Self-Organizing Traffic Signals Using Secondary Extension and Dynamic Coordination Rules

Burak Cesme <bcesme@coe.neu.edu> (corresponding author)
Department of Civil and Environmental Engineering
Northeastern University
360 Huntington Avenue, 400 Snell Engineering
Boston, MA, 02115
617.373.3991

Peter G. Furth <pfurth@coe.neu.edu>
Department of Civil and Environmental Engineering
Northeastern University

Submitted for Presentation & Publication to the Transportation Research Board
July 31, 2012; revised November 12, 2012

Word Count:
6,435 + 4 figures / tables @ 250 = 7,435

Abstract

Actuated control is very efficient for isolated intersections, but along arterials it lacks the means to synchronize signals, leading to frequent stops. Industry practice in the United States is to use fixed cycle coordination for arterial control, with signals running with a common cycle and offsets that help ensure good progression. However, fixed cycle coordination has many drawbacks, among them increased delay for pedestrians, promoting speeding, and a low degree of flexibility for accommodating transit priority. This research proposes a new paradigm for traffic signal control, "self-organizing signals," that begin with a base of actuated control, but with rules added that can lead signals to synchronize with their neighbors. Rules were developed for secondary extensions to serve an imminently arriving platoon and, for closely spaced intersections, dynamic coordination which forces signals to synchronize with a critical intersection. Simulation tests in two corridors indicate the success of this self-organizing logic, with delay reductions of up to 14% compared to an optimized coordinated-actuated scheme where there is no transit priority. With transit signal priority applied aggressively, transit delay was reduced by more than 10 seconds per train per intersection (more than 50% reduction in transit delay) with almost no impact to traffic.

Signal control can play an important part in making urban transportation more sustainable. Efficient control can allow streets to serve motor traffic with smaller footprint, and can be a means of reducing pedestrian delay, promoting traffic safety, and promoting transit use by facilitating short cycle lengths and transit priority.

Actuated control is very efficient for isolated intersections, but along arterials it lacks the means to synchronize signals, which can lead to frequent stops. Industry practice in the United States is to use fixed cycle coordination for arterial control, with signals running with a common cycle and offsets that help ensure good progression. However, fixed cycle coordination has many drawbacks, among them increased delay for pedestrians, promoting speeding, and a low degree of flexibility for accommodating transit priority. This research is a step toward developing a new paradigm for traffic signal control, "self-organizing signals," that begins with a base of actuated control, but with rules added that can lead signals to synchronize with their neighbors. The goal is that self-organizing signals should combine the benefits of actuated control with the benefits of progression while providing a flexible framework that can allow aggressive transit signal priority.

This paper describes the basic framework of self-organizing control with algorithms for control when arterials are below saturation. A companion paper [1] describes the development of self-organizing logic when traffic is near or beyond saturation.

Actuated Control

The operating principle of actuated control is to keep a phase green until the phase is no longer discharging at or near the saturation flow rate. This operating strategy keeps cycle lengths to a minimum and prevents overflow (a remaining queue when a signal turns red) due to random fluctuations in demand. Because control delay tends to increase with cycle length and to increase dramatically when there is overflow, actuated control is more efficient than fixed-time control, and results in near-minimal delay [2].

Actuated control is amenable to transit signal priority. If a phase is delayed or shortened due to a priority interruption, actuated control will naturally compensate by allowing it, on its next realization, to run until the queue is discharged. In this way, the signal can quickly recover from priority interruptions, making their impact on traffic delay small, and thus allowing aggressive priority tactics that result in near-zero delay for transit [3]. Actuated operation with aggressive priority is commonly used in Zurich, the Hague, and many other European cities known for fast and reliable tram and bus service.

Detector configuration and control settings can affect how quickly the controller detects when a traffic movement is no longer discharging at or near the saturation flow rate and switches control to the next phase. The lost time that occurs from not detecting and switching efficiently can substantially increase cycle length and delay [4]. Three important elements of efficient control are upstream detectors, short unit extensions, and non-simultaneous gap-out [2] [4], features that are routine in Dutch practice but less common in American. On multilane approaches, Cesme and Furth [5] have shown that using multi-headways (e.g., on a three-lane approach, the time between vehicle *n* and vehicle *n*+3) makes it possible to more precisely detect when saturation flow has ended, reducing lost time, cycle length, and delay compared to traditional gap detection.

Drawbacks of Fixed Cycle Coordination

Along arterials, actuated operation has an important drawback – it doesn't provide good progression (green waves) except under special circumstances described under the heading "Actuated Control's Self-Organizing Potential". If signals can be coordinated so that the platoon of traffic discharged from one intersection arrives when the signal is green at the next intersection, delay will be very little regardless of cycle length. The easiest way to provide this kind of progression is to use fixed cycle coordination, in which every signal along the arterial operates with a common, pre-timed cycle length, with pre-timed splits (allocation of the cycle to conflicting phases) and offsets (difference in green start time for the arterial movement) that correspond to the travel time between intersections.

Fixed cycle coordination has become the dominant method used for arterial control in the United States. It can easily be configured to provide a green wave in one arterial direction, giving the through traffic in that direction near-zero delay, making this scheme especially suitable for one-way arterials. On two-way arterials, green waves can be provided in both directions if intersection spacing is such that travel time between intersections equals a multiple of half the cycle length. Where spacing is not ideal, finding a compromise between the two directions and using flexibility in whether left turn phases lead or lag can sometimes lead to reasonably good progression. As a result, it has been estimated that coordination can reduce delays and stops by between 10 and 40 percent, depending on the prior method of signal control, traffic flows, and road layout [6].

