

5th IMA Conference on Mathematics in Transport 12 - 15 April 2010

Comparison of signalized junction control strategies using individual vehicle position data

By **Simon Box, Ben Waterson, Nick Hounsell**

Transportation research group, University of Southampton, UK

Abstract

This paper is concerned with the development of control strategies for urban signalized junction that can make use of individual vehicle position data from localization probes on board the vehicles. Strategy development involves simulating the behaviour of vehicles as they negotiate junctions controlled by prototype strategies and evaluating performance.

Two strategies are discussed in this paper, a simple *auctioning agent* strategy and an extended auctioning agent strategy where a machine learning approach is used to enable agents to be trained by a human expert to improve performance.

The performance of these two strategies are compared with each other and with the MOVA algorithm in simulated tests. The results show that auctioning agents using individual vehicle position data can out perform MOVA, but that this performance can be improved further still by using learning auctioning agents trained by a human expert.

1. Introduction

Urban signalized junction control is a task that requires sensors to monitor the state of the network, a processing system to analyse the sensor data and make control decisions and traffic lights to implement the control. Sensors that are commonly used in signalized junction control are inductive loops (Sreedevi 2005), microwave emitter/detectors (Wood et al 2006) and traffic monitoring cameras. Examples of automated control algorithms that are in use to process data from these sensors and set signal timings are MOVA (Vincent & Peirce 1988) for isolated junctions and SCOOT (Hunt et al 1982), which can coordinate multiple adjacent junctions.

The sensors mentioned above all collect *census* data, that is counts of vehicles passing a specific point in space. The type of data that can be collected using on board vehicle localization sensors is *probe* data and this different type of data can present a fundamentally different view of the state of the network (Rose 2006). Probe data allow an analysis of the system that tracks each vehicle individually and can provide a higher resolution of position data. While several on board vehicle localization technologies exist, for example GPS (Trimble 2008), LIDAR (Levinson et al 2007), and computer vision (Wang et al 2007), a barrier to their implementation in signalized junction control is the requirement to communicate the localization data between the vehicles and infrastructure.

However a number of large European Commission funded projects (CVIS (Kompfner 2008), SafeSpot†, Coopers‡) have recently focussed on the development of technologies

† www.safespot-eu.org/

‡ <http://www.coopers-ip.eu/>

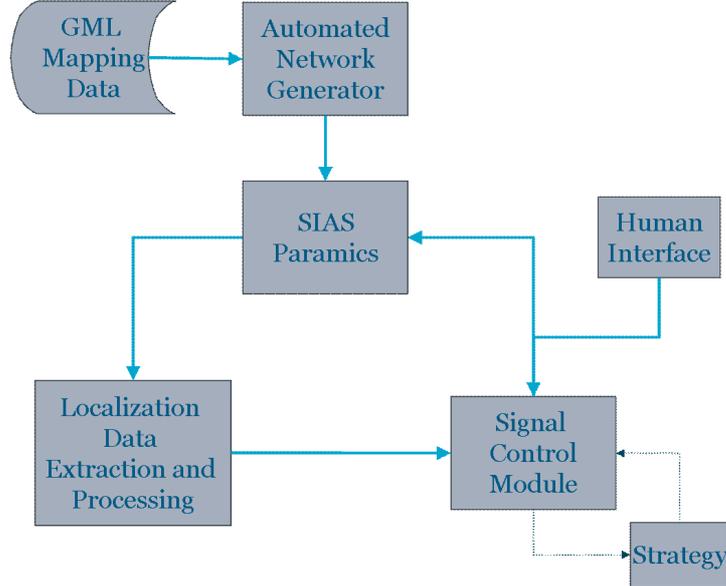


FIGURE 1. Block diagram showing the simulation software architecture

and standards for Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications systems. Furthermore common European protocols have now been established for these type of communications (IEEE 802.11 (WAVE) and IEEE 802.11p), making it possible for these technologies to become commonplace in vehicles in the near future.

This has led to interest in research that examines the use of V2I communications and localization systems in signalized junction control. For example the iBus project (Hounsell et al 2008) uses GPS based localization systems on London buses to give them priority at signalized junctions.

