Self-organising traffic control inspired by inventory theory and the process of osmosis

A.P. Burger\textsuperscript{a}, M.D. Einhorn\textsuperscript{b,\ast}, J.H. van Vuuren\textsuperscript{c}

\textsuperscript{a}Department of Logistics, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, email: apburger@sun.ac.za

\textsuperscript{b}Department of Logistics, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, email: einhorn@sun.ac.za, cell phone: (0027) 82 854 2251

\textsuperscript{c}Department of Industrial Engineering, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, email: vuuren@sun.ac.za

Abstract

Decentralised, self-organising traffic signal control is an attractive alternative to present day centralised traffic control techniques and has been shown to be superior in terms of reducing vehicle delay time in a road network. Recent self-organising traffic signal control algorithms, however, incorporate numerous user-defined parameters, which, if not selected appropriately, can render the algorithms ineffective. Three novel, robust self-organising traffic signal control algorithms are presented in this paper which are free of any predeter- mined parameters. Instead, all three assume the use of novel radar detection technology to provide all the data necessary to inform signal switching policies. The first is inspired by inventory theory, the second by the process of osmosis and the third is a hybrid of the two. The algorithms are all tested and compared to previously proposed self-organising traffic signal control algorithms from the literature in a purpose-built microscopic traffic simulation framework under various traffic demand scenarios. The various algorithmic performances are measured in terms of their ability to reduce total vehicle delay time and facilitate natural coordination among adjacent intersections. The hybrid algorithm is found to be superior and is able to best implement flexible traffic signal control without sacrificing natural coordination between adjacent intersections.

Keywords: Traffic, Transportation, Simulation, Self-organisation

\ast Corresponding author
1. Introduction

Signalised traffic control and its attempted optimisation within a traffic network have been the focus of many a study across several different scientific disciplines, including engineering, operations research, physics and statistics. Today, two distinct types of predominant traffic signal control exist: fixed-time control and vehicle actuated control. Fixed-time control was the earlier of the two approaches. It involves the optimisation of several traffic signal cycle parameters, such as the duration of the cycle itself, the duration of the various green times which comprise the cycle, and the offset of green times of adjacent intersections, in an attempt to facilitate coordination in a traffic network [1, 2]. These parameters are typically optimised off-line for assumed average traffic flows, such as morning and afternoon rush hours [3, 4]. A disadvantage of this approach, however, is that traffic signal timings are set for assumed mean traffic demands which are rarely actually met and as a result are typically too rigid to respond to sudden fluctuations in vehicle demand away from an assumed mean. The traffic signal timings are therefore often too long or too short, resulting in an inefficient utilisation of intersection capacity and thus avoidable vehicle delays.

Vehicle actuated control, on the other hand, seeks to adapt to variations in the average traffic demand over a given time horizon by employing some form of vehicle detection mechanism\(^1\) to provide input to the traffic signal control algorithm [3]. These data are then used to determine when to switch between signal phases. Two prominent examples of such control techniques are the Split Cycle Offset Optimisation Technique (SCOOT) [7] and the Sydney Coordinated Adaptive Traffic System (SCATS) [8]. While these vehicle actuated control techniques are able to perform on-line or real-time optimisation operations, they remain largely centralised, attempting to determine optimal cycle lengths, green time splits and cycle offsets of adjacent intersections based on prevailing traffic conditions as interpreted by upstream vehicle detectors. A disadvantage to this method of traffic signal control is that the problem of optimal control of switched network flows is known to be NP-hard [9].

\(^1\)Currently the most common form of detection used is an inductive loop detector installed beneath the road surface [5, 6].
Self-organising, decentralised traffic signal control has been proposed as an attractive alternative to overcoming the disadvantages mentioned above. In self-organising traffic signal control, traffic signal timings of individual intersections are governed by a predetermined set of algorithmic rules. There is no communication between adjacent intersections and no explicit attempt is made to achieve coordination among intersections. Instead, a natural coordination emerges as a result of effective vehicle detection and traffic signal switching operations. Self-organising traffic signal control strategies have been shown to outperform both optimised fixed time control strategies [3, 4, 10, 11, 12] and state-of-the-art centralised traffic responsive systems [3] in terms of minimising vehicle delay.

The self-organising traffic signal control approach by Lämmer and Helbing [3, 4] employs an optimisation strategy which seeks to serve alternate intersection approaches as quickly as possible based on approach priority values, as well as a stabilisation strategy which ensures that queues along intersection approaches do not grow exceedingly long before receiving service. The self-organising traffic signal control approach of Gershenson [11], Gershenson and Rosenblueth [10] and Zubillaga et al. [12] also serves intersection approaches in a priority-based manner, with platoons of vehicles receiving a higher priority in an attempt to facilitate the formation and propagation of green-waves.

While both self-organising approaches described above can be very effective at minimising vehicle delay under certain prevailing traffic conditions, they rely on several user-defined parameters to ensure their effective operation. Poorly selected parameter values can render the self-organising traffic control algorithms ineffectual, as was shown in [13]. Both self-organising traffic control approaches assume the use of some vehicle detection mechanism, but they do not categorically specify what form of detection is used, nor do they state the capabilities or short-comings of the detection equipment. Finally, both self-organising traffic control approaches are based on a number of simplifying assumptions such as that no vehicle acceleration or deceleration takes place (they either move at a constant speed or are stationary), that all vehicles travel at the same speed and that vehicles are assumed to be of uniform size. These assumptions are made due to the fact that the self-organising traffic control algorithms rely more on the presence of vehicles along an intersection approach rather than on their individual characteristics, such as speed, size and distance from the intersection.

In this paper, three novel self-organising traffic signal control algorithms
are put forward. All three algorithms assume the use of recently developed radar detection equipment [14]. When mounted at an intersection, this equipment is capable of detecting and tracking individual vehicles, measuring their speeds, lengths and distances from the intersection, and of estimating the times before the vehicles reach the intersection. The radar has a detection range of more than 250 metres. The algorithms proposed in this paper are free of any user-defined parameters and instead rely on the individual vehicle characteristics and space available along road sections to determine the phase switching policies, rather than functioning based on the mere presence of vehicles along an intersection approach. The algorithms are implemented and tested in a custom-built microscopic traffic simulation framework which allows for the incorporation of variable vehicle speeds, vehicle acceleration and vehicles of variable size [15]. This allows for the full potential of the radar detection equipment to be realised by the novel self-organising traffic control algorithms which are sensitive to changes in vehicle speed, acceleration and deceleration, as well as lane changes. The existing two self-organising traffic control algorithms referred to above are also implemented and investigated in the same traffic simulation framework, allowing for a direct comparison of the various new and existing self-organising traffic control algorithms.

