Arrival time based Traffic Signal Optimization for Intelligent Transportation Systems

Vamsi Paruchuri  
Computer Science  
University of Central Arkansas  
Conway AR 72035 USA  
vparuchuri@uca.edu

Sriram Chellappan  
Computer Science,  
Missouri University of S&T  
Rolla, MO 65409, USA  
chellaps@mst.edu

Rathinasamy B. Lenin  
Department of Mathematics  
University of Central Arkansas  
Conway AR 72035 USA  
rblenin@uca.edu

Abstract—Road Transportation is a crucial component of today’s society, which drives several facets of our lives. The goal of intelligent transportation systems (ITS) is to improve the effectiveness, efficiency, and safety of the transportation system. Traffic signals are an elementary component of all road transportation systems. In order to maximize the productivity of a city, traffic signals must be able to efficiently control the flow of vehicles. Traditionally, current traffic signal optimization is based on traffic arrival rates, either estimated or forecasted.

In this paper, we illustrate that arrival time based solutions can outperform arrival rate based approaches. To the best of our knowledge, this is the first work that exploits arrival times of vehicles to improve traffic signal efficiency in order to reduce stopped delays and fuel consumptions, thus in turn reducing greenhouse gases and emissions. We show that arrival time knowledge can be utilized in realizing drastic gains in sparse load scenarios and significant gains in moderate load scenarios. The performance improvement translates to reducing stopped delays by over 40,000 hours daily and in reducing fuel consumption by over 650 gallons/signal/day.

Key Words: Scheduling; Traffic Signals; Vehicular Networks; Optimization;

I. INTRODUCTION

Road Transportation networks are a critical backbone of society today. With increasing demands of mobility, we are seeing an unprecedented growth in the number of vehicles on our roads. For example, in the decade between 1993 and 2003, there has been a 19% increase in the number of vehicles on US roads. This roughly translates to an increase of about 37 million vehicles in that decade alone. According to the U.S. Census Bureau, more than 87% of the American workforce today drive to work, with more than 75% driving alone. In fact, it has been reported that the average American spends more than 100 hours per year commuting to work alone, which roughly translates to about 23 minutes per working day. With increasing density in number of vehicle per road, traffic congestion is increasingly posing significant burdens on society [1-4]

The US Department of Transportation estimates that, in 2007, the annual cost of delays was 87.2 billion in urban areas alone, with about 2.8 billion gallons of wasted fuel consumption in 439 urban areas [1] and 53.2 billion pounds of carbon dioxide, considering that every gallon of gas produces about 19 pounds of CO2 [5]. In urban areas, road traffic contributes most to greenhouse gas (GHG) emissions: in the 10-county metropolitan Atlanta area, on-road cars and trucks account for 58% of emissions of nitrogen oxides and 47% of hydrocarbon emissions [2]. From an environmental perspective, road transportation related energy use and CO2 emissions per capita in US are about three times higher than compared to their European counterparts [6]. A report by the World Bank [7] estimated that 650,000 people died prematurely from urban air pollution in developing countries in 2000. The report also states that efficiency improvements in the urban transport system are critical in order to reduce transport sector emissions. Other impacts of traffic congestion include road rage [8, 9], obesity [4], delays in emergency vehicle responses [1], wear and tear of vehicles [10, 11] etc. Mitigating congestion in urban transportation networks is hence a critical need.

Road transportation accounts for about 9.9% of GHG emissions globally, and in United States, the contribution is about 21.6% [12]. During the 1996 Summer Olympics Games in Atlanta, when peak morning traffic decreased 23% and peak ozone levels decreased 28%, emergency visits for asthma events in children decreased 42% [3]. Non-point source pollution that include transportation-related pollutants is the biggest factor affecting Arkansas' water quality [13] and is linked to adverse health conditions like cancer and chronic illnesses.

Several works exist that aim at optimizing the traffic signal control and some of them are also being used in practice. Almost all of these works are based on arrival rate estimates based on the past measurements or based on forecasting. Few recent efforts that use arrival times to compute optimal time phases, and they only use predictions of very small time windows [18, 19]. Furthermore, in most of the existing approaches the short term forecasts is translated into fine-grained arrival for achieving real time adaptation of traffic signals. An added limitation of almost all current works is that
they focus only on high traffic conditions. However, majority of durations, traffic conditions are moderate or low when considering a city wide scenario. We postulate that significant fuel and emission reductions can be achieved even in low and medium traffic scenarios.