This paper is concerned with the design of signal control algorithms that could be used in the hypothetical scenario where some or all of the vehicles on the network are equipped with localization sensors and can transmit localization data to signalized junction controllers via V2I communications. In earlier work (Box & Waterson 2010a,b) the authors have presented a description of the simulation test bed used in the development of algorithms and some simple algorithms for control of both isolated junctions and pairs of connected junctions. This paper presents a brief summary of this earlier work before going on to present a new control strategy where a machine learning approach is taken to design a control algorithm that can learn control strategies from a human expert.

In this new approach a simple isolated T-junction is simulated and the setting of the traffic lights is done by a human expert who *plays* the simulation much like a computer game. During this process a multi-class logistic regression algorithm is used to identify patterns between the state of the system and the decisions made by the expert. The performance of the new *expert trained* algorithm is compared with the performance of auctioning agent algorithm developed previously (Box & Waterson 2010b) and with the MOVA algorithm (Vincent & Peirce 1988), which is commonly in use on isolated junctions today.

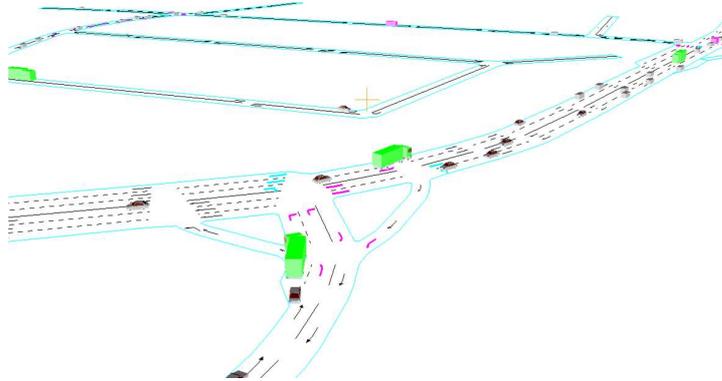


FIGURE 2. Example of the SIAS Paramics graphical output

2. Background Information

This section describes the simulation test bed that is used to develop signal control algorithms and the auctioning agent architecture that has been developed for urban signalized junction control using localization probe data.

2.1. Simulation Test Bed

Figure 1 shows the architecture of the simulation test bed. At the centre of this is a module for simulating vehicle movements and interactions through signalized junctions at the individual vehicle level (microscopic). The approach used in this research was to employ an existing commercially available microscopic traffic simulator (SIAS Paramics) to fill the roll of this module. The main advantage of this approach is expedience, allowing us to develop a test bed for control strategies relatively quickly.

A further advantage is that Paramics generates rich graphical visualizations (Figure 2), which can be employed for human interface control as will be discussed shortly. A disadvantage of this approach is that Paramics is a “black box” in our simulator where we are not aware of all the processes occurring between the input and output of data. The calibration setting for Paramics used in this research are the default setting in Paramics version 2007.1. Care must be taken with this approach that control strategies developed in the test bed are not too highly tuned to behaviour in Paramics, which may not be representative of the real world. It will be discussed later (Section 6) how this problem can be overcome in principle using a learning approach.

As can be seen in Figure 1, the simulation test bed has several additional modules that are built around Paramics.

The Network Generator module is used generate Paramics network models from Ordinance Survey mapping data.

The Localization Data Extraction and Processing module interrogates the Paramics simulation to obtain localization data for all the vehicles in the simulation. These data are perfectly accurate when extracted from the simulation and therefore not representative of data that would be obtained from real localization sensors. Therefore the *processing* stage of this module adds random noise to the data to simulate the performance of a real localization system (Box & Waterson 2010a).

The Signal Control module receives the processed localization data and uses them to inform decisions about signal control. The decisions are implemented directly in the Paramics simulation by setting the colour of the traffic lights. The Signal Control module

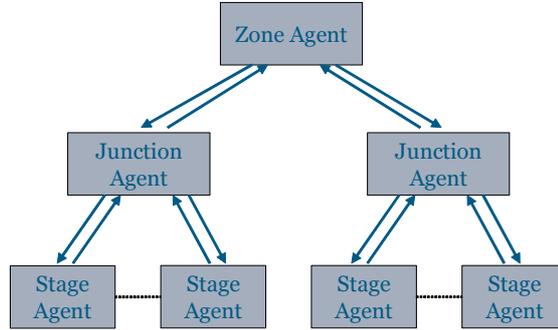


FIGURE 3. Structure of the agent hierarchy tree

is designed such that prototype control strategies can be easily ported in and out for testing. The module also has a human interface connection where a human controller can observe the graphical output from the Paramics module and, through keyboard input, directly set the colour of the lights in the simulation.