2. Prevailing traffic signal control paradigms in the literature

Traditional traffic signal control (i.e. both the classical fixed-time control and the more recently developed vehicle actuated control) as well as self-organising traffic signal control strategies are described in this section. The mechanisms and logic behind each strategy are considered, as well as the associated advantages and limitations of these strategies.

2.1. Traditional traffic signal control

In a classical fixed-time traffic signal control paradigm, constant predetermined parameters are implemented in a constantly repeating cyclical manner. For an isolated intersection, these parameters include the cycle time and the duration of the various green times within this cycle [1, 2]. Coordination among adjacent signalised intersections may be achieved by implementing the same cycle time at each intersection and adjusting the offset between them for displaying a green signal [16]. This is usually implemented so as to facilitate the formation and propagation of green waves (i.e. a platoon of vehicles which encounters a green signal at several adjacent intersections
without having to stop). The problem of optimising these traffic signal control parameters for a given traffic network is NP-hard [9].

Advances have been made in respect of the classical cycle-based fixed-time traffic signal control strategy. These more advanced concepts attempt to adjust the traffic signal control parameters in response to slow or systematic variations of the average traffic demand [3]. Two such examples are the commercially available SCOOT [5,7] and SCATS [6,8]. Both SCOOT and SCATS are a form of centralised traffic signal control which are intended for use in conjunction with a broader traffic control system which typically involves some form of human supervision. While more effective at reducing delay time than classical fixed-time control, these two examples of attempted online optimisation of traffic signal timings are essentially only variations on the theme of a mainly cyclic mechanism of control. Alternative heterogeneous phase plans, which vary the order and the frequency of service of the different intersection approaches, are not considered [3].

In an attempt to improve on the responsiveness of traffic control strategies to fluctuations in prevailing traffic conditions, various alternative approaches have been introduced in the literature for accommodating acyclic traffic control and permitting the alternate orderings of green phases. Examples of these strategies include PRODYN [17], Optimised Policies for Adaptive Control (OPAC), Urban Traffic Optimisation by Integrated Automation (UTOPIA) [18], and Adaptive Limited Lookahead Optimisation of Network Signals — Decentralised Version (ALLONS-D) [19]. A limitation shared by all of these strategies is that they rely on vast amounts of data collection and processing, and require considerable processing power [20]. Furthermore, for the case of global coordination in a traffic network, the data required (such as dynamic origin-destination matrices of vehicles as well as turning probabilities) are difficult to quantify and obtain. [20].

2.2. Self-organised traffic signal control

An attractive alternative to solving the complex combinatorial optimisation problem of signalised traffic intersection control by enumerating large solution spaces or traversing complex decision trees and involving possible switching sequences has been proposed in the form of decentralised self-organising traffic control. A self-organising system functions through contextual local interactions without any central control. The constituent components each achieve a simple task individually, but a complex collective behaviour emerges from their mutual interactions. Such a system modifies
its structure and functionality so as to adapt to changes to requirements and to the environment based on previous experience [21]. The self-organising traffic control algorithm proposed by Gershenson [10], as well as that proposed by Lämmer and Helbing [4], are briefly reviewed here.

Gershenson claims that improving vehicle flow at signalised intersections and obtaining coordination among adjacent intersections is more a problem of adaptation rather than optimisation [11]. The self-organising traffic control algorithm proposed by Gershenson and Rosenblueth [10] comprises six rules which are implemented independently at each signalised intersection to regulate traffic, with no direct communication among intersections. The model assumes that all vehicles are of the same size and all travel at an equal, uniform speed. Time is assumed to be discrete, and all cells are updated synchronously. A further assumption is that traffic may travel in one direction along a roadway only and that no turning is allowed at intersections.

The self-organising traffic control algorithm proposed by Lämmer and Helbing [3, 4] differs from that proposed by Gershenson and Rosenblueth [10] in that it comprises both an optimisation strategy and a stabilisation strategy. Traffic flow through the road network is described by a fluid dynamic model which considers vehicle flow rates rather than the characteristics of individual vehicles. The traffic dynamics of an intersection approach are characterised by the arrival rate and departure rate of vehicles to and from the intersection. Again it is assumed that all vehicles travel at a fixed, uniform speed. Traffic models which do not explicitly incorporate vehicle accelerations and decelerations adjust the setup time (usually 3 to 8 seconds in duration and typically comprising an amber and all-red period determined according to safety requirements [4]) in an attempt to implicitly account for delays caused by driver reactions and finite acceleration times, as is the case with the model of Lämmer and Helbing [4]. This practice has been shown to lead to discrepancies in the accuracy of the actual delay times experienced by vehicles [22].

The optimisation strategy relies on the effective anticipation of traffic flows and platoons in order to predict the effects on future delay times of vehicles in terms of starting, continuing, or terminating service processes [23]. It assigns each approach a priority index which may be interpreted as the average service rate of the approach, i.e. the anticipated number of vehicles expected to be served during the resulting setup period and green phase. The approach achieving the largest priority index is awarded service.

Lämmer and Helbing [4] couple the aforementioned optimisation strategy
with a stabilization strategy so as to ensure stability\(^2\) along all intersection approach lanes. The stabilisation strategy functions according to two key requirements. The first requirement is that each traffic flow should be served once on average within a desired service interval \(Z\). Here, \(Z\) is the cycle time of an associated stable, fixed-time control programme. This requirement ensures a desired regularity of service to each traffic flow. The second requirement is that each traffic flow should be served at least once within a service interval \(Z_{\text{max}}\).

The optimisation and stabilisation strategies described above are combined as follows: The optimisation strategy is responsible for assigning service to the various phases as long as none of the approaches served during any of the phases is considered to be critical\(^3\). As soon as an approach is identified as being critical, the phase during which the approach receives service is moved forward as soon as possible. An approach remains in a critical state until either the vehicle queue along it has been cleared completely or until it has received a green time equal to that which a stable, fixed-time control programme would have awarded it. If it so happens that more than one flow is considered to be critical, then the flows are served in the order in which they became critical.

The two self-organising traffic control strategies described above have been shown to outperform both optimised classical fixed-time control programmes [3, 4, 10, 11, 12] and state-of-the-art centralised traffic responsive systems [3] in terms of minimising vehicle delay. They are, however, not without limitations. Both algorithms rely sensitively on the appropriate selection of several parameter values in order for them to function effectively. The algorithm of Gershenson and Rosenblueth [10] contains seven parameters and while the optimisation strategy of Lämmer and Helbing [4] is free of parameters, their stabilisation strategy relies on the appropriate selection of two parameters, \(Z\) and \(Z_{\text{max}}\), which are chosen according to a stable, fixed-time control strategy based on average vehicle arrival and departure rates. While Gershenson and Rosenblueth [10] maintain that their algorithm is “robust” in that it is “not affected by modest parameter variations,” they do not provide

\(^2\)A controlled queueing network is deemed to be stable if all queue lengths remain bounded at all times [24].