In this paper, we utilize traffic arrival times for improving the traffic signal performance and reducing vehicle delay times and correspondingly, fuel consumption and greenhouse gas (GHG) emissions, thus reducing the carbon footprint. Most of existing work in the past has used traffic arrival rate as a performance parameter for transportation networks. While it is relatively easier to compute signal transition times as a function of arrival rates, we illustrate that these approaches can be inefficient. The contributions of the paper are two-fold: (i) utilize vehicle arrival times rather than arrival rates for signal optimization and (ii) illustrate the performance benefits signal optimization can achieve even in low-traffic and moderate traffic conditions.

The rest of the paper is organized as follows: Section II presents a survey of related literature; Section III presents background and motivations; Section IV elaborates the arrival time based signal cycle optimization methodology; Section VI presents details about the simulator and presents performance evaluation results; and finally, Section VI presents concluding remarks.

II. RELATED WORK

US Congress launched the Intelligent Transportation Systems (ITS) program in 1991 [14]. One of its goals is to develop Advanced Traffic Management Systems (ATMS) that will rely on the consolidation of information, automotive, and highway technology. A wide range of small, complementary systems from electronic route guidance to pre-emptive signal control will essentially automate highways. Sensors and communication devices will be along the roads, as well as in the vehicles. Thus, the road will know its operational status including real-time information of the vehicles. Wide range of methodologies where a vehicle can access this information to make intelligent decisions have been proposed [14, 16]. Multiple efforts are underway within United States [14] to realize the benefits of ITS including Hattiesburg MS, Los Angeles, and British Columbia.

The use of sophisticated surveillance technologies, including inductive loop detectors and surveillance cameras at signalized intersections, enables traffic signal controllers to make use of real-time traffic information. This information, including, but not limited to, vehicle counts, link volume and link occupancy, proved to be very useful in computing real-time signal timing plans for both isolated and coordinated signal control.

Most modern traffic signal control technologies belong to this category. For the isolated control case, it was Miller [17] who first proposed a control strategy based on online traffic information. Other more recent methods include SCATS [18], PRODYN [20], OPAC [21], UTOPIA [22], SPPORT [23], COP [19]. It should be noted that although many of the above control strategies (e.g., OPAC, PRODYN and SCATS) are also used in coordinated control, the co-ordinations are mostly done heuristically due to the combinatorial complexity of the problem.

Other notable research that focuses on the coordinated control problem includes SCOOT [24], CRONOS [25], REALBAND [26], Lin and Wang [27], and Heung et al. [28]. Methodologies in systems like RHODES [29] and OPAC were presented to use real time traffic arrival time information, but the time horizons for which there are considered is very small (typically 1-2 minutes).

III. BACKGROUND AND MOTIVATION

In this section, we first present the traffic signal model we adopt and then review the major traffic signal control paradigms. Finally, we present a motivating example for intuitively understanding the performance improvements arrival time information can result.

A. Signal Model

We consider a typical four-legged isolated traffic intersection as shown in Figure 1 in which various possible traffic flows are shown in arrows according to National Electrical Manufacturers Association (NEMA) convention.

![Figure 1: A typical traffic signal intersection](image)

We denote four directions of an intersection by 0, 1, 2, and 3 as shown in Figure 1. There are four phases in traffic flow at any intersection: Phase 0: 0 → 2 and 2 → 0; Phase 1: 2 → 1 and 0 → 3; Phase 2: 1 → 3 and 3 → 1; Phase 3: 1 → 0 and 3 → 2. Let S denote the set of these 4 phases. We define a signal cycle of a traffic intersection as a starting time of green
light in phase 1, followed by the starting time of yellow light in phase 1, followed by green light in phase 2, so on until the time the green light of phase 4 starts. Let $\delta_y$ denote the yellow phase duration which is constant for all phases at an intersection.

### B. Existing Traffic Signal Control Paradigms

There are three major signal control paradigms of which two are currently used in practice and the third one is majorly a research topic.

i. **Actuated Control:** Actuated signal control requires actuation by a vehicle in order for a phase or traffic movement to be serviced. Actuation is achieved by vehicle detection devices like inductive loop wires in the pavement at or near the intersection and video detection [32]. Efficiency of actuated control primarily depends on settings and detector placement that minimize lost time, features that are very commonly absent in practice. Further actuated control is not well suited for arterials.

ii. **Fixed cycle coordinated control:** The traffic signal assigns right-of-way at intersection according to a predetermined schedule. The sequence of phases and time durations for each phase in the cycle is fixed, often based on historic traffic patterns. This paradigm is also often used along arterial and in scenarios where signalized intersections are closely spaced. By adopting a common cycle length and setting offsets equal to travel time between intersections, **platoons** of vehicles can be a given a green wave with almost no delay at any intersections. However, most often, it is feasible only to provide a green wave in only one direction. Furthermore, fixed-cycle strategies allow little deviation from a given phases’ schedules green durations and do not perform well in presence of long queues.

