1.1 Introduction

InSync is an adaptive traffic signal system developed by Rhythm Engineering that utilizes state-of-the-art sensor technology, image processing and artificial intelligence. These elements are integrated into a system that automatically optimizes local traffic signals and coordinates signals along roadway arterials according to real-time traffic demand. The use of InSync eliminates the need for static signal coordination plans. A six- intersection arterial can be fully installed, initially configured and operational in 2-3 days. Within two weeks, after some initial remote monitoring, the arterial can be running in a fully-optimized mode. A low level of maintenance is required after this initial monitoring and inputting of minor reconfigurations. This system is relatively inexpensive; typically equal to or slightly above the cost of a high-quality video detection system.

1.2 Hardware Overview—Cameras, Processor, Detector Cards, Ethernet Communications

InSync utilizes high-end *IP digital cameras* in weatherproof enclosures that are mounted normally on mast arms with standard brackets. The cameras are connected to an *InSync processor* installed within each local traffic cabinet through a CAT-5 Ethernet cable and a 24-Volt electrical wire that provides power. The InSync processor is placed in the local traffic cabinet and interfaces with the local signal controller using *detector cards* that are plugged into existing detector card racks. Other system hardware includes a 110/24-Volt transformer, surge protectors, an un-managed Ethernet switch and a pigtail cable for red/green returns to be fed back into the processor from the controller's leads. Except for an I/O board within the processor and the associated detector cards that are each required to communicate with the various proprietary controllers, the system uses "off-the-shelf" components. In order for arterial coordination to take place, *Ethernet communications* must exist between the networked intersections. Since InSync uses distributed network architecture, an unlimited number of signalized intersections may be coordinated.

1.3 Configuration Procedures—Standard Web Interface

InSync is both Ethernet and web-centric in its functionalities. Each processor and every camera has an IP address. These components can be accessed directly via the network without any proprietary software. All the necessary configurations of and any software upgrades to the system software can be accomplished remotely over the Ethernet network. The onsite cameras are properly aimed, zoomed, focused and tightened to effectively view vehicles arriving at and progressing through a traffic signal. A web page associated with the InSync system is accessed via a standard Internet browser (Internet Explorer or Firefox) that leads a user through the process of drawing all the necessary detection, count and contrast zones that quantify the traffic data generated by approaching vehicles. It also provides a dropdown menu page for all the adaptive system parameters.

2.1 Sensor Functionalities—Real Time Measurement/Reporting of Queue Lengths & Percent of Occupancy for All Approaches

The video/data collection sensors (the IP cameras) capture and communicate real-time images of vehicles approaching an intersection to the InSync processors. The processor reads and interprets these images for its optimization processes. This kind of image tracking provides a sufficient estimation of real-time queue lengths and the percentage of occupancy of each lane/approach for optimization purposes. Advanced detection is not essential to create an effective traffic-adaptive dynamic, though it can be incorporated seamlessly into the system. These data are updated by the processor every second. (Similar traffic data could also be input using other kinds of sensors.)

2.2 Optimization Techniques—Global Coordination; Local Intelligent Actuation

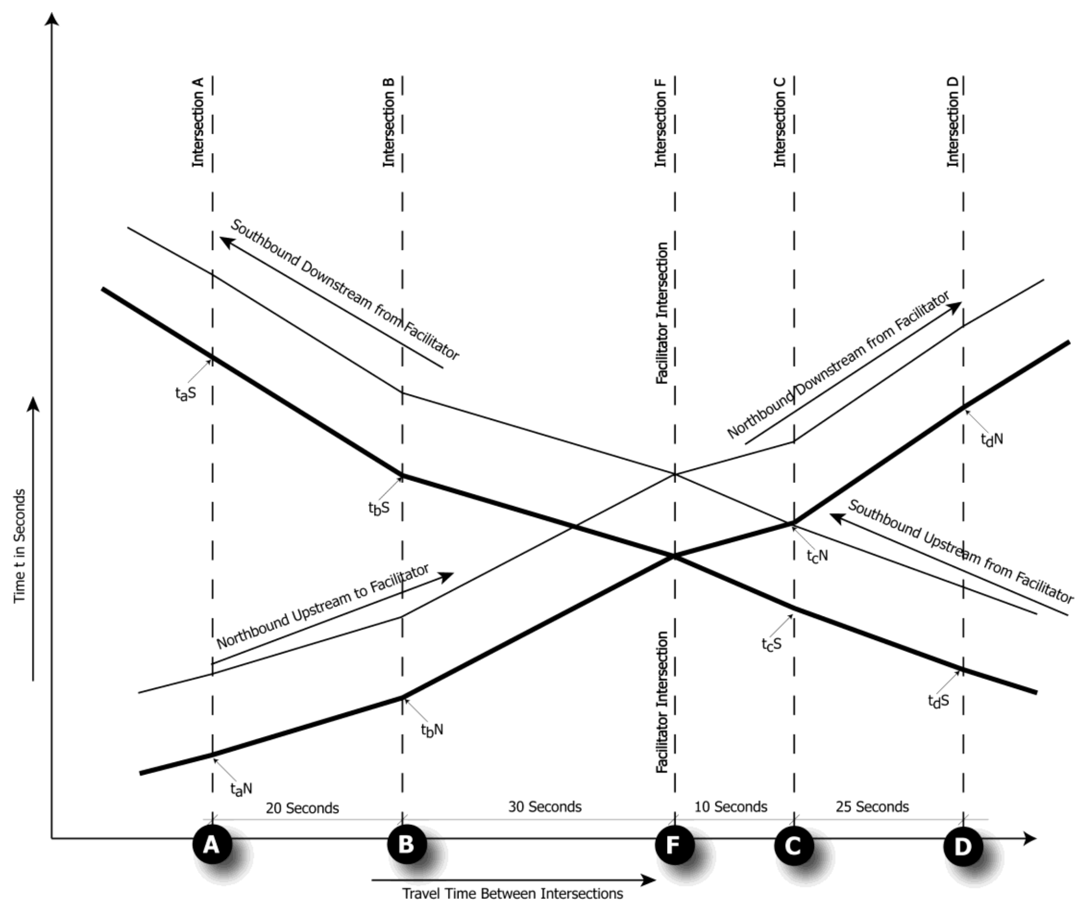
There are two aspects to InSync's signal optimization that deal with the tensions between providing progression of platoons of vehicles along a main arterial and the clearance of vehicles involved with secondary traffic movements within the grid: the global and the local. (InSync operates and optimizes signals within the minimum/maximum parameters that users have input to the initial configuration settings of the system.)