\(^3\)An approach is considered to be critical as soon as a critical number of vehicles can be served. This critical number of vehicles is determined according to a pre-defined threshold function.
a procedure for selecting appropriate parameter values. We consider this a vital drawback as it has been shown that appropriate parameter selection is integral to the effectiveness of self-organising traffic control algorithms [13]. Lämmer and Helbing [4], on the other hand, provide a method for parameter selection for their stabilisation strategy, but their method relies heavily on an assumed stable, fixed-time control programme designed for constant average traffic flows which they do not elaborate upon. A potential drawback of this approach is that the stabilisation strategy may therefore be subject to the same limitations as the fixed-time control approach on which it is based in that it may not be flexible enough to adapt adequately to sudden changes in traffic demand.

The self-organising traffic control algorithms presented in the following two sections provide an alternative to those described above in that they are free of any user-defined parameters, relying instead only on individual vehicle characteristics, such as speed, length and distance from the intersection as input data, as well as on the physical dimensions of the road network itself. For this reason, characteristics such as vehicle acceleration and lane selection are incorporated into the algorithms, rather than omitting them for simplification purposes as is usually the case.

3. A self-organising control algorithm inspired by inventory theory

The well-known economic order quantity (EOQ) in inventory theory seeks to determine two key decision variables: determining inventory reorder points in time and determining the associated reorder quantities that minimise the total costs incurred in maintaining inventory levels and failing to meet customer demand. Parallels may be drawn between this inventory control problem and the problem of allocating green time to the various phases of a traffic signal cycle. Green time may be considered as the product in inventory, while the phases of the signal cycle and their associated approach lanes may be interpreted as customers seeking to replenish their green time stock. The vehicles served during the specific phases may be viewed as the cause of demand for inventory stock. In this analogy, the two key inventory theory variables mentioned above translate to the following questions: When should a particular phase receive green time? How much green time should this phase receive?
3.1. The costs involved in basic inventory control models

In inventory theory, all costs are typically expressed in monetary terms, while in the context of traffic control at signalised intersections, it is more natural to express costs in terms of vehicle delay time. The following four costs are usually associated with basic inventory models [25]: the ordering and setup cost, the unit purchasing cost, the holding or carrying cost, and the stockout or shortage cost.

The ordering and setup cost is incurred when placing an order for inventory stock, or setting up a production run if the stock is produced internally. Costs of this nature usually do not depend on the size of the order or length of the production run; they are typically fixed costs incurred each time stock is ordered or a production run is initiated. In terms of traffic signal control, the ordering and setup cost may be interpreted as the delay time incurred by vehicles waiting at an intersection during the necessary amber and all-red phases preceding the onset of a green signal.

The holding or carrying cost measures the expenses associated with holding one unit of stock in inventory for a single time period. This cost normally incorporates storage cost, insurance cost and inventory taxes. The most significant component of holding cost, however, is the opportunity cost associated with tying up capital in inventory. In the context of traffic signal control, a certain phase may be considered to be holding inventory while it is receiving service. This cost may be interpreted as the delay time incurred by all vehicles at an intersection which are required to wait while the particular phase receives service (i.e. all vehicles along intersection approaches not currently receiving service).

The stockout or shortage cost is the result of demand not being met on time and is comparatively harder to quantify in inventory theory. It includes the cost of lost sales as well as the cost of loss of customer goodwill. In the context of traffic signal control, stockout cost may be interpreted as the vehicle delay time resulting from the termination of a green signal (i.e. the delay time experienced by all vehicles along an intersection approach to which a green signal has just been terminated while they wait for service to resume).

3.2. The inventory traffic signal control algorithm

The newly proposed inventory traffic signal control algorithm (I-TSCA) is a self-organising, adaptive traffic signal control algorithm which seeks to minimise the total delay time experienced by vehicles passing through an
intersection. It is inspired by the inventory theory analogy described above and seeks to minimise the total cost associated with basic inventory models.

The I-TSCA functions by calculating this total cost (i.e. the sum of the setup cost, the holding cost and the stockout cost) associated with awarding green time to each phase of the traffic signal cycle. Green time is then awarded to the phase which will result in the lowest total cost, thus minimising the total delay time of all vehicles which are to pass through the intersection. Appropriate radar detection equipment, as described in §1, is assumed to be mounted at the intersection, allowing for the detection and tracking of individual vehicles in terms of their speeds, distances from the intersection and the headways between vehicles and the vehicles in front of them. These data are taken as input to the I-TSCA, allowing it to calculate the effective “customer demand” for each approach to the intersection and thereby for each phase of the traffic control cycle.

Three sets are associated with each lane of approach to an intersection. The first of these sets, \( C_j(t) \), contains all vehicles present along approach lane \( j \) at time \( t \). The second set, \( S_j(t) \), contains all stationary, queued vehicles along approach lane \( j \) at time \( t \) and is a subset of \( C_j(t) \). The third set, \( Q_j(t) \), contains both currently queued vehicles and vehicles which have not yet stopped, but will become queued, either behind an existing queue, or behind an amber or red traffic signal along approach lane \( j \) at time \( t \). Central to determining whether or not a vehicle will become queued, is the ability to predict the location of the queue position \( \rho_j(t) \) along approach lane \( j \) at time \( t \). This queue position indicates how far upstream from the intersection a vehicle queue reaches along approach lane \( j \) at time \( t \). Suppose that at time \( t \) the speed of vehicle \( i \in C_j(t) \setminus Q_j(t) \) is \( v_i(t) \) and that its distance to \( \rho_j(t) \) is \( d_{i,\rho_j(t)}(t) \), as illustrated in Figure 1. Should it be determined at time \( t \) that vehicle \( i \) will become queued, it is assigned a stopping point \( \mu_i = \rho_j(t) \), as well as a position \( \epsilon_i(t) \) in the predicted vehicle queue. The I-TSCA continually updates the predicted vehicle queue length \( |Q_j(t)| \) along approach lane \( j \). Vehicles are assumed to depart from a queue at a constant rate of \( \eta \) vehicles per second.