The Paramics module, the localization module and the signal control module are all synchronized to allow real time simulation signal control.

2.2. Auctioning agent algorithms

A signalized junction controller that uses localization probe data from all vehicles in the local area may have to process significant amounts of data in order to set signal timings. Previous research on signal control strategies where a large amount of loop data needs to be processed has demonstrated the advantage of a hierarchical agent structure (Choy et al 2003). Here individual software agents process small amounts of raw data, which they then pass on in a significantly refined form to another agent above them in the hierarchy. In this research we have adopted an agent hierarchy with the structure shown in Figure 3.

The lowest level agents in the hierarchy are called *stage agents*. There is one stage agent for each signalling stage of the junction. The stage agent receives data from vehicles whose position reveals that they are on a road, or in a lane, that will be given a green light during that stage. These data are refined by the stage agent into a simplified form, which constitutes a *bid* for priority. These bids are received by the *junction agent*, which will then assign priority to the stage with the winning bid. In a situation where a number of closely connected signalized junctions need to coordinate signal timings the junction agent will communicate with a *zone agent* above them in the hierarchy before assigning priority. The role of zone agents is examined more closely in Box & Waterson (2010b).

2.2.1. Bidding Algorithm

A bidding algorithm is used by the stage agent to refine the raw vehicle position data. The resultant bid should be descriptive of the state of the network on the approach to the signal stage and, in some sense, be a measure of the *need* for priority. One possible approach is to calculate the bid as a simple linear function of the number of vehicles being considered by the agent D , the speed of each vehicle V_i and the distance of each vehicle from the junction X_i .

$$B = \sum_{i \in D} 1 - \alpha V_i - \beta X_i \quad (2.1)$$

where α and β are coefficients that can be tuned, speed V_i is taken as the derivative of vehicle i 's position over the previous two time steps.

Having received bids from all stage agents the Junction agent simply needs to select the stage with the highest bid and assign priority. To avoid changing the stages too rapidly the junction agents performs these *auctions* only at a fixed time interval δt known as the auctioning rate.

Results from simulated tests of this approach presented in (Box & Waterson 2010a) showed that good values for the parameters were $\alpha = 0.01$ and $\beta = 0.001$ and a good auctioning rate was $\delta t = 10$ s. These are also the values used in the tests presented here. The simulated tests showed that the auctioning agent algorithm outperforms the MOVA algorithms, in terms of mean *delay*, mean *speed* and mean *queuing time*.

In Box & Waterson (2010b) it was shown that this auctioning approach can be extended to coordinated pairs of junctions by allowing the *zone agent* to examine selected bids from the stage agents and, if necessary, weight them to encourage coordination.

3. Expert trained auctioning agents

Here we present an extension to the auctioning agent approach that allows the junction agent to be trained by a human expert to make better decisions.

The simple approach described above where the junction agent picks the *highest bid* can be thought of as a simple division of the *bid space* into regions, one corresponding to each stage, the green light is given to a stage if the point defined by the bids from all stages falls into its region. This approach is valid as long as the bid is a true representation of the stage's *need* for priority, but with the naive bidding algorithm in (2.1) this is unlikely to be the case. One way to improve performance may be to design a more sophisticated bidding algorithm that closer represents the need for priority. However an alternative approach is to consider the bid less as a measure of need and more as a provider of information and allow the junction agent to move the regions in bid space around to optimize the junction's performance. Of course, the difference between these two approaches is only conceptual but thinking in terms of regions in bid space should make the following sections easier to understand.

The simulation test bed described in Section 2.1 can produce a rich animated graphical output of the simulation while it is running (Figure 2). This combined with the *human interface* module allows a human to control the junction as if they are playing a computer game. A human player controlling a simple isolated signalized t-junction can outperform the automated MOVA algorithm and the auctioning agent algorithm described in Section 2.2.

