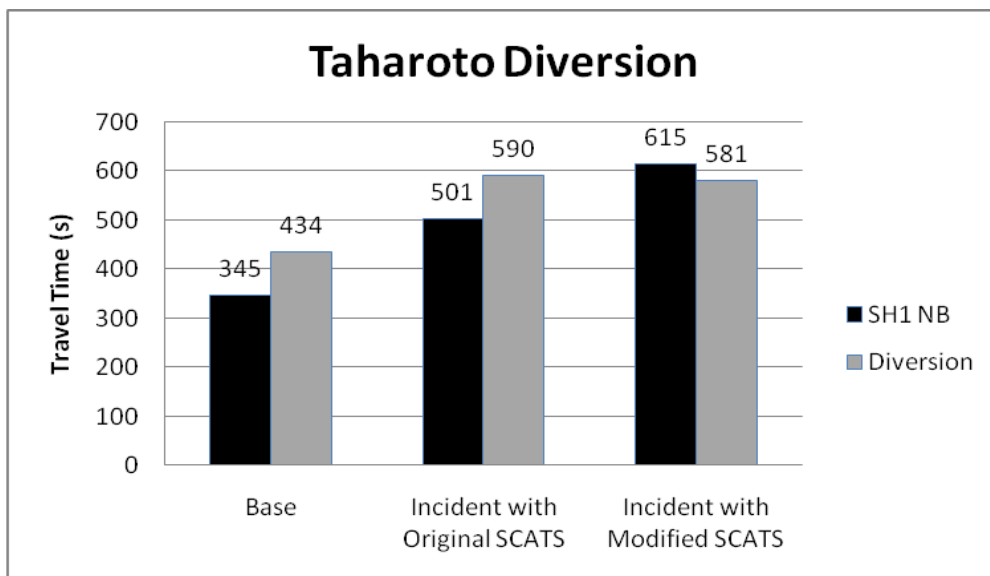


Figure 4. Taharoto diversion results.

The second diversion route (the Northcote diversion) is shown in Figure 5. This shows vehicles travelling northbound on the motorway (SH 1) with destinations at (or beyond) the north end of the study area. These motorists have the option of staying on the motorway (the black line) or diverting to the parallel SCATS-controlled arterial (the grey line), and re-joining the motorway at the Tristram interchange if travelling further north.

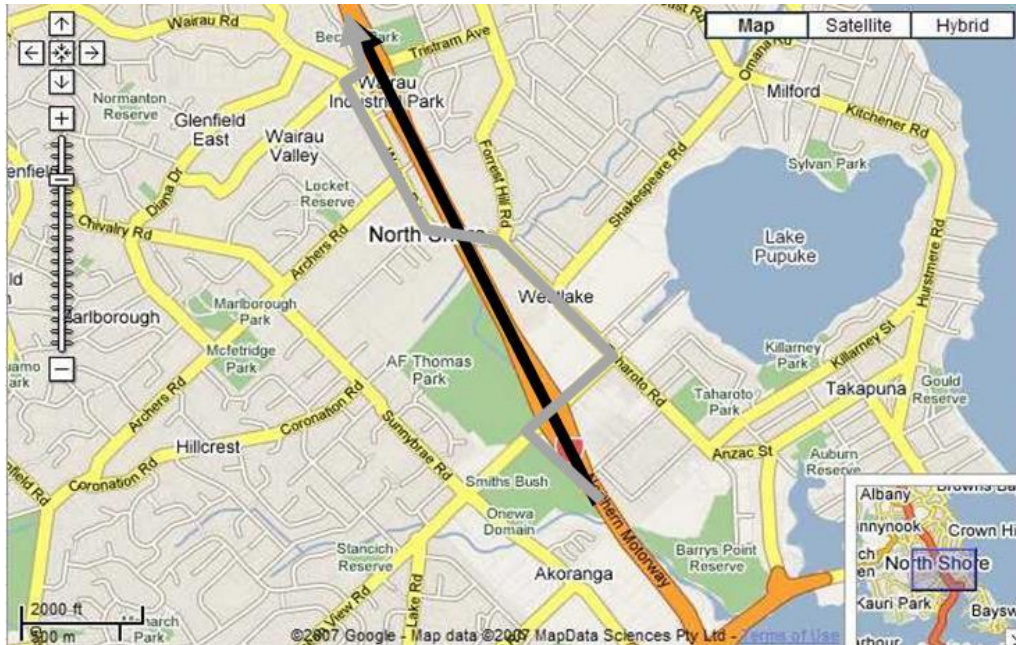


Figure 5. Northcote diversion route.