However, fixed cycle coordination, which includes coordinated-actuated control and common forms of adaptive control, has significant drawbacks. First, the need for a common cycle length results in longer cycle lengths, because it forces all the intersections to adopt the cycle length needed by a single "critical" intersection. This is not a problem for traffic that benefits from a green wave, but for road users that don't – pedestrians, transit vehicles that make stops, cross traffic, and turning traffic – longer cycles generally mean longer delay. Because each intersection along an arterial has different cross-traffic and other characteristics, each has its own natural cycle length, meaning the cycle length it would have if operated in isolation. For example, one intersection may need a cycle at least 60 s long, another 70 s, and other 100 s; with fixed cycle coordination, all three would have to run with a 100 s cycle.

Second, with pre-timed operation, there is always a considerable chance of overflow unless slack is built into the cycle, and so the "needed" cycle length with pre-timed control will be longer than the average cycle length that will occur with efficient actuated control. Again, the longer cycle increases delay for pedestrians and others not benefitting from a green wave. Coordinated-actuated control allows unneeded green to be reallocated from the uncoordinated movements to the coordinated movements (usually the arterial through movements), but it does not usually allow reallocation in the other direction, and so it, too, needs slack in the cycle length to limit the probability of overflow.

Third, coordinated schemes are not usually designed to deal with periods of oversaturation, when the critical objective should be to maximize capacity (or utilization of capacity), not progression. During periods of oversaturation, as overflow becomes predominant, progression benefits disappear anyway.

Fourth, fixed cycle coordination is poorly suited to transit signal priority because it has a limited ability to handle priority interruptions. In order to provide effective signal priority without causing dramatic impacts on traffic, a control system should be inherently interruptible [7]. Fixed cycle coordination offers a limited range within the signal cycle in which transit vehicles can get green extension or an early green start. Equally importantly, it lacks mechanisms needed to compensate movements affected by a priority interruption, say, by giving them a longer green period in the next cycle [3]. Coordinated-actuated operation offers some flexibility for compensating a coordinated phase interrupted by a cross-street, but it has no mechanism to compensate the cross street or a turning movement hurt by a priority intervention favoring transit operating along the arterial [3]. And if a priority interruption is so severe that it "knocks" an intersection out of coordination, the process used to recover synchronization can lead to capacity shortfalls and long delays. Because of this inherent lack of flexibility, many transit signal priority applications are so timid (e.g., allowing only short green extensions, not allowing priority in consecutive cycles) that they show little benefit – often less than 3 seconds in transit delay reduction per signalized intersection.

Fifth, the longer cycles that accompany fixed cycle coordination can lead to long periods of unsaturated green, when the signal remains green even though the queue has discharged. Dutch traffic engineers consider this a safety problem, believing that it promotes speeding and pedestrian non-compliance, and is one of the reasons that Dutch practice favors actuated control over coordinated control.

Sixth, the timing plan being operated is usually optimized for traffic volumes collected at a previous date and time, and therefore can run much (or all) of the day with a plan that is poorly adapted to current traffic conditions. Adaptive traffic control systems such as SCOOT and ACS-Lite overcome this weakness by constantly measuring traffic volumes and updating the signal timing parameters frequently [8].

Finally, local road and traffic characteristics can make it such that the coordination benefits of fixed cycle coordination fall far short of the ideal of everybody enjoying a green wave. The irregular and non-ideal intersection spacing prevalent on many arterials makes it such that good two-way coordination is not achievable; progression will either be good for one direction and poor for the other, or a compromise solution will be found in which stops are frequent in both directions. In urban settings, a large fraction of arterial traffic does not continue through, but turns on or off at some point. Traffic turning off the arterial reduces platoon density, undermining the efficiency of a green wave, and suggesting that optimal control should involve stopping mainline platoons in order to restore their density.

Actuated Control's Self-Organizing Potential

A system is described as self-organizing when local elements interact with each other in order to achieve dynamically a global function or behavior. Self-organizing is a process whereby the pattern at the global level of a system emerges solely from the interactions among the lower level or local components of the system.

On the single intersection level, actuated control is self-organizing. Its only global control is establishing the phase sequence to be followed every cycle (and this, too, can be relaxed); otherwise, rules govern each signal phase directly, telling them when to start and stop based on information from phase-specific detectors. What emerges from this decentralized control is cycle lengths and splits that adjust themselves to actual demand as if they had been globally optimized.

Gershenson and Rosenblueth [9] have further demonstrated that actuated control has the potential for achieving self-coordination in street networks. Their test network is a grid of one-way streets, with alternating northbound and southbound streets crossed by alternating eastbound and westbound streets. A set of rules governs each intersection approach based on local detectors. They include an extension rule (hold the signal green while a dense platoon is passing), a spillback prevention rule (end the green if the downstream segment is blocked), a gap-out rule (end the green if there is no more queued traffic and the competing traffic stream has a queue), and an interrupt rule (switch control to a direction if the accumulated waiting time in its queue is long). Visualizations show the strong coordination that is achieved, with platoons advancing with little or no delay through successive intersections in all directions. Summary statistics show lower average delay than with a globally optimized coordination plan that provides green waves for all directions, because the self-organizing system adapts to random fluctuations in demand.

One mechanism by which actuated signals coordinate with their neighbors comes through the information carried downstream in a platoon. Actuated signals are programmed to end their green upon detection of the end of a passing platoon. If the offset between two neighboring intersections ever becomes such that a platoon discharged from the first intersection arrives the second intersection while the platoon phase is green, the second intersection will hold its green until the platoon passes, and thus the two intersections will end their green with an ideal progression offset. Suppose now that the cross-traffic demands at the two intersections are such that the red interval at the first intersection is just as long as or a little longer than the second intersection's red interval. Then the next platoon released from the first intersection will again arrive on green at the second intersection, maintaining an ideal end-of-green offset. However, if the first intersection's red period is shorter than the second intersections, or is much longer, then this synchronization can be lost.

Clearly, this coordination mechanism is rather limited. Furthermore, it is not clear to what extent the remarkable coordination achieved by Gershenson and Rosenblueth is an artifact of their simplified model with one-way streets, equal average traffic demand on all streets, no turning traffic, and no lost time when switching control. Still, their results raise the question of whether other coordination mechanisms can be developed for actuated control that will lead to self-coordinating behavior while maintaining the essential benefits of actuated control.