iii. **Cycle-free optimized control:** By using upstream detectors, arrival flow profiles at every approach can be predicted. Phaseswitching policies that minimize expected delay without the constraint of fixed cycle lengths or fixed phase sequences can be adopted. Optimized switching problem is shown to be exponential complete problem [33]. Some research efforts in U.S. using this logic include [34, 35]. This paradigm holds the most promise in minimizing the delays and improving the performance of traffic signals.

To the best of our knowledge, all the existing efforts focus on improving traffic signal efficiency by using arrival rates and most of them further focus on improving the performance in saturated conditions or highly congested conditions.

### C. Motivation

Fixed time signal control methods use historically recorded or observed volumes while several recent works use short term forecasting to measure real-time arrival rates [18, 19] as a performance parameter for transportation networks. While it is relatively easier to compute signal transition times as a function of arrival rates, there is a clear problem. For example, consider a simple traffic arrival pattern at an intersection in two conflicting directions say North to South and East to West, as shown in Figure 2. A vehicle has right-of-way if it arrives in green phase, while it has to stop and encounters delays, if it arrives in red phase. Based on the arrival rate based approach, allocation of green times is equal in both directions resulting in four vehicles stops (Figure 2.a). Exploiting arrival times for signalling can significantly improve the efficiency in the above example, wherein the resulting allocation of green times is optimal and results in only one vehicle stop (Figure 2.b).

Access to real-time location information of vehicles is envisioned to be a characteristic of ITS. Two notable efforts
are the Real Time Rome project [30] that uses cell phones and GPS devices, and RHODES [31].

IV. SIGNAL CYCLE OPTIMIZATION USING VEHICLE ARRIVAL TIMES

In this section we first present the notation and then the optimization problem of traffic signal control that utilizes vehicle arrival times. Later, we present a greedy solution that utilizes the arrival times.

A. Notation

Following is the notation used in this paper.

\( v_{ij}^k \) \( k \) th vehicle in direction \( i \) to \( j \); \( i, j \in \{0, 1, 2, 3\}, i \neq j \)

\( n^{i-j} \) Total number of vehicles travelling through the signal in direction \( i \) to \( j \) during a given time duration

\( a_{ij}^{k} \) Arrival time of \( k \) th vehicle in direction \( i \) to \( j \) at the traffic signal

\( g_{p}^{\rho} \) Start time of \( \rho \) th green phase in phase direction \( \rho \)

\( m^{\rho} \) The number of green phases in phase direction \( \rho \)

\( d_{ij}^{k} \) Departure time of \( k \) th vehicle in direction \( i \) to \( j \) at the traffic signal

\( D \) Set of departure times leaving the traffic signal in direction \( i \) to \( j \)

\( w_{ij}^{k} \) Queue length in direction \( i \) to \( j \) at a given arrival time \( a_{ij}^{k} \) at the traffic signal

\( q_{ij}^{k} \) Queuing (waiting) delay corresponding to vehicle \( v_{ij}^{k} \)

\( \mu_{ij}^{k} \) Fitness corresponding to vehicle \( v_{ij}^{k} \)

B. Optimization Problem:

Let \( \alpha = \{ a_{ij}^{k} \} (0 \leq k \leq n^{i-j}) \) be the arrival times of vehicles \( v_{ij}^{k} \) in the direction \( i \) towards \( j \). The set of beginning times of green phases (sorted in ascending order) is denoted by:

\[ \alpha = \bigcup \{ g_{p}^{\rho} \} \text{ where } 0 \leq \rho < m^{\rho} \]

Let \( M \) be \( | \alpha | \). We assume four green phases:

\[ \{ g_{m}^{0-2}, g_{m}^{2-3}, g_{m}^{3-2-1}, g_{m}^{1-0-3-2} \} \]

Given \( \alpha \), duration of green phase \( k \) can be computed as \( gd_{m}^{k} = \text{duration}(g_{k}) = g_{k+1} - g_{k} - \delta_{y} \), where \( \delta_{y} \) is the transition time from green phase to red phase (i.e., duration of yellow phase). \( F = \{ \mu_{ij}^{k} \} (0 \leq k \leq n^{i-j}) \) be the fitness corresponding to each vehicle, and primarily depends on amount of fuel consumed during decelerating, idling and accelerating phases while crossing a signal. By incorporating multiple factors as fuel consumption, greenhouse gas emissions and delays, the fitness can be modified easily so that the optimization algorithm can yield an appropriate solution as desired.