This motivates the idea to use an expert human signal controller to train a *learning* junction agent. The training involves getting the expert to play the signal control game, while they are playing the stage agents are simultaneously calculating bids using (2.1). Thus stage decisions made by the expert are associated with points in bid space. Over time the bid space becomes populated with a large number of data points. The technique of multi-class logistic regression can then be used fit suitable probability distributions for each of the stages to the data points and thus for any new point the probability that the expert would pick a given stage can be estimated. This allows the *trained* junction agent to define regions of bid space where each stage has the highest probability. The multi-class logistic regression technique used in this research is described in detail in the following section.

4. Multi-class Logistic Regression

Each time the expert makes a decision a pattern is recorded which links a point in bid space \mathbf{b} to a specific stage S_k . Taking each of the points in bid space which are associated with a given stage we define the stage-conditional probability over these points.

$$p(\mathbf{b}|S_k) = \mathcal{N}(\mathbf{b}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}) \quad (4.1)$$

which is a multivariate Gaussian with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. A prior probability for each stage $p(S_k)$ can be defined simply as the fraction of patterns connected with that stage. The posterior probability of the stage S_k given the pattern data is given by Bayes' theorem.

$$p(S_k|\mathbf{b}) = \frac{p(\mathbf{b}|S_k)p(S_k)}{\sum_{k \in K} p(\mathbf{b}|S_k)p(S_k)} \quad (4.2)$$

where K is the set of all k stages. From (Bishop 2006) (4.2) can be written as

$$P(S_k|\mathbf{b}) = \frac{e^{a_k}}{\sum_{k \in K} e^{a_k}} \quad (4.3)$$

where a_k is

$$a_k = \ln p(\mathbf{b}|S_k)p(S_k) \quad (4.4)$$

$$= \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_k \mathbf{b} - \frac{1}{2} \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_k + \ln p(S_k) \quad (4.5)$$

$\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}$ and $p(S_k)$ are parameters that can be learned from the pattern data \mathbf{b} . Alternatively, for simplicity, we can combine the terms in (4.5) to give

$$a_k = \mathbf{w}_k^T \mathbf{b} + w_{k0} \quad (4.6)$$

and simply learn the new parameters \mathbf{w}_k and w_{k0} from the data. Thus when given a new point in bid space it is possible to estimate the probability that the expert would choose stage S_k using (4.3).

4.1. Learning the parameters

In this section we show how to learn the parameters \mathbf{w} and w_0 from the expert training data. First we make a simplification to the notation by appending w_0 to the end of the vector \mathbf{w} and append 1 to the end on the vector b . Thus equation (4.6) becomes

$$a_k = \mathbf{w}_k^T \mathbf{b} \quad (4.7)$$

We now define a large vector of parameters \mathbf{W} which is made by appending the vectors \mathbf{w}_k for each of the K stages.

For each pattern \mathbf{b}_n in the set of N patterns we can define a target vector \mathbf{t}_n which has K elements t_{nk} , which equal 1 if \mathbf{b}_n is associated with S_k and 0 otherwise. All N vectors \mathbf{t}_n can be concatenated to form a large (NK) vector \mathbf{T} . We define the likelihood function.

$$p(\mathbf{T}|\mathbf{W}) = \prod_{n \in N} \prod_{k \in K} p(S_k|\mathbf{b}_n)^{t_{nk}} \quad (4.8)$$

The aim is to find the values of \mathbf{W} which maximize (4.8), this can be achieved equivalently by minimizing the negative logarithm of (4.8), giving us the following error function.

$$E(\mathbf{W}) = -\ln p(\mathbf{T}|\mathbf{W}) = -\sum_{n \in N} \sum_{k \in K} t_{nk} p(S_k | \mathbf{b}_n) \quad (4.9)$$

To calculate the derivatives of this error function we follow the method described by Li (2009). First we define a large KN vector \mathbf{Y} whose sub-vectors \mathbf{y}_k have N elements

$$y_{kn} = p(S_k | \mathbf{b}_n) = \frac{\exp(\mathbf{w}_k^T \mathbf{b}_n)}{\sum_{k \in K} \exp(\mathbf{w}_k^T \mathbf{b}_n)} \quad (4.10)$$

where we have made use of (4.3) and (4.7). Then we define a large matrix of scalars $\tilde{\mathbf{B}}$ that is a concatenation of $K \times K$ sub-blocks

$$\tilde{\mathbf{B}} = \begin{bmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B} \end{bmatrix} \quad (4.11)$$

where \mathbf{B} is a matrix with N rows of vectors \mathbf{b}_n^T and $\mathbf{0}$ is a matrix of zeros with the same dimensions as \mathbf{B} .