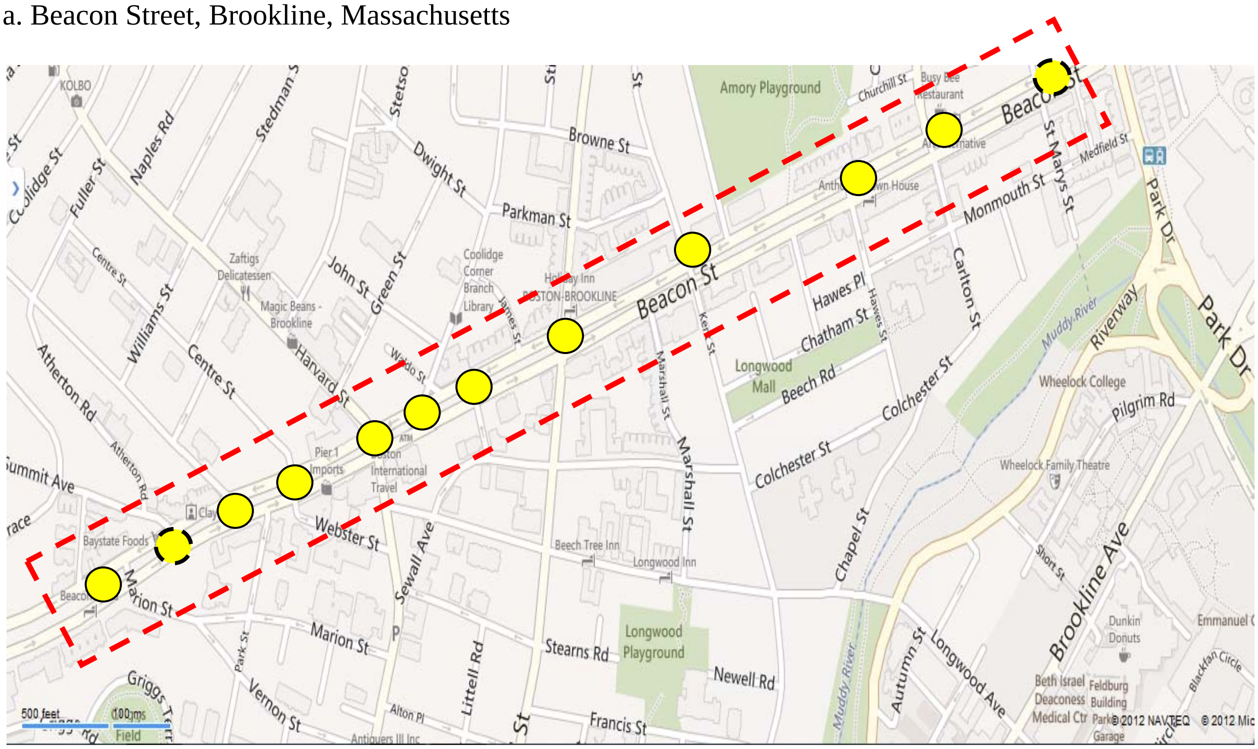
Objective and Experimental Environment

The goal of this research was to develop and test rules that lead actuated control to be self-organizing at an arterial level in a standard traffic microsimulation environment with realistic traffic flows. An important feature should be that the system is self-healing, recovering quickly and easily from interruptions due to transit priority. This approach is fundamentally different from optimizing methods used in adaptive control schemes such as Rhodes and OPAC [10] [11].

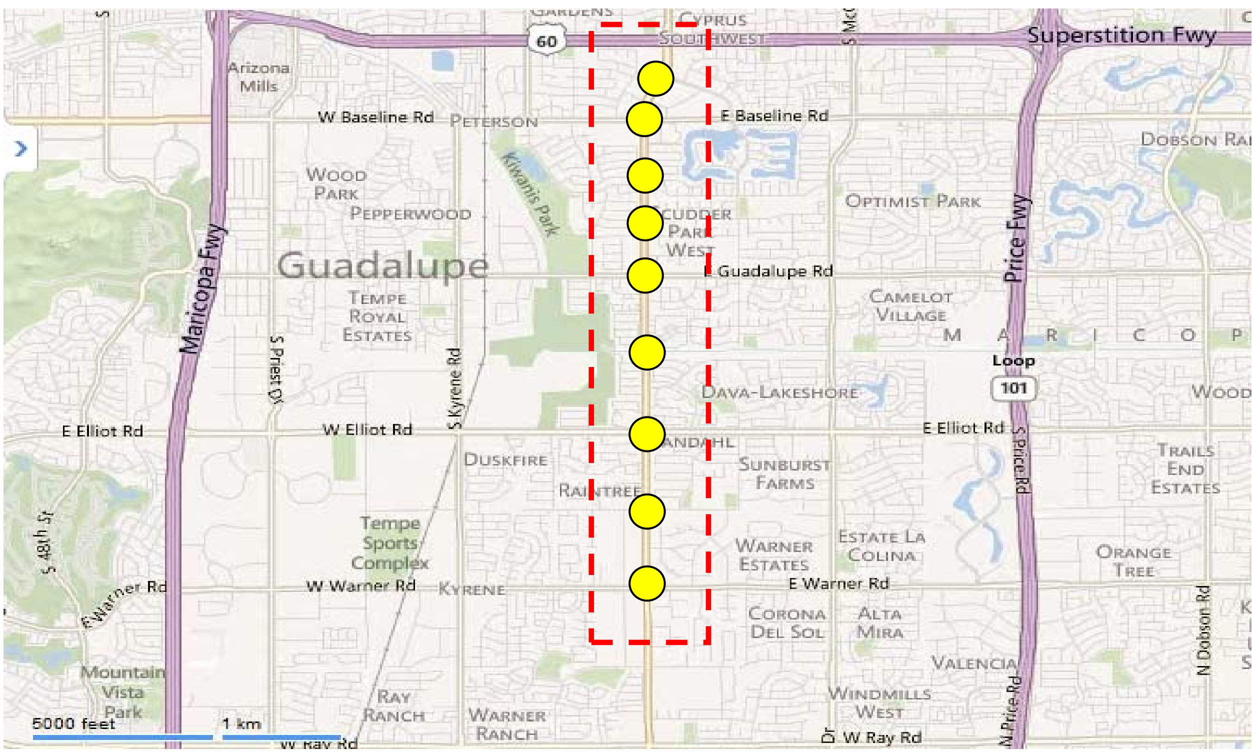
The microsimulation environment used was VISSIM. Control logic was programmed in C++ and run in a parallel program using VISSIM's application programming interface in which at every time step, detector information is passed from the simulation to the control program, which in turn passes information about signal states back to the simulation program.