Each traffic signal cycle is assumed to comprise a set of \( \mathcal{P} \) phases and associated setup periods which precede each phase. A remaining green time \( \chi_m(t) \) and a remaining setup time \( \tau_m(t) \) are associated with phase \( m \) of the traffic signal control cycle. Suppose that vehicle \( i \) is travelling along intersection approach lane \( j \) which receives service during phase \( m \) of the traffic signal cycle. To determine whether or not vehicle \( i \) will become queued,
and hence added to $\mathbf{Q}_j(t)$, the I-TSCA assesses whether the time it will take vehicle $i$ to reach $\rho_j(t)$ is either less than the sum of the remaining red time together with the time required to clear the current predicted queue (if the traffic signal displayed is not green) or less than the time required to clear the current predicted queue of vehicles (if the signal displayed is green). If this is indeed the case, then vehicle $i$ will become queued, hence delayed and thus added to $\mathbf{Q}_j(t)$. In other words, if the traffic signal currently displayed at time $t$ is not green and $d_{i,\rho_j(t)}(t)/v_i(t) < \sum_{p \in \Phi \setminus \{m\}} (\chi_p(t) + \tau_p(t)) + |\mathbf{Q}_j(t)|/\eta$, then vehicle $i$ is immediately added to $\mathbf{Q}_j(t)$ and $\rho_j(t)$ is incremented by the effective length $\ell_i$ of vehicle $i$ (i.e. the actual length of vehicle $i$ together with the minimum space gap required between stationary vehicles). Similarly, if the traffic signal displayed at time $t$ is green and $d_{i,\rho_j(t)}(t)/v_i(t) < |\mathbf{Q}_j(t)|/\eta$, then again vehicle $i$ is immediately added to $\mathbf{Q}_j(t)$ and $\rho_j(t)$ is incremented by $\ell_i$. For the case in which $\mathbf{Q}_j(t)$ is empty (i.e. the $\rho_j(t)$ is located at the stop line of the intersection) vehicle $i$ is added to $\mathbf{Q}_j(t)$ under one of two conditions: If the traffic signal displayed at time $t$ is green and $d_{i,\rho_j(t)}(t)/v_i(t) > \chi_m(t)$ or if the traffic signal displayed at time $t$ is not green and $d_{i,\rho_j(t)}(t)/v_i(t) < \sum_{p \in \Phi \setminus \{m\}} (\chi_p(t) + \tau_p(t))$.

The I-TSCA iterates through three steps. The first step is the calculation

Figure 1: Vehicle characteristics. At time $t$, vehicle 1 is waiting at the intersection (i.e. $v_1(t) = 0$) along an approach lane. Vehicle 2 will become queued behind vehicle 1 within $d_{2,\rho_2(t)}/v_2$ seconds. Vehicle 3 is $d_{3,\rho_j(t)}$ metres from the currently predicted queue position.
of the required green time for each phase of the traffic control cycle. During the second step, the associated costs (in terms of vehicle delay time) are calculated for each phase of the traffic control cycle, should the previously calculated green time be implemented. The total cost of all the possible phase options are compared during the third step, and the phase achieving the lowest total cost is selected for service. More detailed descriptions follow of how these steps are performed.

**Step 1: Calculating the required green time of a phase.** The I-TSCA calculates the required green time $\gamma_j(t)$ for approach lane $j$ as the amount of time required to clear all vehicles contained in $Q_j(t)$. If it so happens that $Q_j(t)$ is empty for approach lane $j$, then the required green time of approach lane $j$ is taken as the time it would take for the front vehicle along the lane to clear the intersection. The required green time for approach lane $j$ is therefore

$$
\gamma_j(t) = \begin{cases} 
  d_{1;p_j(t)}(t)/v_1(t) & \text{if } |Q_j(t)| = 0 \text{ and } |C_j(t)| > 0, \\
  |Q_j(t)|/\eta & \text{if } 0 < |Q_j(t)| < \infty \text{ and } |C_j(t)| > 0, \\
  \infty & \text{if } |C_j(t)| = 0.
\end{cases}
$$

(1)

For the case where $|Q_j(t)| = 0$ in (1), $d_{1;p_j(t)}(t)$ is simply the distance between the front vehicle along approach lane $j$ and the intersection. The reason for selecting the vehicle closest to the intersection in determining the required green time is to maximise intersection utilisation. If $\mathcal{A}_m$ is the set of all approach lanes served during phase $m$ of the traffic signal cycle, then the required green times for lane $j$ are compared, and the minimum required green time over all the lanes is selected as the required green time for phase $m$, i.e. $\Gamma_m(t) = \min_{j \in \mathcal{A}_m} \gamma_j(t)$. The reason for selecting the minimum green time as representative for the phase is to ensure that a lane associated with a relatively large green time does not prevent other lanes with shorter associated green times from receiving service, since providing a large amount of green time may prove to be uneconomical in terms of minimising the total delay of all vehicles utilising the intersection. For this reason, if there are no vehicles approaching the intersection along a lane, then the associated green time of that lane is set to infinity so as to ensure that the lane is not considered for receiving a green signal.

**Step 2: Calculating the associated vehicle delay.** If it is determined that vehicle $i$ along approach lane $j$ will become queued, the I-TSCA calculates an expected delay term $\phi_{ij}(t)$, indicating how long this vehicle may expect
to be delayed. This delay term depends on whether or not approach lane \( j \) is receiving service, indicated by a binary parameter \( \kappa_j(t) \), where \( \kappa_j(t) = 1 \) if approach lane \( j \) is currently receiving service at time \( t \), or \( \kappa_j(t) = 0 \) otherwise. If approach lane \( j \) is served during phase \( m \) of the traffic control cycle, then

\[
\phi_{ij}(t) = \begin{cases} 
\tau_m(t) + \frac{\epsilon_i(t)}{\eta} - \frac{d_{i,\mu_i}(t)}{v_i(t)} & \text{if } \kappa_j(t) = 1, \\
\sum_{p \in \mathcal{P}} \tau_p(t) + \sum_{p \in \mathcal{P} \backslash \{m\}} \chi_p(t) - \frac{d_{i,\mu_i}(t)}{v_i(t)} + \frac{\epsilon_i(t)}{\eta} & \text{if } \kappa_j(t) = 0.
\end{cases}
\]

(2)

For the case where \( \kappa_j(t) = 1 \) in (2), the first term in the expression for \( \phi_{ij}(t) \) is the remaining setup time at time \( t \) before a green signal is displayed to approach lane \( j \) during phase \( m \). The second term is the amount of time it will take for the queue of vehicles in front of vehicle \( i \) (at position \( \epsilon_i(t) \) in the queue) to dissipate at a rate of \( \eta \) vehicles per second. The third term is the time that will elapse before the vehicle comes to rest in the queue. If the vehicle is already stationary, \( d_{i,\mu_i}(t) = 0 \) and this last term falls away. For the case where \( \kappa_j(t) = 0 \), the first three terms represent the time that vehicle \( i \) will be required to wait before a green signal is displayed to approach lane \( j \) and equals the sum of all remaining setup times of the traffic control cycle, as well as the remaining green times of all the other phases of the traffic control cycle, less the time it will take for vehicle \( i \) to become queued. The last term is again the amount of time vehicle \( i \) will be delayed while the queue in front of it dissipates upon receiving a green signal. In the case where vehicle \( i \) is not predicted to become queued, it will not be delayed and will therefore have an associated delay term of \( \phi_{ij}(t) = 0 \). The set \( \mathcal{J} \) contains all approach lanes to the intersection. To calculate the cost \( \Phi_m(t) \) associated with implementing the required green time for phase \( m \), the I-TSCA sums the delay terms of all detected vehicles approaching the intersection, \( i.e. \)

\[
\Phi_m(t) = \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{C}_j(t)} \phi_{ij}(t).
\]