Then we define a large matrix of scalars $\tilde{\mathbf{Y}}$ that is a concatenation of $K \times K$ sub-blocks $\tilde{\mathbf{Y}}_{jk}$, which are $N \times N$ diagonal matrices where the n^{th} diagonal element is given by

$$\tilde{Y}_{jkn} = y_{jn} (I_{jk} - y_{kn}) \quad (4.12)$$

where I_{jk} equals 1 if $j = k$ and 0 otherwise.

Using the matrices defined above we now define the gradient of the error function (4.9) as

$$\nabla E(\mathbf{W}) = \tilde{\mathbf{B}}^T (\mathbf{T} - \mathbf{Y}) \quad (4.13)$$

and the Hessian matrix is given by

$$\nabla \nabla E(\mathbf{W}) = \tilde{\mathbf{B}}^T \tilde{\mathbf{Y}} \tilde{\mathbf{B}} \quad (4.14)$$

4.2. Iteratively Re-weighted Least Squares IRLS

The values of \mathbf{W} that minimize $E(\mathbf{W})$ are found iteratively using the Newton-Raphson update formula.

$$\mathbf{W}^{\text{new}} = \mathbf{W}^{\text{old}} - \nabla \nabla E(\mathbf{W}^{\text{old}})^{-1} \nabla E(\mathbf{W}^{\text{old}}) \quad (4.15)$$

This approach is known as iteratively re-weighted least squares. \mathbf{W} is initialized randomly using the approach recommended by (Nabney 2002) shown below.

$$w \sim \mathcal{N}\left(0 \left| \frac{1}{\sqrt{N+1}} \right.\right) \quad (4.16)$$

Typically the IRLS algorithm is run a number of times to avoid a result in a poor local minimum.

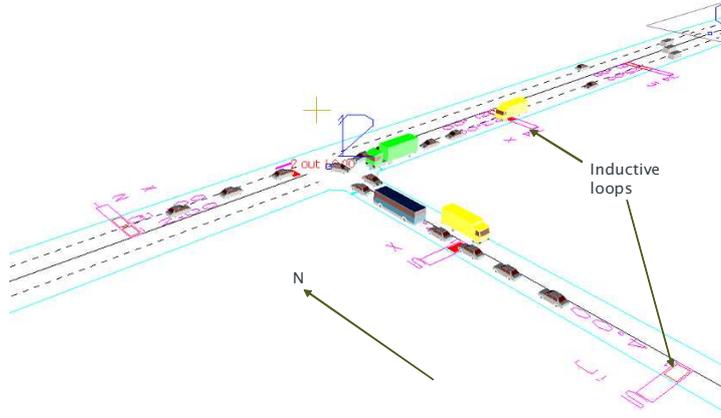


FIGURE 4. Paramics T-junction model used in the tests

	West	East	South
West	–	12.5	3.3
East	15.8	–	0.8
South	2.7	2.7	–

TABLE 1. Matrix of demands (Vehicles per minute) across the junction shown in Figure 4

5. Tests

Simulated tests were carried out on the T-junction shown in Figure 4. This junction has three signal stages: stage 1 gives priority to vehicle on the Eastern and Western arms of the junction, stage 2 is a right turn filter giving priority to right turning vehicles on the Western arm and stage 3 gives priority to the Southern arm.

Four tests were carried out on the junction using four different control methods, each test lasted for a simulated time of 1 hour during which time statistics for delay were recorded. Delay is defined here as the time between a vehicle entering and leaving the region of the junction shown in Figure 4. The vehicle demand levels were constant in all tests, the demand matrix is shown in Table 1. The four control methods used in the tests are summarised below.

- The first control method tested was MOVA, here simulated loop detectors (shown in Figure 4) were used to provide data for the MOVA algorithm.
- The second control method used was the *highest bid* algorithm described in Section 2.2.1, here simulated probe detectors with an accuracy of $\sigma = 2$ m were used in all vehicles.
- The third control method was human control, in this case the expert human was Dr Nick Hounsell who has many years experience in traffic analysis and highway design.
- The fourth and final method used was auctioning agents with a *trained* junction agent that has learned from the expert as described in Section 3. Again this test used simulated probe detector data with accuracy of $\sigma = 2$ m.