Traffic was modeled on two arterials, shown in **Figure 1**. One is Beacon Street in Brookline, Massachusetts, from just west of the Marion Street intersection to just east of the signalized crossing at St. Mary's Street, a distance of 1.3 miles with 12 intersections with irregular spacing varying from 280 to 970 feet. Light rail transit (LRT) operates in a median reservation, with six stops in each direction. Traffic counts and signal timing parameters came from a 2009 study by Vanasse Hangen Brustlin (VHB), Inc., and transit schedule and passenger on-off data were obtained from the transit agency and field studies. To model the random arrival of trains at the first stop of the study segment (which is not the beginning of the line in either direction), dummy stops were placed at the start of the segment with random dwell times. To model the effect of crowding on boarding time, dummy traffic signals that applied only to trains were placed just after each stop to hold the train for a computed crowding penalty, as in Wadjas et al. [12].

a. Beacon Street, Brookline, Massachusetts



b. Rural Road, Tempe, Arizona



Signalized Intersection   Midblock Signal (signalized)   Study Area

Figure 1: Study Intersections along Modeled Arterials

The second test-bed was a segment of Rural Road with nine signalized intersections in Tempe, Arizona, a suburb of Phoenix. Traffic counts and signal timing information were obtained from Maricopa Association of Governments (MAG). Unlike Beacon Street, Rural Road has little traffic other than automobiles, and its intersection spacing is more regular and, except the last segment, at least 1300 ft, making it well suited to fixed cycle coordination. This test allows one to see what might be lost from abandoning fixed cycle coordination where it works relatively well.

The proposed self-organizing logic builds on standard full actuation logic, extending the green while there is no gap in the traffic, subject to minimum and maximum green times. On Beacon Street, minimum greens were long enough for pedestrians to cross in every cycle; on Rural Road, pedestrian phases were actuated. All approaches use upstream detection with the extension detector (detector used to detect gap-out) about 2 seconds' travel time upstream of the stop line, non-simultaneous gap-out, and multi-headway gap-out logic on multilane approaches.

To capture information about queue sizes, our logic uses pairs of detectors (i.e., a pulse detector upstream of the stopline and a pulse detector after the stopline) bracketing a "trap" within which the number of vehicles can be determined from the difference between cumulative entries and cumulative exits.

Rules for Secondary Extension

At isolated intersections, an efficient actuated controller cycles as quickly as possible, avoiding wasted green time by holding the green only until saturation flow has ended, subject to minimum greens. A short cycle minimizes delay for pedestrians and non-arterial traffic. However, if a platoon is expected to arrive shortly after an arterial approach's phase gaps out, it may be worth giving that phase a "secondary extension," meaning some additional green time beyond gap-out, because such an extension will drastically reduce delay for the vehicles in the platoon while only slightly increasing cycle length (and therefore delay) for other traffic.

From detectors placed upstream of approaches (ideally 20 seconds of travel time upstream), the controller monitors the arrival profile for an approaching platoon. The arrival profile is the cumulative expected arrivals at a time *t* in the future for *t* = 1, 2, … seconds. With every detector actuation, the controller updates the profile of arrivals to the intersection using a given travel time offset, without any platoon dispersion. Vehicles in an upstream intersection's queue can be included in the profile from the beginning of the change interval preceding its green start (in our experiment, they are included in the profile wherever intersection spacing is less than 20 seconds). Depending on intersection spacing, the expected arrival profile can often be known for 20 seconds or more.

Willingness to grant a secondary extension should increase as the set of arriving vehicles becomes larger, denser, or more imminent. A proposed measure that combines these three features is $L*$, the lost time per vehicle, defined as the ratio of wasted green time during the secondary extension to the number of arrivals during the secondary extension, and minimized by considering different potential lengths of secondary extension. Let time *t* be initialized so that the current time is 0, and let *L(t)* = lost time per vehicle if the secondary extension's length is *t*, given by

$$L(t) = \frac{t - n(t) * h_{sat}}{n(t)} \quad (1)$$

where *n(t)* is the number of vehicles expected to pass the stopline if the green phase is extended by *t* and h_{sat} is the saturation headway. The numerator is the wasted green time, i.e., green time in excess of what would be needed to serve the arriving vehicles if they discharged at saturation headway. *L(t)* is then calculated for discrete values of *t* up to SX_{max}, the maximum allowed length of secondary extension, to find an optimizing value of *t*:

$$L^* = \min_{t=2,4,\dots,SX_{max}} L(t) \quad (2)$$

In the common case in which the arrival profile includes a gap followed by a dense platoon of uniform density, $L*$ will be minimized at the time *t* at which the last vehicle in the platoon reaches the stopline. Observe that in such a case, $L*$ will be smaller if the platoon is larger, denser, and more imminent (shorter gap before it arrives).

Whether a secondary extension should be granted depends on how much excess capacity an intersection has. **Figure 2** displays affordable lost time per vehicle as a function of intersection *v/c* (volume to capacity) ratio, which is given by

$$v/c = \frac{\sum_{critical} \frac{v_i}{s_i}}{1-L/C} \qquad (2)$$

where v_i and s_i are a movement's arrival rate and saturation flow rate, the sum in the numerator is over critical movements only, L = sum of the lost time for the critical movements, and C = maximum desirable cycle length. Our tests used C = 90 s, assumed that L was 4 s per critical phase, and measured v and s adaptively (i.e., using detectors within the simulation environment, updating estimates every 5 minutes).