The optimization problem can now be formulated as:

\[
\begin{align*}
\text{Given } \alpha, \text{ find a schedule } \alpha \text{ such that } \\
\text{Objective Function: } \min \sum_{\forall k \in N} \mu_{ij}^{k} \\
\text{Constraints: } g_{m} \geq g_{m-1} + \delta_{y}
\end{align*}
\]

The only constraint above ensures that two green phases are apart by a time duration of at least \( \delta_{y} \). We note that there are no constraints on the minimum or maximum green phase durations. While it is desirable to have such constraints especially on the maximum duration so as to avoid starvation-like situations, here we are interested in the optimal solution and do not consider it. At the same time, we also note that it is not difficult to incorporate such constraints and the methodologies presented further in the paper are not affected.

\( D = \{ d_{ij}^{k} \} (0 \leq k \leq n^{i-j}) \) be the corresponding vehicle departure times. \( N \) is the total number of vehicles arriving at the traffic signal in a given duration.

Given a green phase schedule \( \alpha \), the departure time \( d_{ij}^{k} \) of a vehicle \( v_{ij}^{k} \) arriving at \( a_{ij}^{k} \) can be computed as follows:

\[
\begin{align*}
&\begin{cases}
\begin{align*}
&d_{ij}^{k} + q_{ij}^{k} \\
&\text{if } \max_{m} \{ g_{m}^{\rho} | g_{m}^{\rho} < d_{ij}^{k} - \Delta \} \leq d_{ij}^{k} \\
&\text{and } d_{ij}^{k} \leq \min_{m} \{ g_{m}^{\rho} | g_{m}^{\rho} < a_{ij}^{k} - \Delta \} + g_{m}^{p} \\
&\min_{m} \{ g_{m}^{\rho} | a_{ij}^{k} - \Delta \leq g_{m}^{\rho} \} + q_{ij}^{k} \text{ else}
\end{align*}
\end{cases} \\
\end{align*}
\]

\( q_{ij}^{k} \) is the queuing delay at the arrival time \( a_{ij}^{k} \) which can be computed based on number of cars waiting \( w_{ij}^{k} \) in the queue.

Optimal signal cycle computation becomes a trivial problem once optimal green phases are computed. However, computing green phases is non-trivial. In fact, the complexity of this problem is \( O(\alpha^{p}) \), where \( \alpha \) is number of phases in the
signal cycle and $\theta$ is the maximum number of total vehicles transiting during a phase across all signal cycles.

During our simulations, we were unable to find an optimal solution for a scenario consisting more than about 45 vehicles, thus prohibiting us from being able to perform any meaningful comparisons with any approximate solutions.

C. An approximate solution

We propose a computationally feasible solution in order to exploit the arrival times to improve traffic signal efficiency. Time is divided into slots of fixed duration $T$. In a given time slot $t_n$, let $\phi_{nj}$ be the number of vehicles arriving in that time slot corresponding to phase $\rho$. Assume that phase $\rho_j$ is being served green in current time slot $t_{n-1}$. At the end of a time slot $t_{n-1}$, the traffic signal computes the phase to be turned green based on the following logic:

if \( g_{\rho_j} < (g_{\max} - T) \) OR \( q_{\rho_j} = 0; \forall \rho_j \in P, i \neq j \)

if \( d_{\rho_j} \) is red in $t_n$ \( > d_{\rho_j} \) is red in $t_n$ \( \forall \rho_j \in P, i \neq j \)

else

continue green phase for $\rho_j$

else

max $\rho_j \in P \left[ d_{\rho_j} \right.$ is red in $t_n$ $\left.] \right.$

The above algorithm ensures that a green duration ($g_{\max}$) for any phase does not exceed the maximum green duration unless there are no vehicles to be served in other phases. A phase’s green duration can be extended to time slot $t_n$ only if the current green phase will not exceed $g_{\max}$. Further, it can only be extended if switching results in an additional delay over the entire phases. This delay is computed greedily, under the assumption that the current phase will be made green in the immediate next time slot itself. If continuation of green phase will not result in overall minimum delay, then the phase that can maximize the reduction in delay is turned green.