The results for this diversion are shown in Figure 6 and also show an improvement in travel time on the diversion route when SCATS is modified for this incident condition. The ‘base’ travel time is not given for this diversion, as it is generally not used when there is no incident.

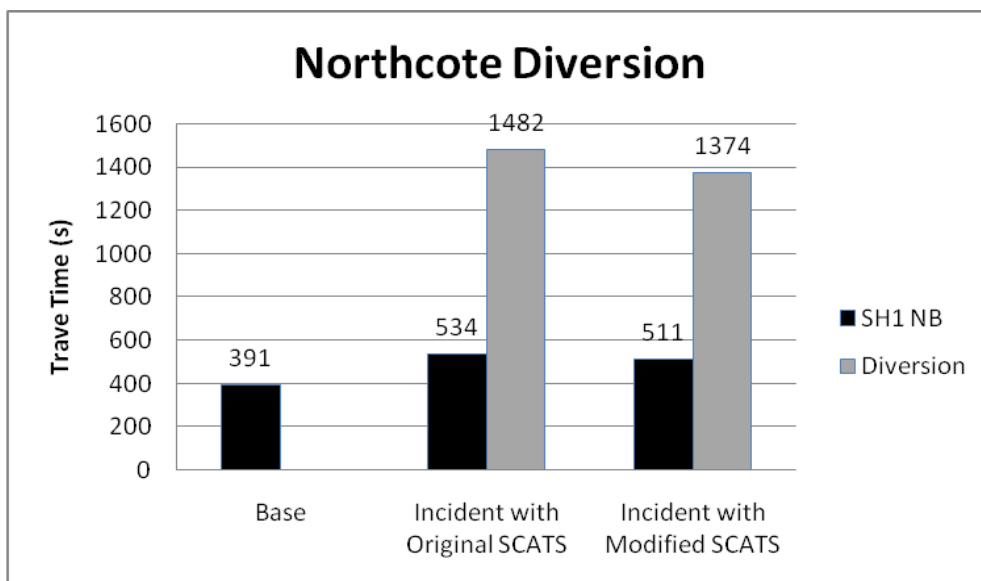


Figure 6. Northcote diversion results.

The modelling shows that an improvement in travel time traffic for diverted traffic due to incidents can be achieved with modifications to SCATS. The modifications made were minor and similar to the modifications a SCATS operator at the traffic control centre would manually make when such an incident is detected on the motorway. The changes were made only for the duration of the congestion related to the incident, with SCATS then reverting to the original settings. Although the improvement in travel times was small, if specific incident plans for particular incidents are developed, it is likely that the SCATS settings could be optimised and the travel time improved further.

CONCLUSION

The research has shown that it is possible to link SCATS signal control software with S-Paramics microsimulation software, so that it is practical to make an assessment of the ability of SCATS to adapt to traffic pattern changes following an incident. The research shows that there are significant benefits from using microsimulation as a 'test bed' for assessing signal control incident management strategies, without inconveniencing road users. In addition, it shows that an improvement in travel time for diverted traffic can be achieved with modifications to SCATS. The modifications made were minor and similar to the modifications a SCATS operator would make when an incident is detected on the motorway.

The results also highlight that an incident management plan is effective when:

- there are sufficient vehicles present to benefit from the plan (and hence justify the work required to implement it);
- there is sufficient capacity on alternative routes to enable them to be an attractive alternative to the original route.

The modelling also showed that the benefits of incident management plans may be limited to motorists on only some paths. For example, travellers already on the motorway generally did not benefit from diverting onto the local road network; their journey still took longer even with the incident in place. However, motorists on the arterial and local roads could benefit from not entering the motorway (or entering it beyond the location of the incident). The modelling provides a means of determining which motorists should be given prioritised corridors and related diversion information (e.g. VMS).

A literature review of techniques and software/systems to manage traffic congestion and respond to incidents found that:

- considerable research has been done in the areas of incident detection and management, ITS methods such as adaptive signal control (e.g. SCATS), and network reliability measures, but there has been little work done to bring all three research areas together.
- while automated incident detection techniques on traffic networks have become increasingly sophisticated, in practice there is still limited use (and trust) of them.

It is not surprising that the benefits of using SCATS as an incident management tool were not large, as SCATS was not developed as an incident management system. However, in

combination with microsimulation tools like S-Paramics, there may be considerable benefits in using this method to determine the most effective treatment option for managing incidents.

ACKNOWLEDGEMENT

The authors thank North Shore City Council for the use of their S-Paramics model, and the NSW Road Traffic Authority, SIAS and baseplus Ltd for academic licences for SCATS, S-Paramics and FUSE, respectively. They also thank the NZ Transport Agency for funding research that formed part of this study.