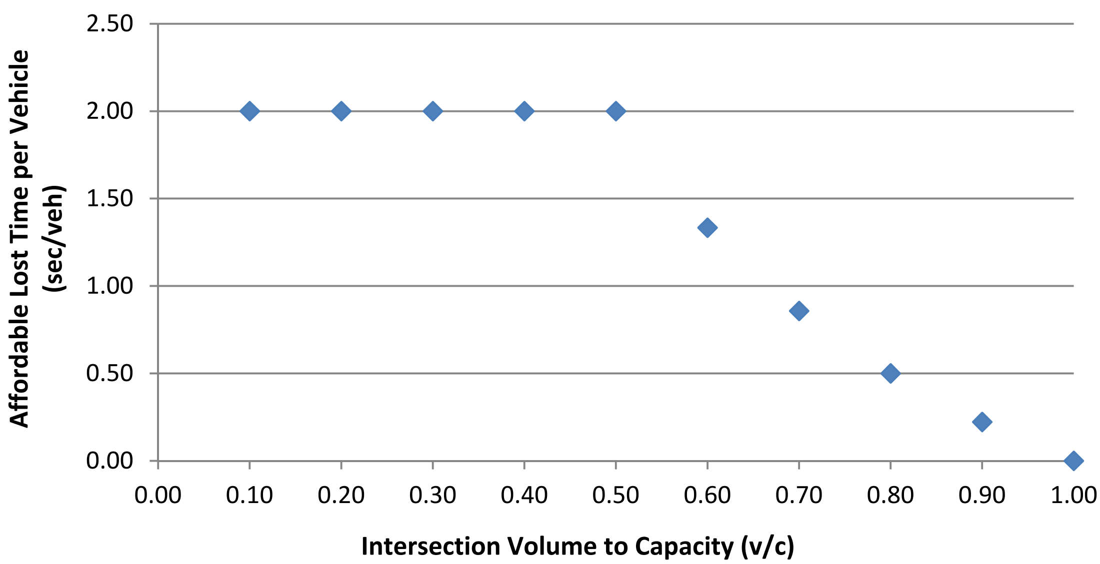


Figure 2: Affordable Lost Time per Vehicle with respect to Intersection Volume to Capacity (*v/c*) Ratio

As the intersection *v/c* increases, affordable lost time per vehicle approaches to zero, making it more difficult for an arriving platoon to secure a secondary extension, and if an intersection is over capacity, secondary extensions will be not allowed. The particular values used in this criterion were selected based on judgment and have not been empirically tested, offering an area for future refinement.

For phases serving the critical arterial through movement, the maximum allowed extension, SX_{max}, is 20 s. This generous limit was allowed because if a critical phase is extended in one cycle to serve a platoon that would otherwise have been served in the next cycle, the longer red interval that conflicting movements will face in the first cycle will be offset by a shorter red in the next cycle. The same is not true for the non-critical arterial through movement, and so a more stringent maximum is specified for it:

$$SX_{max} = \min\{\max(10, \Delta C_n), 20\} \quad (3)$$

where ΔC_n is the difference between intersection's natural cycle length and neighboring intersections' natural cycle length. Equation 3 limits the secondary extension for the non-critical arterial phase to 10 seconds when the local intersection's natural cycle length is close to that of its neighboring signals. But if a neighboring signal has a much longer cycle length than the local intersection, then a longer extension will be allowed, because it will improve progression for the critical as well as non-critical direction.

Each time an arterial phase gaps out, lost time per vehicle in the arrival profile is calculated. If it meets the criteria for affordable lost time, a secondary extension will be granted, with its length being the value of t that minimizes L^*. When the secondary extension ends, normal gap-out logic governs, allowing the controller to extend the green even further as long as the flow remains heavy. That way, an entire platoon can be served even if all of it was not detected because of a limited horizon or a limit in maximum allowed extension. A phase may receive only one secondary extension per cycle.

Dynamic Coordination for Coupled Intersections

When intersections are closely-spaced (up to 500 or 600 ft, depending on expected cycle length), a failure to provide good progression during higher demand periods can cause spillback or starvation, dramatically reducing intersection capacity. Closely spaced intersections also offer an opportunity for good two-way progression by starting their arterial greens at the same time, and on arterials with transit priority, there will often be no transit stop between them. Closely spaced intersections are therefore coupled and use different control logic that forces a coupled group of intersections to cycle synchronously, pushing platoons through the zone of coupled intersections with little or no interruption. This approach was inspired by Zurich's traffic signal control, which uses "dynamic coordination" within zones of two or three closely-spaced intersections, with zones separated by segments long enough to serve as buffers. Using dynamic rather than fixed cycle coordination makes the control logic flexible in order to give priority to their frequent trams, with green waves for cars often simply following the tram, which is usually allowed to progress through with no delay at all. Our logic likewise assumes that coupled zones are small, and is therefore not suitable for tight downtown grids with many closely spaced intersections.

Each coupled zone is governed by a critical intersection, the intersection with the greatest natural cycle length. Determination of which intersection in the zone is critical is done adaptively based on volume and saturation flow measurements over periods of five cycles; volume-capacity (v/c) ratio is likewise measured adaptively by tracking green and red times. Dynamic coordination within the zone aims to follow the critical intersection's critical arterial through phase ("mainline" phase). When the v/c ratio at the critical intersection is below 0.9, it aims for simultaneous green start for arterial phases in the coupled zone. Above 0.9, it aims for progression in the mainline direction with offsets designed to prevent starvation and spillback at the critical intersection by making sure that the upstream intersection(s) release in time for the critical intersection to have a ready queue, and that the downstream intersections clear in time for the platoon arriving from the critical intersection, as described in [1].