5.1. Results

Figure 5 shows the average delay for each of the control methods tested, measured across all vehicles for the duration of the test. The *highest bid* method can outperform MOVA

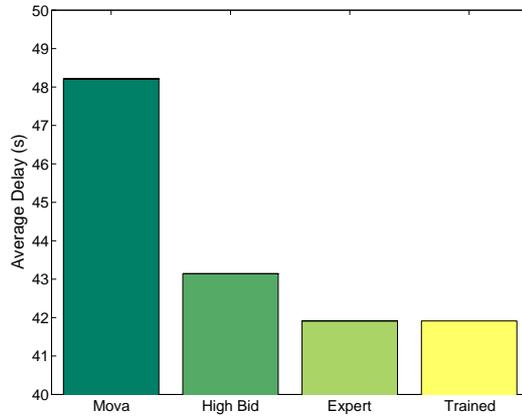


FIGURE 5. Average delay recorded for each of the four control methods

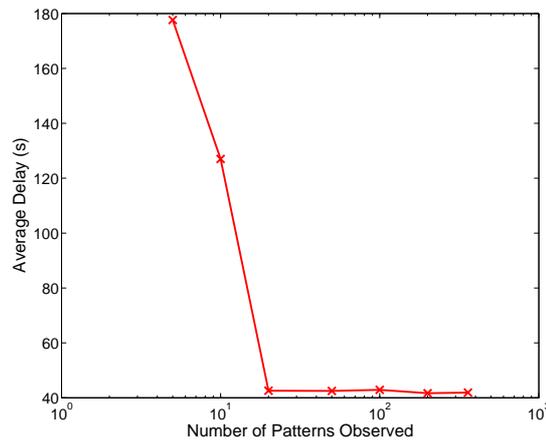


FIGURE 6. Average delay for tests using the trained junction agent after it has been exposed to various amounts of the expert data.

despite its simplicity due to the fact that it has access to a much richer source of information from the localization probe data. But it cannot perform better than the human expert who beats the highest bid method by about 1.2s. However after learning from the expert data the trained junction agent can equal the performance of the expert in terms of delay as shown in the final bar in Figure 5.

It is also informative to study how the performance of the trained junction agent varies when it has been trained on different amounts of data. In the test whose results are presented in Figure 5 all the expert data was used which consists of 354 patterns. Figure 6 shows the average delay measured in seven tests of the junction agent where it has been trained on different amounts of data ranging from just 5 patterns up to 354. This shows that a good performance is achieved after just 20 patterns have been observed and after this point delay continues (in general) to reduce very slightly with some variation.

6. Discussion

The results in Section 5.1 provide encouraging evidence that localization probe data can provide useful information that allows signal control strategies based on these data to outperform traditional strategies based on loop data.

It has also been shown that using learning algorithms to emulate a human expert traffic controller can lead to improvements in performance. It is a slightly surprising result that the trained junction agent performs equally as well as the expert, because the employed technique of logistic regression captures only broad trends in the expert's behaviour. This suggests that the performance of the strategy learned from the expert is relatively insensitive to small changes in the strategy.

The simulated junction on which these tests were performed was relatively un-complex. Complexity was further limited by restricting the tests to a constant level of demand. This lack of complexity is reflected in the small amounts of data required to train the junction agent (Figure 6). The advantages of the learning approach presented here is that, through finding patterns in the human generated data, it may be able to capture something that a human does when solving this problem that is difficult to encode in an algorithm. It is encouraging that this learning approach was able to yield improvement on this simple problem but it is anticipated that the advantages of this approach will become more pronounced when testing problems of greater complexity, for example coordinating multiple junctions, junctions with high demand and dealing with rare events, such a breakdown on the junction. Exploring these scenarios with the learning approach is a goal for future work.