The setup cost is accounted for by summing the required setup time \( \tau_m(t) \) for each vehicle that will be delayed during this period. The holding cost is accounted for by summing the delay terms for each vehicle along all lanes \( j \) for which \( \kappa_j(t) = 0 \). The stockout cost is accounted for by summing the delay terms of vehicles along all lanes \( j \) currently receiving service, should \( \kappa_j(t) = 0 \) (\( i.e. \) service be terminated) at time \( t \).

**Step 3: Assigning service.** Service is assigned to the phase \( m \) for which \( \Phi_m(t) \) is a minimum. If phase \( m \) is selected for service at time \( t \), it is assigned a green time of \( \Gamma_m(t) \). After assigning service to phase \( m \), the I-TSCA continues to re-evaluate required green times and associated delay costs, but
service is not switched until the assigned green time of $\Gamma_m(t)$ has elapsed. This is to prevent too frequent switching between phases. Once the assigned green time has elapsed (i.e. $\chi_m(t) = 0$), the I-TSCA once again compares all $\Phi_m(t)$ values for $m \in \mathcal{M}$, assigning service to the phase achieving the minimum resulting delay cost. It may occur that the phase which has just received service continues to be associated with the minimum resulting delay cost in which case it is assigned a new, extended green time and continues to receive service.

4. A traffic control algorithm inspired by the process of osmosis

When two liquids of differing solute concentrations are separated by a semi- or partially permeable membrane through which the solute cannot pass, the solvent passes by osmosis from the liquid with the lower solute concentration, through the membrane, and into the liquid with a higher solute concentration [26]. This movement is said to occur down a concentration gradient [27]. The liquid with a high water concentration and thus lower solute concentration has a higher osmotic potential or osmotic pressure [26] than that of the liquid with a higher solute concentration and solvent molecules continue to move from a region of high osmotic pressure to a region of low osmotic pressure until the osmotic pressures of the liquids on both sides of the membrane are equal [27].

The notion of osmotic pressure has inspired our second traffic signal switching control algorithm. In the context of traffic control at a signalised intersection, the solvent may be likened to the vehicles travelling along the intersection approach lanes while the solute may be considered as the space along the lanes not occupied by vehicles. The intersection itself may be considered as the partially permeable membrane. Each approach lane $j$ to an intersection is coupled with an exit lane $j'$ leading away from the intersection. The presence of vehicles requiring service along an approach lane $j$ may be considered to exert a push pressure on an intersection which effectively pushes vehicles through the intersection, in much the same way as a liquid with a high osmotic pressure would force solvent through a membrane into a liquid of lower osmotic pressure. The presence of space on an exit lane $j'$, on the other hand, would exert a pull pressure on the intersection which effectively pulls vehicles through the intersection, in much the same way as a liquid with a high solute concentration (and thus low osmotic pressure) would draw solvent through a membrane from a liquid of higher osmotic pressure.
4.1. The osmosis traffic signal control algorithm

The newly proposed osmosis traffic signal control algorithm (O-TSCA) does not attempt to minimise vehicle delay time at a signalised intersection explicitly, nor does it factor vehicle delay time into any algorithmic calculations when considering which phase should receive service. Instead, the O-TSCA relies on the sum of push and pull pressures exerted on an intersection by respective pairs of approach and exit lanes. As with the I-TSCA, the O-TSCA assumes the installation of suitable radar detection equipment at the intersection so that vehicles approaching the intersection along lane \( j \) and departing from it along the associated exit lane \( j' \) may be detected. Physical vehicle lengths are taken as input data to the O-TSCA, as well as the space which is not occupied by vehicles along the intersection approaches.

Suppose the length of approach lane \( j \) is \( \alpha_j \) (in metres) and that vehicle \( i \) travelling along lane \( j \) has an effective length \( \ell_i \), which is the actual length of the vehicle together with a minimum safety gap that has to be maintained between stationary vehicles (see Figure 2). At time \( t \) the O-TSCA calculates a demand \( \delta_j(t) \) associated with lane \( j \) and an availability \( \omega_j(t) \) associated with lane \( j \). The demand \( \delta_j(t) = \sum_{i \in C_j(t)} \ell_i \) is a measure of the total effective space occupied by the vehicles travelling along approach lane \( j \) at time \( t \). The availability \( \omega_j(t) = \alpha_j - \delta_j(t) \) represents the amount of space available for vehicles currently travelling along approach lane \( j \) to occupy once they have crossed the intersection onto the adjoining exit lane \( j' \). The pressure \( \pi_j(t) \) exerted by approach lane \( j \) at time \( t \) on an intersection, is the sum of its demand and availability at time \( t \), i.e. \( \pi_j(t) = \delta_j(t) + \omega_j(t) \). The implications of calculating the pressure exerted by an approach lane in such a manner are summarised in Table 1. It may be the case that a vehicle passes from approach lane \( j \) through the intersection onto one of any of a number of possible exit lanes. In this scenario, the exit lane \( j' \) associated with approach lane \( j \) is taken to be the exit lane with the least space available along it\(^4\). While approach lane \( j \) is receiving service, the O-TSCA calculates the throughput \( \theta_j(t) \) of the lane. This is the total space occupied by the vehicles which have passed through the intersection from lane \( j \) onto all adjoining

\(^4\)This is done so as to avoid the possibility of an approach lane receiving service in spite of insufficient space to accommodate vehicles on the adjoining exit lane \( j' \). An adjoining exit lane \( j' \) with little space available will not contribute significantly to the pressure exerted by approach lane \( j \) on the intersection and so the approach lane will not carry a high priority for service.
exit lanes.

<table>
<thead>
<tr>
<th>$\delta_j(t)$</th>
<th>$\omega_j(t)$</th>
<th>$\pi_j(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 1: Lane pressures relative to demand and availability. Both the demand and availability of approach lane $j$ results in an effective *pushing* and *pulling* of vehicles through an intersection. A large demand results in pressure which effectively pushes vehicles through the intersection (provided that there is sufficient space available along the adjoining exit lanes to meet the demand) while a large availability effectively pulls vehicles through the intersection (provided that there is sufficient demand to fill the availability).