To achieve dynamic coordination, an "earliest activation time" is calculated for each arterial phase at the critical intersection at every phase transition based on the current signal state and commitments made by the local controller (e.g., minimum green, pedestrian clearance). A phase is "activated" when the clearance interval preceding its green is initiated. Using trap counts, minimum greens are calculated for every green start based on the time needed to discharge the standing queue; that helps limit the difference between minimum green and the green interval at which the phase will gap out.

Earliest activation time accounts for the possibility that an intervening phase that is not on recall might be skipped, unless a call has been registered. It is then adjusted by adding the expected time to clear any standing queue at the critical intersection (whose length is known from trap counts). The critical intersection communicates it to its neighboring intersections, who continue to pass it along peer-to-peer until the entire coupled zone has been reached, and, when critical intersection v/c is less than 0.9, this adjusted earliest activation time then becomes the scheduled mainline activation time for the non-critical intersections in the zone.

Normally, the non-critical intersections will be ready to activate their mainline before their scheduled activation time, since non-critical intersection naturally cycle faster than the critical intersection. Where the mainline phase is immediately preceded by a cross-street phase, that cross-street phase will be held after it gaps out until the scheduled activation time. This allows the slack time forced by dynamic coordination to be well used, both in reducing delay to cross-street vehicles who would otherwise have to wait for the next cycle, and by reducing cross-street demand for the next cycle to help ensure that a non-critical intersection won't become critical.

If the mainline phase is immediately preceded by an arterial opposing leading left, using the slack time to hold this leading left is not an efficient use of time unless it has an unusually high demand. Instead, the controller estimates the needed split for the left turn using a queue count amplified by 10% to allow for late arrivals, and subtracts it from the scheduled mainline activation time to schedule an activation time for the leading left. The cross-street phase preceding that leading left will then be held after gap-out until this scheduled activation time.

If the scheduled activation time arrives before the cross street has gapped out, the cross street is not truncated; it is allowed to complete its phase normally, in which case the arterial phase will begin later than scheduled.

Because scheduled activations in a coupled zone are not guaranteed, a further rule to help ensure dynamic coordination is giving arterial phases a "green hold" beginning at their scheduled activation time, not allowing them to give up the green they are about to begin until the upstream platoon (if there is an upstream signal in the coupled zone) has arrived, in which case normal gap-out logic will keep the phase green until the platoon has passed. This is especially important for any coupled signal downstream of the critical intersection, so that discharge from the critical intersection will not spill back.

Where the arterial has left-turn phases, two-way progression can be improved by sequencing phases with leading lefts entering a coupled zone (implying, for the opposite direction, a lagging through movement leaving the zone) and lagging lefts leaving the zone (implying a leading through movement entering the zone). This sequencing reduces the chance of starvation at the critical intersection. Lagging lefts at the end of a zone also provide better progression for cars turning left at the end of the zone.

Test Case I: Simulation of Beacon Street, Brookline, Massachusetts

The first test case is the Beacon Street corridor described earlier. It was modeled with three coupled zones, one with two signals (Webster-Winchester), and one with three signals (Harvard-Pleasant-Charles), and the last one with two signals (Hawes-Carlton). The remaining three signals were uncoupled, and of course the zones are not coupled to each other. LRT has no control at the two signals that are midblock pedestrian crossings; therefore, results for LRT should be interpreted as pertaining to ten signals.

Two important features of this corridor are heavy pedestrian use and the frequent LRT service operated in a median reservation. All pedestrian phases were put on recall. Conditional priority was applied for transit, giving priority to late vehicles but not to early vehicles in order to try to maintain even headways [13]. Priority tactics programmed into the control logic include green extension (up to 10 seconds of green extension), early green, phase rotation (switching leading and lagging lefts), and phase insertion (inserting a short transit-only phase that is not constrained by pedestrian timing needs), as described in [7]. All transit stops were modeled as far-side in order to make them more amenable to aggressive transit priority, even though the existing layout has a mix of far-side and near-side stops.

Because the Beacon Street LRT is not in conflict with the arterial phases, priority interventions for LRT can actually benefit arterial traffic. Therefore, in order to further test the system's ability to heal from interruptions that take time away from the arterial, the transit priority case also includes granting a 20 second extension for cross-street traffic on Harvard Street, starting when it gaps out, once every 6 minutes, as might occur from giving priority to a crossing bus route. In case of a conflicting call (i.e., LRT calls for priority at the same time that Harvard Street calls for interruption), priority was given to LRT.

Traffic counts for the morning peak show the critical intersection operating at 79% of capacity. In order to test model performance under more demanding conditions, the volume profile was scaled so that the critical intersection *v/c* in successive 15-minute increments took on the values {0.66 (warm-up), 0.66, 0.79, 0.93, 0.93, 0.66, 0.66}. Five simulation runs for each self-organizing and coordinated-actuated control were run, with results reporting averages for the 90-minute period following the warm-up. Coordinated-actuated control used timing parameters optimized using Synchro (with a cycle length of 80 s) with manual adjustments to offsets based on observing the simulation.

Table 1 provides a comparison. Without transit priority, self-organizing signals reduce overall vehicular delay by 14% compared to an optimized coordinated-actuated scheme. There was no appreciable change in transit delay. During the period with existing traffic volumes (t = 30 to 45), average network cycle length falls from 80 to 69.2 seconds, showing a strong benefit to pedestrians.

When aggressive signal priority was applied, the self-organizing signals were able to reduce transit delay by more than 10 seconds per vehicle per intersection, which is slightly more than 50% of the delay that occurs under coordinated-actuated logic. This is a substantial reduction, especially considering that only 69% of trains (those behind schedule) requested priority. If only the trains requesting priority are accounted for, average delay reduction for transit is 14 s per intersection, or 70% of their non-priority delay.