The Global Element: Time Tunnels and Adjustable Periods for Optimizing Progression

Users must determine the main directions within the grid, but can also redefine and automatically toggle between arterials by TOD/DOW. Special parameters can be set in for intersecting main arterials that provides effective coordination within the grid.

Time Tunnels: Green waves/time tunnels are guaranteed by successively turning each light green, at the expected arrival time of vehicles from upstream intersections. This can be illustrated using speed lines. Speed lines, are configured starting with a chosen facilitator intersection. By default, the speed lines for the main two directions of travel intersect at this facilitator intersection. Time tunnels are made to occur at this intersection by requiring the simultaneous initialization of green lights for both directions. The facilitator intersection, decides a time at which it will serve a green band for the coordinated tunnel phases and communicates that time with the adjacent intersections. Each adjacent intersection uses the expected vehicle travel time between it and the intersection it received the tunnel message from to decide when it needs to turn the tunnel phases green for both its downstream tunnel phase from the facilitator intersection and its upstream tunnel phase to the facilitator intersection. Start times for downstream tunnel phases *to* the facilitator (or upstream *from* the facilitator) at the adjacent intersection are adjusted by (-travel_time) so that vehicles are released from the upstream intersection in time to reach the facilitator intersection when it initiates its tunnel time. Start times for downstream tunnel phases *from* the facilitator (or upstream *to* the facilitator) at the adjacent intersection are adjusted by (+travel_time) so that vehicles are released from the downstream intersection as they arrive from the facilitator intersection.

Using the travel times between each adjacent intersection, these tunnel times are calculated by each intersection as it receives the dynamic tunnel phase timing messages from each adjacent intersection along the artery. Visually, the speeds lines for an artery with two main directions of North Bound and South Bound with four "listener" intersections and one facilitator intersection in the middle would look like this:



Expected travel times between intersections are listed on the x-axis, and elapsed time is the y-axis. The slope of a speed line is always 1, since the expected travel time between intersections has a 1 to 1 relationship with wall time (or time on a clock).

Tunnel start times at each intersection for the North and South phases are relative to the tunnel start times at the facilitator. If we say that the North and South tunnel phases start at t=0 at the facilitator intersection, then tcN (the time of the start of the North Bound downstream tunnel phase at Intersection C) = 0 + travel_time_to_facilitator = 0 + 10 = +10. This says that Intersection C must force a green light for the North Bound phase 10 sec. after the facilitator intersection does. tcS (the time of the start of the South Bound upstream tunnel phase at Intersection C) = 0 - travel_time_to_facilitator = 0 - 10 = -10. This says that Intersection C must force a green light for the South Bound phase, 10 sec. before the facilitator intersection does.

To complete the example, the other start times at Intersections A, B, and D are:
tdN: +35
tdS: -35
tbS: +30
tbN: -30
taS: +50
taN: -50

Additionally, the speed lines are not actually required to intersect at the facilitator, or even intersect at an intersection. The tunnels can be offset to allow any arrangement of speed lines as desired. This flexibility can provide for more efficient progression. Also, travel times for both directions between adjacent intersections do not have to be the same. The South Bound direction may take 5 seconds longer to travel than the North Bound direction, between 2 intersections. It is for simplicity that we start with the above configuration. The minimum tunnel bandwidth (or guaranteed green time for a tunnel phase) is configurable.