The duration of the time slots, $T$ plays a critical role in the performance of this algorithm. A very small $T$ will result in very fine granularity and too many switches, resulting in possible loss of precious time in terms of yellow phases. A large $T$ can result in unnecessary delays for vehicles waiting in queues. In this paper, we set $T$ to $3*t_c$, where $t_c$ is duration of yellow phase. The rationale behind this choice is that it corresponds to a phase turning red (one $t_c$), a minimum green phase equivalent to one $t_c$ for the corresponding green in a different phase and turning back to green (one $t_c$). A different minimum green phase duration can be easily accommodated by adjusting $T$ correspondingly. From our performance analysis, we found that this choice of $T$ consistently resulted in minimizing average vehicle delays.

V. SIMULATION RESULTS

We study the performance improvement at different traffic loads. As a first performance measure, we introduce and study stop probability, the probability that a vehicle has to stop at the signal before crossing the signal. The vehicle stops either if it does not have right-of-way (i.e., red phase) or if the queue is being discharged. The second performance measure is average delay per vehicle. The delay of a vehicle is computed as the additional travel time for the vehicle when compared to the case of uninterrupted movement of the vehicle with no traffic signal being present. Thus, it comprises of both stopped delay and delays resulting due to slowing down and acceleration.

We consider three different scenarios: Arrival Time based approach (AT) where the traffic signal phases are optimized to minimize delay using prior information of arrival times; Arrival Rate (AR) based approach with a signal cycle duration of 60 seconds (AR ($C = 60$ sec)) and Arrival Rate based approach with a signal cycle duration of 90 seconds (AR ($C = 90$ sec)). We assume a yellow phase duration and minimum green phase duration of 5 seconds.

A. Stop probability

The stop probability for an isolated two-lane two-phase intersection with no turning movements is shown in Figure 3 for the three scenarios.

Arrival rates are considered to be equal in all directions and vehicle arrivals follow a Poisson distribution. Yellow (amber) duration is assumed to be 5 seconds. The most notable aspect is that for rate based approaches the stopping probability remains above 0.5 even for near-zero arrival rates. This can be understood by noting that the green phase duration is 25 seconds for cycle duration of 60 seconds and a vehicle has to arrive during the green phase to cross without stopping (assuming that no vehicles are in the queue). However, for AT based approaches, since arrival times are known, the phases can be scheduled to minimize the delays and the stops. The stop probability of AT based scheme converges to that of AR based schemes as the arrival rate of vehicles increases.

B. Average delay per vehicle

The second performance measure, average delay per vehicle for both AT and AR based approaches is presented in Figure 4. We note that even though the stop probabilities for AT based approach is similar to AR based approaches when the arrival rate is more than 3 vehicles per minute (vpm), the average delays are still significantly lower. The reason is that the AT approach adapts phase durations to minimize the delays based on arrival times while the phase durations remain constant for AR based approaches.

We conclude that arrival time knowledge can be utilized in realizing drastic gain in sparse load scenarios (< 6 vpm, i.e. < 360 vehicles per hour (vph)) and significant gains in moderate
load scenarios (< 10 vpm, i.e. < 600 vph). The performance improvement would translate to reducing stopped delays by over 40,000 hours compared to AR (C = 60 sec) for arrival rate of 360 vph in each direction. Considering that the average idling car consumes about 0.156 gallons of gas per hour [5] and that each gallon of gas produces about 19 pounds of CO2 [5], the savings result in reducing fuel consumption by over 650 gallons/signal/day and reducing CO2 emissions by over 12,500 pounds/signal/day.

C. Applicability of the performance improvement

We analyzed the hourly volumes data of over 90 locations across the state of Utah [20] for the month of June 2010. For all the rural locations, less than 10% of collected hourly volumes were greater than 360 and less than 2% were greater than 600 vph. For locations functionally classed as "Urban Minor Arterials", about 65% hourly volumes were less than 360 vph and over 84% were less than 600 vph. For 32 locations functionally classed as "Urban Major Arterials" and primarily located in Salt Lake City, 35% hourly volumes were less than 360 vph and over 58% were less than 600 vph. While the above conclusions are based on data collected in only one state, we strongly believe that arrival time information can significantly improve performance of traffic signals in almost all rural areas, majority of semi-urban locations and significant number of urban locations.

VI. CONCLUSION

While arrival rate based approaches have been vastly studied, arrival time based approaches can result in vast improvements in terms of delay, fuel consumption and greenhouse gas emissions. Furthermore, great savings can be improved by optimizing the traffic signal control in moderate and low traffic scenarios.

ACKNOWLEDGEMENTS

This research is partially supported by University Research Council at University of Central Arkansas by grant numbers URC 364R03 and 364R04. This research is also partially supported by the National Science Foundation under Grant No. 1205695. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. Support from University of Missouri Research Board is also acknowledged.

REFERENCES