Table 1: Simulation Results for Beacon Street with and without Transit Signal Priority

	Optimized Coordinated-Actuated	**Self-Organizing**	
	No TSP	No TSP	With TSP
Average Network Delay(s/vehicle)	68.4 (0.0%*)	58.6 (-14.3%*)	70.5 (3.1%*)
Total Network Delay (vehicle-hours)	219.0 (0.0%)	187.8 (-14.2%)	227.0 (3.7%)
Inbound Train Delay (s/train)	197.8	216.9	98.9
Inbound Train Delay Change (s/train/intersection)	-	1.9	-9.9
Outbound Train Delay (s/train)	206.1	206.7	99.7
Outbound Train Delay Change (s/train/intersection)	-	0.1	-10.6
Percent of Trains Requesting Priority (only late trains request priority)	0%	0%	69%
Inbound train headway coefficient of variation at first stop in modeled network	0.426	0.426	0.426
Inbound train headway coefficient of variation at last stop in modeled network	0.580	0.562	0.437
Average Cycle Length (s) during base period (v/c = 0.79)	80.0	69.2	Not Measured

Note: Values in parentheses indicate percentage change in average network delay compared to optimized coordinated-actuated plan.

Equally important for transit is the large reduction in headway coefficient of variation (cv) for inbound (peak direction) service at the downstream end of the study segment. Without priority, headway cv tends to increase along a transit line as late trains, facing more than the usual number of passengers wanting to board, tend to become later, leading to serious crowding. The test shows headway cv remaining almost constant along the segment when priority is applied, rather than growing from 0.43 to around 0.57 without priority under both self-organizing and coordinated-actuated control. At the same time, the self-organizing logic showed its self-healing properties by keeping traffic delay within 3% of the optimized fixed cycle plan. Results for the signal-priority case include every-six-minute interruptions at the Harvard Street intersection as well as priority interventions for the Beacon Street LRT.

Test Case II: Simulation of Rural Road, Tempe, Arizona

Unlike Beacon Street, the Rural Road corridor has a geometry and traffic mix well suited to fixed cycle arterial progression, with intersections more regularly and distantly spaced, few pedestrians, and little transit service. The arterial was modeled with one coupled zone with two signals (Baseline-Minton); the remaining five signals were uncoupled. To test the algorithms for periods with volume below capacity, a.m. peak traffic volumes were scaled down to make the v/c ratio at the critical intersection equal 0.9. (Tests with a greater degree of saturation are described in [1].) Simulation results were obtained for a 60 minute period following a 15 min warm up period, with averages reported from five simulation runs. Timings for coordinated-actuated control were optimized for the adjusted volumes using Synchro; they include an 80 s cycle length except at Carver, where the cycle length is 40 s.

Simulation results indicated than self-organizing signals reduced overall vehicular delay by 7.1% compared to an optimized actuated-coordinated plan (average vehicle delay was reduced from 49.5 seconds to 46.0 seconds). Total vehicle delay was reduced from 163.6 vehicle-hours to 151.7 vehicle-hours, which resulted in a 7.3% reduction in total delay. This example shows that self-organizing signals can achieve synchronization at least comparable to that obtained using fixed cycle coordination even when the arterial has intersection spacings well suited to fixed cycle coordination. In addition, with self-organizing signals, the average cycle length (simple average over intersections) is 17% lower than with coordinated-actuated control (63 versus 76 s), indicated a reduction in pedestrian delay. At three major intersections, self-organizing signals had an average cycle length within 3 s of the coordinated cycle length of 80 s; the Carver intersection cycled on average every 39 s, and the other intersections between 49 and 63 s.

Figure 3 shows a record of signal state changes (red and green intervals) for the critical direction arterial phase under self-organizing control at four intersections on Rural Road – two major intersections with two minor intersections in between. In the coordinated plan, all four run with an 80 s cycle. With self-organizing control, one can see how the cycle and the green interval length varies within a given intersection as well as between intersections.