To determine the pressure $\Pi_m(t)$ exerted by phase $m$ of a traffic signal cycle at time $t$, the O-TSCA sums together the pressures exerted by all the lanes served during phase $m$, *i.e.* $\Pi_m(t) = \sum_{j \in \mathcal{A}_m} \pi_j(t)$. The phase with the largest pressure is awarded service. If phase $m$ was most recently selected for service at time $t^*$, the combined demand and availability of all the approach lanes $j \in \mathcal{A}_m$ at time $t^*$ are stored as the variables $\Delta_m = \sum_{j \in \mathcal{A}_m} \delta_j(t^*)$ and $\Omega_m = \sum_{j \in \mathcal{A}_m} \omega_j(t^*)$, respectively. Phase $m$ continues to receive service until either

$$\sum_{j \in \mathcal{A}_m} \theta_j(t) \geq \Delta_m$$

(*i.e.* the cumulative throughput of all the approach lanes served during phase $m$ is at least as large as the cumulative demand of all the approach lanes served during phase $m$ at time $t^*$) or

$$\sum_{j \in \mathcal{A}_m} \theta_j(t) \geq \Omega_m$$

(*i.e.* the cumulative throughput of all approach lanes served during phase $m$ is at least as large as the cumulative availability associated with lanes served during phase $m$ at time $t^*$). Condition (3) ensures that all vehicles which were initially detected as requiring service do, in fact, receive service, while
Figure 2: Calculating the demand $\delta_j(t)$, the availability $\omega_j(t)$ and the pressure $\pi_j(t)$ of approach lane $j$ at time $t$. At time $t$, $\delta_j(t) = \sum_{i \in C_j(t)} \ell_i = \sum_{i=1}^{5} \ell_i$ and $\omega_j(t) = \alpha_{j'} - \delta_{j'}(t) = \alpha_{j'} - \sum_{i \in C_{j'}(t)} \ell'_i = \sum_{i=1}^{3} \ell'_i$. The pressure exerted by approach lane $j$ at time $t$ is therefore $\pi_j(t) = \delta_j(t) + \omega_j(t) = \sum_{i=1}^{5} \ell_i + \alpha_{j'} - \sum_{i=1}^{3} \ell'_i$.

Condition (4) ensures that no phase provides service to vehicles which cannot be accommodated on adjoining exit lanes due to a lack of availability of space. When either condition (3) or (4) holds, the O-TSCA again compares the pressures exerted by each phase, assigning service to the phase exhibiting the largest pressure. It may well occur that the phase which has just received service continues to achieve the largest pressure, in which case the phase will continue to receive service until the pressure of another phase exceeds it. If service is terminated to phase $m$ at time $t$, the throughput of all approach lanes $j$ served during phase $m$ is reset to zero (i.e. $\theta_j(t) = 0$ if $j \in \mathcal{A}_m$).

4.2. A hybrid self-organising traffic signal control algorithm

Initial testing of both the I-TSCA and the O-TSCA revealed that under lighter prevailing traffic demands the I-TSCA tends to outperform the O-TSCA, while under heavier prevailing traffic demands, the opposite is true in terms of minimising average vehicle delay. The reason for this is that under lighter traffic demands, the O-TSCA often provides green times that are longer than necessary, resulting in unnecessary intersection underutilisation.
On the other hand, during heavier traffic demand, the I-TSCA tends to switch too frequently between phases, resulting in green times that are too short to adequately clear queues of arriving vehicles. A hybrid self-organising traffic signal control algorithm was therefore developed in a bid to overcome these shortcomings by combining the I-TSCA and the O-TSCA so as to exploit the best attributes of both traffic control algorithms.

The hybrid algorithm runs both the I-TSCA and the O-TSCA concurrently together with an intersection utilisation maximisation supervisory mechanism (IUMSM). The IUMSM ensures (a) that the intersection is not underutilised due to green times that are too long, and (b) that service is not terminated prematurely to a platoon of moving vehicles ready to cross the intersection. This is achieved by considering the proximity of the nearest vehicles to the intersection along all intersection approach lanes. If, for example, a request is made by the I-TSCA for a signal change, the IUMSM determines whether or not there is a vehicle within a distance equal to the respective vehicle’s safe following distance of the intersection along any approach lane currently receiving service. If there is no vehicle within this specified range from the intersection, then the signal change takes place. If, on the other hand, there is a vehicle within this specified range, the current service is continued until there are no longer any vehicles within the specified range of the intersection along any approach lanes currently receiving service, or until the O-TSCA issues a request for a signal change. This ensures that green times are not terminated prematurely. Analogously, if there is at least one vehicle approaching a red signal, and this vehicle will reach the intersection within a time equal to or less than the duration of an amber and all-red phase, while at the same time there are no vehicles approaching a green signal which will reach the intersection within a time less than or equal to the duration of an amber and all-red phase, the IUMSM issues a change in service if neither the I-TSCA nor the O-TSCA has done so yet. This ensures that intersection underutilisation is avoided.

5. Algorithmic comparison and evaluation methodology

All algorithmic testing reported in this paper was carried out in a microscopic traffic simulation environment [28], facilitating the detailed analysis and evaluation of the performance of each algorithm and thus allowing for fair comparisons to made.
5.1. Traffic simulation modelling framework

A microscopic traffic simulation framework was designed and implemented in the Java-based simulation suite Anylogic 6.8 [29] for the purpose of comparing the efficacies of the algorithms described in §2.2, §3 and §4. This framework facilitates the construction of microscopic traffic simulation models in a customisable manner. The framework accommodates real-world data, as recorded by radar detection equipment, as input values and provides for the accessibility of these data to a variety of traffic control algorithms incorporated into the model. In particular, dynamically observed vehicle-specific data such as the length, position along a roadway, speed, and rate of acceleration and deceleration of individual vehicles are explicitly accommodated\(^5\). Due to the modelling complexity associated with these characteristics, they are often omitted from traffic simulation models in favour of the adoption of simplifying assumptions (e.g. that all vehicles are of uniform length and travel at constant, uniform speeds). In an attempt to better gauge the proficiency of the various traffic control algorithms at reducing vehicle delay and facilitating coordination among intersections in as realistic a manner as possible, the framework also allows for the incorporation of user-specified performance measure indicators into the model which make use of individual vehicle data and characteristics — this feature is usually also not available in commercial traffic simulation packages.