REFERENCES

- baseplus (2007). baseplus FUSE v2.0 User Manual. baseplus Ltd, Christchurch, NZ.
- Berdica, K., Andjic, Z. and Nicholson, A.J. (2001). Simulating Road Traffic Interruptions – Does it Matter What Model we Use?
- Eddington, R. (2006). The Eddington Transport Study: Main Report. HM Treasury, London, UK.
- FHWA (2006). Travel Time Reliability: Making It There On Time, All The Time. FHWA Publication FHWA-HOP-06-070, Washington, DC, USA.
- FHWA (2009) Best Practices in Traffic Incident Management. FHWA Publication FHWA-HOP-09-DRAFT. Washington, DC, USA.
- Goodwin, P.B. (1992). A Quality Margin in Transport. *Traffic Engineering and Control*, 33 (12), 661-665.
- Hidas, P. (2001). Simulation Modelling For Incident Management. Proceedings 24th Australasian Transport Research Forum, Hobart, Australia.
- Jong, G.C. de, Kouwenhoven, M., Kroes, E.P, Rietveld, P. and Warffemius, P. (2009). Preliminary Monetary Values for the Reliability of Travel Times in Freight Transport. *European Journal of Transport and Infrastructure Research*, 9(2), 83-99.
- Koorey, G.F., McMillan, S. and Nicholson, A.J. (2008). The Effectiveness of Incident Management on Network Reliability. Research Report 346, Land Transport NZ, Wellington, NZ.
- Nicholson, A.J. and Du, Z.P. (1997). Degradable Transportation Systems: An Integrated Equilibrium Model. *Transportation Research B*, 31(3), 209-223.
- Nicholson, A.J., Schmöcker, J.D., Bell, M.G.H. and Iida, Y. (2003). Assessing Transport Reliability: Malevolence and User Knowledge. *The Network Reliability of Transport* (Eds. Bell, M.G.H. and Iida, Y.), Elsevier Science, 1-22.
- Nicholson, A.J. (2007). Road Network Unreliability: Impact Assessment and Mitigation. *International Journal of Critical Infrastructures*, 3 (3/4), 346-375.
- NZTA (2009). Roads of National Significance. NZ Transport Agency, Wellington, NZ, <http://www.nzta.govt.nz/network/rons/>.
- NZTA (2010). Economic Evaluation Manual. NZ Transport Agency, Wellington, NZ.
- Parkhurst, G., Kenny, F. and Goodwin, P.B. (1992). Quality Attributes of a Journey: Their Identification and Impact on Travellers. Transport Studies Unit, Oxford University, Oxford, UK.

- SACTRA (1999). Transport and the Economy: Full Report. Standing Advisory Committee on Trunk Road Assessment, Department for Transport, London, UK.
- Samant A. and Adeli H. 2000. Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis. *Computer-Aided Civil Infrastructure Engineering*, 15 (4), 241-250.
- Schrank, D., and Lomax, T. (2009). 2009 Urban Mobility Report. Texas Transportation Institute, Texas A&M University, College Station, Texas, USA.
- Scottish Executive (2005). Congestion on Scottish Trunk Roads 2003. Scottish Executive, Edinburgh, UK.
- SIAS (2007). S-Paramics 2007 Reference Manual. SIAS Ltd, Edinburgh, UK.
- Srinivasan D., Jin X. and Cheu R.L. (2004). Evaluation of Adaptive Neural Network Models for Freeway Incident Detection. *IEEE Transactions on Intelligent Transportation Systems*, 5 (1), 1-11.
- Transfund NZ (2004). Project Evaluation Manual. Transfund NZ, Wellington, NZ.
- US DoT (2000). Incident Management Successful Practices - A Cross-Cutting Study. US Department of Transportation, Washington, DC, USA.
- Williams B. and Guin, A. (2007). Traffic Management Center Use of Incident Detection Algorithms: Findings of a Nationwide Survey. *IEEE Transactions on Intelligent Transportation Systems*, 8 (2), 351-358.
- Wolff, R.W. (1989). *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, New Jersey, U.S.A.
- Yaguang K. and Anke, X. (2006). Urban Traffic Incident Detection Based On Fuzzy Logic. Proc. 32nd IECON Annual Conference on Industrial Electronics, Paris, 772-775.
- Zhang, K. and Taylor, M.A.P. (2006). Effective arterial road incident detection: a Bayesian network based algorithm. *Transportation Research C*, 14 (6), 403-417.