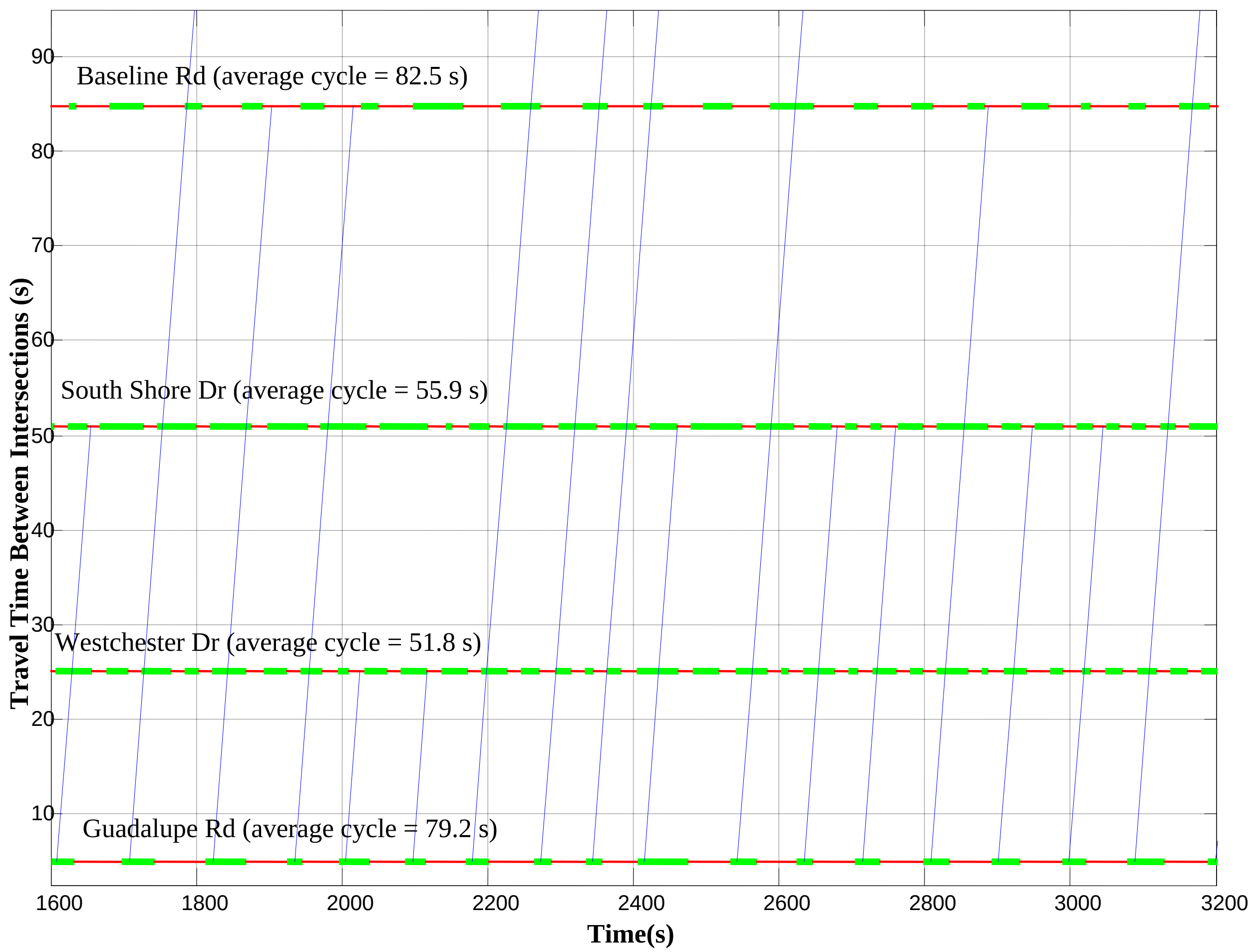


Figure 3: Signal State Changes and Green Bands for the Critical Arterial Direction (Northbound, Shown by Diagonal Lines) for four Intersections on Rural Road under Self-Organizing Control Logic

Diagonal lines show the leading edge of green bands at the progression speed in the critical (northbound) direction; one can also see that the opposite direction gets reasonable progression. The frequency and extent of green bands show how the self-organizing method leads to partial synchronization while allowing less busy intersections to cycle at a faster pace.

Conclusions

Adding rules for secondary extensions and for dynamic coordination of short zones of closely spaced intersections appears to create the coordination mechanisms needed to make decentralized, actuated signal control achieve a degree of synchronization that leads to efficient traffic control on arterials, while preserving the benefits of shorter cycles and responsiveness to traffic changes, and flexibility for applying aggressive transit priority and for recovering from signal priority interruptions. Self-organizing logic offers a new paradigm for arterial traffic signal control that can help overcome the limitations of fixed cycle coordination.

Simulation results indicated that self-organizing logic yielded considerably less average vehicle delay (reduction of approximately 14% and 7% along Beacon Street and Rural Road, respectively). Providing aggressive signal priority for transit along Beacon Street reduced delay by more than 10 seconds per train per intersection (50% reduction in transit delay) with negligible impacts to non-priority traffic (average delay increased by approximately 3% compared to the optimized coordinated-actuated operation).

Acknowledgment

This research is supported by the National Science Foundation under Grant No. CMMI 1130936. The authors also acknowledge the support received from Maricopa Association of Governments and Caliper Corporation.

References

1. Cesme, B., and P. G. Furth, "Self-Organizing Control Logic for Oversaturated Arterials", Submitted for Presentation and Publication to *Transportation Research Board*, 2012.
2. Newell, G. F. and E. E. Osuma. "Properties of Vehicle-Actuated Signals: II Two-Way Streets." *Transportation Science* 3, No. 2, 1969, pp.99-125.
3. Furth, P.G., Cesme, B., and T. Rima, "Signal Priority Near a Major Bus Terminal: Case Study of Boston's Ruggles Station," *Transportation Research Record,* Vol. 2192, 2010, pp. 89-96.
4. Furth, P.G., Cesme, B., and Th.H.J. Muller, "Lost Time and Cycle Length for an Actuated Traffic Signal," *Transportation Research Record,* Vol. 2128, 2009, pp.152-160.
5. Cesme, B. and P.G. Furth, "Multiheadway Gap-Out Logic for Actuated Control on Multilane Approaches," *Transportation Research Record,* 2012.
6. D. I. Robertson and R. D. Bretherton, "Optimizing Networks of Traffic Signals in Real Time - The SCOOT Method," *IEEE Trans. on Vehicular Technology.* Vol. 40, No. 1, pp. 11-15, 1991.
7. Janos M., P.G. Furth, "Bus Priority with Highly Interruptible Traffic Signal Control," *Transportation Research Record,* Vol. 1811, 2002, pp. 157-165.
8. Stevanovic, A., "*Adaptive Traffic Control Systems: Domestic and Foreign Practice, NCHRP Synthesis 403,* Transportation Research Board, Washington, D.C., 2010.
9. Gershenson, C. and D.A. Rosenblueth, "Modeling Self-Organizing Traffic Lights with Elementary Cellular Automata," C3 Report 2009.06, Universidad Nacional Autonoma de Mexico, 2009.
10. Mirchandi, P. and L. Head, "A Real-Time Traffic Signal Control System: Architecture, Algorithms, and Analysis," *Transportation Research Part C: Emerging Technologies,* Vol. 9, No. 6, Dec. 2011, pp. 415-432.
11. Gartner, N. H., "Demand-Responsive Decentralized Urban Traffic Control," *Report DOT/RSPA/DPB-50/81/24*, U.S. Department of Transportation, Washington D.C., 1982.

12. Wadjas, Y., and P. G. Furth, "Transit Signal Priority Along Arterials Using Advanced Detection," *Transportation Research Record*, Vol. 1856, 2003, pp. 220-230.

13. Furth, P.G. and Th.H.J. Muller. "Conditional Bus Priority at Signalized Intersections: Better Service Quality with Less Traffic Disruption, "*Transportation Research Record* 1731, pp. 23-30. 2000.