5.2. Purpose-built simulation model

A purpose-built microscopic traffic simulation model was implemented in the framework of §5.2 for the purposes of comparing and evaluating the efficiencies of the traffic control algorithms described in §2.2, §3 and §4. The road network topology adopted in this model is that of a corridor of four adjacent intersections. Each intersection comprises four approach roads, consisting of three lanes each at the intersection, as may be seen in Figure 3. Vehicles may turn left or travel straight through the intersection from the left-most lane of each approach, or travel straight through the intersection from the middle lane, or turn right at the intersection from the right-most lane (which is an exclusive right-turn lane). One of four phases may be implemented at any one time at each intersection. These four phases are shown in Figure 3. For

\(^5\)Pre-existing functions built into the road traffic library in Anylogic 6.8 allow for these data to be gathered in real time.
more detailed information on the design of the simulation model described in this section, the reader is referred to [15].

Figure 3: Four possible phases. (a) All vehicles travelling from West to East and East to West receive a green signal. Vehicles turning right do so on a permissible basis. (b) Exclusive right-turn phase for vehicles travelling from West to East and East to West. (c) All vehicles travelling from North to South and South to North receive a green signal. Vehicles turning right do so on a permissible basis. (d) Exclusive right-turn phase for vehicles travelling from North to South and South to North.

5.3. Performance measures

Several performance measure indicators were incorporated into the model of §5.2 in order to gauge and compare the effectiveness of the various traffic signal control algorithms described in §2.2, §3 and §4. These include the actual and normalised mean and maximum delay times experienced by vehicles in the system, the actual and normalised mean and maximum number of stops made by vehicles in the system, and the saturation levels of the system.

Since the desired speed and destination of each vehicle entering the traffic network is generated stochastically and known in advance in the model of §5.2, it is possible to calculate the time it would have taken each vehicle to complete its journey through the road network were it to not be impeded by any other vehicles or red traffic signals, by dividing the distance it has to travel by its desired speed. Subtracting this ideal travel time from the actual time the vehicle spends completing its journey results in the delay time of the vehicle. Upon exiting the traffic network, the delay time and number of stops made by each vehicle are recorded, allowing for further detailed statistical analysis at the end of each simulation run.
5.4. Experimental design

Six traffic signal control algorithms were compared and evaluated according to the performance measures described in §5.3 in the context of the simulation model of §5.2. These included the traffic signal control algorithms of Lämmer and Helbing (henceforth referred to as LH), Gershenson (henceforth referred to as Gersh), our traffic signal control algorithm inspired by inventory theory (henceforth referred to as I-TSCA), our traffic signal control algorithm inspired by the process of osmosis (henceforth referred to as O-TSCA), our hybrid traffic signal control algorithm (henceforth referred to as Hybrid), and a fixed-time traffic signal control algorithm (henceforth referred to as Fixed) for which the cycle length and green times at each intersection were calculated according to the procedure for equalising the degree of saturation on all intersection approaches, proposed in the Highway Capacity Manual [30]. The offset between green signals of adjacent intersections for the last algorithm was calculated as the time it would take vehicles to travel (at a speed limit of 60km per hour) from one intersection to the next and was implemented so as to facilitate the propagation of green waves along the corridor [10].

Each algorithm was tested under three prevailing traffic density conditions: light, medium and heavy traffic flow conditions [30]. For the light traffic flow scenario, the mean arrival rate for each entry approach to the traffic network was taken as 10 vehicles per minute. For the medium and heavy traffic flow scenarios, these mean arrival rates were taken as 20 vehicles per minute and 30 vehicles per minute, respectively.

For each traffic flow scenario and traffic control algorithm combination, the simulation model was run thirty times. Each simulation run was the equivalent of one hour of simulated traffic flow and was preceded by a sufficiently large\(^6\) warm-up period. For the cases of low and medium traffic flows, a sufficient warm-up period was found to be the equivalent of 30 minutes, as opposed to 15 minutes for the case of heavy traffic flow. A (pseudo) random number generator was used to generate travel plans for the vehicles during the simulation (including the arrival times of vehicles, their destinations, their desired speeds, etc). While the seed of this random number generator

---

\(^6\)The method proposed by Law [31] was employed to determine the lengths of appropriate warm-up periods. The model was considered to be warmed-up when the average number of vehicles present on the road network had converged to a steady state mean.
is different for each of the 30 simulation runs, the same 30 seeds were used for the testing of each algorithm in order to ensure a fair comparison between all the algorithms tested. The results obtained from these simulations are presented and analysed in the following section.

6. Simulation results

In this section, results are presented in the form of box and whisker plots for each of the six traffic control algorithms tested with respect to the mean delay time experienced as well as the mean number of stops made by vehicles under light, medium and heavy prevailing traffic flow conditions. An analysis and interpretation of the graphical results follows.

In each case an analysis of variance (ANOVA) test was carried out with respect to each set of results in order to determine whether or not there was a significant difference in the performances of the various algorithms (at a 95% confidence level). This was followed by employing the Tukey Honest Significant Difference (HSD) method [32] in order to determine which algorithms differed significantly from each other in terms of their results and by what margin (again at a 95% confidence level), thus effectively ranking the algorithms in terms of their performance in respect of the four-intersection case study.

6.1. Simulation results for light traffic flow conditions ($\lambda = 10$)

For the case in which the mean delay time was considered it was found that there was a significant difference between at least two of the traffic control algorithms, suggesting that the null hypothesis (that all the traffic control algorithms are equivalent) may be rejected. Hybrid was found to be the best performing algorithm (achieving an average mean delay time of 11.89 seconds), while the worst performing algorithm was Fixed (achieving an average mean delay time of 17.48 seconds) as may be seen in Figure 4a. At a 95% confidence level, Hybrid significantly outperformed Gersh, LH and Fixed by 6.2%, 26.7% and 32%, respectively. Hybrid also achieved the lowest average maximum delay time of 89.02 seconds as well as the lowest maximum delay time of 99 seconds. When considering the mean number of stops, it was found that O-TSCA was the best performing algorithm with a value of 0.87 stops, followed by LH and Hybrid, which it outperformed by 7.5% and 12.8%, respectively as illustrated in Figure 4b.
By nature, Gersh, I-TSCA and Hybrid are the more flexible traffic signal control algorithms and as a result yielded comparatively short green times, with those of Hybrid being on average approximately 1.4 seconds longer than those of Gersh and I-TSCA, resulting in a decrease in both vehicle delay and the number of stops made by vehicles. However, Fixed, employed the shortest green times of all the algorithms and it performed the worst in terms both performance measures. LH and O-TSCA employed considerably longer green times and while these were clearly less effective at reducing vehicle delay, they were very effective at enabling green waves as indicated by their superior performances at reducing the average number of stops made by vehicles. Overall, Hybrid may be considered to be the best performing algorithm under light traffic flow conditions as it achieves the best balance between flexibility and intersection coordination.