In learning from human data the junction agent makes the implicit assumption that all the decisions that the human makes are good ones. Because the junction controller is receiving localization probe data from all vehicles it is possible for it to track the positions of the vehicles after a decision has been made and evaluate if the decision was good or bad. This raises the possibility of extending the learning approach to use in-situ learning (*reinforcement learning*) following an approach similar to that discussed in Tesauro (1992). In this scenario: in situations where the probability of two stages are close the controller could choose (on some occasions) to pick the stage with the lower probability and if this is evaluated as good decision a new point can be added to the training data. Thus the junction agent can continue to learn and improve while in operation. This also is a goal for future work.

In Section 2.1 we briefly discussed the disadvantages of using a simulator to design signal control algorithms. A further advantage of the learning algorithm approach is that, while agents can be trained on simulated data, as we have done here, they are equally happy to be trained on data from the real world and can be tuned using in-situ learning as described above. Therefore using this approach can avoid the trap of designing a control strategy which is tuned to the simulator but does not perform well in the real world.

7. Conclusions

A simulation test bed that can model the generation and processing of localization probe data from all vehicle on the simulated network has been used to develop urban signalized junction control strategies that employ these data. An auctioning agent hierarchy has been presented for these strategies and simulated tests show that using a simple highest bid approach, delay across an isolated signalized T-junction can be reduced by an average of 5 seconds per vehicle over MOVA control.

The auctioning agent approach has been extended to allow the junction agent to be trained by human expert using the technique of multi-class logistic regression. Simulated results show that with this approach delay across the T-junction can be reduced by an average of 6.3 seconds per vehicle over MOVA control.

REFERENCES

- BISHOP, C. M. 2006 Pattern Recognition and Machine Learning *Springer*.
- BOX, S. & WATERSON, B 2010a Signal control using vehicle localization probe data. *42nd Annual UTSG Conference, University of Plymouth, 5-7 January 2010*.
- BOX, S. & WATERSON, B 2010b Comparison of signalized junction control strategies that use localization probe data. *The IET Road Transport Information and Control Conference, 25 - 27 May 2010, Dexter House, London, UK*.
- CHOY, M. C., SRINIVASAN, D. & CHEU, R. 2003 Cooperative, hybrid agent architecture for real-time traffic signal control. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, **33**(5), 597–607.
- HOUNSELL, N., SHRESTHA, B., HEAD, J., PALMER, S. & BOWEN, T. 2008 The way ahead for london's bus priority at traffic signals. *IET Intelligent Transport Systems*, **2**(3), 193–200.
- HUNT, P., BRETHERTON, R., ROBERTSON, D. & ROYAL, M. 1982 Scoot on-line traffic signal optimisation technique. *Traffic Engineering and Control*, **23**, 190–192.
- KOMPNER P. 2008 Cvis cooperative for mobility. *Online at http://www.cvisproject.org/download/cvis_brochure_May2008_Final.pdf*.
- LEVINSON, J., MONTEMERLO, M. & THRUN, S. 2007 Map-based precision vehicle localization in urban environments. *Proceedings of Robotics: Science and Systems, Atlanta, GA, USA*.
- LI, J. 2009 Logistic regression. *online at <http://www.stat.psu.edu/~jiali/course/stat597e/notes2/logit.pdf>*.
- NABNEY, I. T. 2002 Netlab Algorithms for Pattern Recognition. *Springer*.
- ROSE, G. 2006 Mobile phones as traffic probes: Practices, prospects and issues. *Transport Reviews*, **26**(3), 275–291.
- SREEDEVI, I. 2005 ITS decision services and technologies loop detectors. *online at: http://www.calccit.org/itsdecision/serv.and.tech/Traffic_Surveillance/road-based/in-road/loop_summary.html*.
- TESAURO, G. 2006 Practical issues in temporal difference learning. *Machine Learning*, **8**(3-4), 257–277.
- TRIMBLE 2008 MTrimble navigation ltd. how gps works: A tutorial. *online at <http://www.trimble.com/gps/>*.
- VINCENT, G. & PEIRCE, J. 1988 MOVA: Traffic responsive, self-optimising signal control for isolated intersections. *TRRL Research Report*, **RR170**.
- WANG, C., HU, Z., KUSUHARA, S. & UCHIMURA, K. 2007 Vehicle localization with global probability density function for road navigation. *Intelligent Vehicles Symposium, 2007 IEEE*, 1033–1038.
- WOOD, K., CRABTREE, M. & GUTTERIDGE, S. 2006 Pedestrian and vehicular detectors for traffic management and control. *TRL Report*.