6.2. Simulation results for medium traffic flow conditions ($\lambda = 20$)

Under medium traffic flow conditions, Gersh achieved the lowest average mean delay time of 21 seconds, demonstrating a statistically significant improvement of 3% over the next best algorithm, Hybrid. The worst performing algorithm with respect to minimising mean delay was LH with an average delay.
mean delay time of 28.6 seconds, as shown in Figure 5a. Hybrid once again achieved the lowest average maximum delay time of 120.27 seconds, but this was not better than that of Gersh and O-TSCA at a 95% level of confidence. O-TSCA achieved the lowest maximum delay time of 137.92 seconds. Neither I-TSCA nor O-TSCA significantly outperformed each other with respect to mean delay time, but they were both outperformed by Hybrid at a 95% level of confidence, illustrating the effectiveness of the IUMSM. Once again, O-TSCA proved to be superior when considering the mean number of stops made, achieving an average of 1 stop, which was a statistically significant improvement of 8.7% over Hybrid, which was the next best algorithm, as may be seen in Figure 5b. Again, Hybrid may be considered the best overall performing algorithm overall in terms of achieving a balance between flexibility and intersection coordination under medium traffic flow conditions. As un-

![Figure 5: (a) Mean normalised delay time for $\lambda = 20$ vehicles per minute and (b) mean normalised number of stops for $\lambda = 20$ vehicles per minute.](image)

under light traffic flow conditions, shorter green times resulted in reduced delay times while longer green times resulted in a reduced number of stoppages. This discrepancy was successfully mitigated at a 95% level of confidence by Hybrid, which demonstrated that under medium traffic flow conditions it may be beneficial to award slightly longer green times. In fact, the average green times implemented by Hybrid (excluding exclusive right turn phases)
were almost double those of Gersh, which would appear to be the next best performing algorithm, overall.

6.3. Simulation results for heavy traffic flow conditions ($\lambda = 30$)

The large variances associated with the performances of Gersh and I-TSCA resulted in a large pooled sample variance and thus large confidence intervals as calculated by the Tukey HSD method. Therefore, although Fixed, with an average mean delay time of 32.91 seconds, outperformed Hybrid, O-TSCA and LH by 4.3%, 9.3% and 17.9%, respectively (as shown in Figure 6a), these improvements were not statistically significant at a 95% confidence level. Fixed also exhibited the lowest average maximum delay time of 164 seconds, but again this superiority was not statistically significant at a 95% level of confidence. As with the mean delay time, Fixed was again the best performing algorithm with respect to minimising the mean number of stops made by vehicles, as shown in Figure 6b with an average of 1.25 stops. Although this represents a 2.2% improvement over Hybrid, the next best performing algorithm, the improvement is not statistically significant at a 95% level of confidence. Based on these results it would appear that Fixed is the best overall performing algorithm for heavy traffic flow conditions, but that its performance is statistically indistinguishable from those of Hybrid and O-TSCA at a 95% level of confidence. The average green times implemented by Gersh were the largest of all six algorithms, while those implemented by I-TSCA were the smallest. As traffic flow and roadway saturation increases, so the demand along intersection approach lanes becomes more constant, lending itself towards fixed-time control. This explains the (expected) superior performance of Fixed as well as those of Hybrid and O-TSCA which were able to consistently implement average green times which were very similar to those of Fixed in duration with very small variance.

7. Conclusion

Based on the results presented in §6 it may be observed that Gersh and I-TSCA performed favourably with respect to reducing vehicle delay time under light and medium traffic flow conditions as they are the two most flexible algorithms. This came at the cost of intersection coordination, however, as they resulted in vehicles having to make stops more often. Under heavy traffic flow conditions, Gersh and I-TSCA implemented green times that were too long and too short, respectively. This resulted in ever-increasing vehicle
queue lengths and in some instances, gridlock occurred at times at some or all of the intersections. This resulted in the large variability in mean vehicle delay for both algorithms which may be seen in Figure 6a.

LH is slightly less flexible due to the fact that preference is given to the phase currently receiving service, resulting in longer green times. These longer green times resulted in vehicles having to make fewer stops, but the duration of these stops were relatively longer. O-TSCA was even less flexible due to the fact that it required all detected vehicles to have cleared the intersection before a change in service could be considered. Shortening the detection range of the radar may rectify this lack of flexibility under light traffic flow conditions.

Fixed was the least flexible of the six traffic signal control algorithms tested in this paper, as illustrated by its poor performance under light traffic flow conditions. As traffic demand increased and became more constant, however, so the performance of Fixed improved. Under heavy traffic flow conditions Fixed was found to be superior to three of the five algorithms, both in terms of reducing vehicle delay and the number of stops made by vehicles, but its performance was statistically indistinguishable from the other two algorithms at a 95% confidence level.
Hybrid performed suitably well with respect to both performance measure indicators for all three traffic flow scenarios. From a statistically significant point of view, Hybrid was never outperformed by the same traffic signal control algorithm with respect to both performance measures under any of the traffic flow conditions. This is due to the fact that it is flexible enough to adjust to small fluctuations in traffic demand (as a result of the influence of I-TSCA and IUMSM) without significantly sacrificing coordination among intersections (as a result of the influence of O-TSCA and again, IUMSM). Moreover, in several instances Hybrid significantly outperform LH and Gersh, which have both been shown to be efficient traffic signal control algorithms [3, 4, 10, 11, 12]. This is noteworthy, because Hybrid does not rely on carefully chosen parameters as do LH and Gersh. Instead, it relies solely on the prevailing traffic conditions and the availability of physical space along the roadways to inform its signal switching policies. It is acknowledged that both LH and Gersh may have been able to yield improved results through alternative parameter selection, but we contend that the time and effort required to find such improved parameter combinations would significantly outweigh the actual improvements that they are capable of achieving. We conclude that Hybrid is not only efficient at both reducing vehicle delay time and facilitating coordination among intersections, but that it is robust enough to do so under various traffic flow scenarios without requiring external intervention in the form of parameter fine-tuning.

8. Future Work

In terms of future work, it is recommended that the six algorithms tested in this paper be compared in the context a larger grid-like network topology comprising more intersections so as to ascertain whether the findings in this paper (which pertain to one-dimensional intersection coordination) can be corroborated in the setting of two-dimensional coordination.

Another desirable extension would be to compare the performances of the algorithms in an experimental setup during which the traffic flow conditions evolve stochastically as a function of time, so as to capture the formation and dissipation of peak conditions as opposed to merely assuming various fixed average arrival rates (as was done in this paper).

We conjecture that the performance superiority of Hybrid demonstrated in this paper will be more pronounced in a combination of the extended experimental setups mentioned above.
References