Vehicles traveling along the main arterial that arrive at an intersection at the beginning of the time tunnel ideally progress unstopped all the way through the coordinated arterial. InSync automatically extends green lights beyond the set parameter if it observes that the moving platoon has not sufficiently gapped out at a user changeable percentage of occupancy (calculated every second by InSync) or set gap time. InSync may, if permitted by user, provide a green light along the main street before the light is guaranteed to be green. Real time traffic data are continually passed to downstream processors by their upstream partners that can also factored into the optimizing process.

Intelligently Adjustable Periods: Through its global interactive communications with the intelligent processors at each intersection (also see next section: The Local Element) the facilitator will, if necessary and as soon as possible, expand or contract the time between tunnels to provide the optimal period lengths to serve each phase along the arterial. At the end of every period, each local processor is "polled" by the facilitator and "reports in" if it needs more time, the same time or less time to clear is local phases. These period length adjustments serve to efficiently progress vehicles and clear out queues, both globally and locally.

The Local Element: Logic and Features of the Optimization Algorithm

Beyond the constraints communicated by the facilitator as "tunnel messages" that guarantee coordinated green lights for the main arterial, the signals operate in "intelligent fully-actuated" mode. The time between tunnels is called a period. If a period is 90 seconds in duration and a green light is guaranteed for the main directions at each intersection for 10 seconds, then 80 seconds are available for the local optimizer to schedule states (phase pairs) at each intersection according to its intelligent scheduling. The local optimizer embodies the dominant logic and algorithm of the adaptive capacities of the system.

Scheduling of States: There are three main factors the optimizer considers in its scheduling logic:

1) If it is close to the initiation of a new tunnel, it will schedule a main street sequence of states. This sequence of main states is only allowed to be scheduled such that after it completes, there is sufficient time to schedule a sequence of cross street states. If the main direction requires a leading left turn, its clearance time is also included in the calculation for time needed on the cross street.
2) If a tunnel has recently ended, it will schedule a cross street sequence as its priority. The amount of time needed for the cross street is based on a balance between the actual amount of clearance time needed and anticipated time needed. If there are no cross street queues it will schedule a miscellaneous main state.
3) A miscellaneous main state is scheduled for phases with queues that have been waiting the longest. Wait times being equal; the phase with the largest queue is scheduled. Any available miscellaneous time is used to schedule any phase with real-time demand including protective permissive left turns on the main directions.

Empty Queues: In the calculation of a state sequence, every vehicle phase is assumed to have a queue of 1 vehicle, if no queue exists. These phases typically find their way into the states scheduled toward the end of the state sequence. This way, if vehicles do arrive on these previously empty phases, they can be served. In this sense, those states act as place holders for vehicles which may arrive. If a phase remains empty when it becomes time to serve that state, that state is either removed or modified to contain phases which do contain a queue of vehicles.

Duration of States: After an initial sequence of states is scheduled, the durations of each state are continually modified to contain enough clearance time to serve vehicles which may have arrived after the state was initially scheduled. As each previous state in the sequence reaches extension time, all states scheduled after adjust their initial durations for newly arrived vehicles. Adjusted durations and extensions are limited by the amount of time left to serve vehicles on pending phases. For a state with two phases, *ph1* and *ph2*, if *ph1* clears out, a phase concurrent with *ph2* that has a queue can be put in place of *ph1*, assuming there is enough time remaining in the scheduled state to fulfill the new phase's minimum green time, amber time, and red time.

Termination of States: States being served are ended or dynamically modified as phases are terminated. Termination of a phase occurs in one of two ways. Both of these methods employ a model of linear change with time, such that a phase will terminate at a higher percentage of capacity as time increases or the phase's gap time will decrease to a point at which, once presence for a detection zone is lost, the phase will terminate, as time increases.

Calculating a Sequence: The optimizer gathers the data for the calculation: current queues, pending pedestrian calls, and any upcoming plans for tunnels. These requirements are converted into restrictions on the beginning and ending times of green lights for phases. A queue is converted into a minimum clearance time: the ending time minus the beginning time for the phase must be at least equal to the time to clear the queue plus the change time required for that light to turn green; similarly for a pedestrian call, except that the clearance time is the time required for pedestrians to walk across the intersection. A plan for a tunnel is converted into restrictions on the beginning and ending times for the phase: the beginning time must be less than or equal to the beginning time of the tunnel minus the change time, and the ending time must be greater than or equal to the ending time of the tunnel. The optimizer considers each permitted sequence as a sequence of transition times at which phases begin and end, with the restrictions transformed into inequalities in these transition times. For each sequence, a minimum-total-time solution satisfying these restrictions and the total waiting time for queues are calculated. The sequence with the least total waiting time is chosen. If no solution satisfying the restrictions exists for any permitted sequence, then the lowest-priority of the restrictions are relaxed until a solution is possible: first, queues on permissive left turns are transferred to their adjacent through movements, then the queues with the least waiting times have their clearance times reduced, then finally only the plans for tunnels are considered. Once a solution is obtained, the times are translated into a schedule of states and the first state of the schedule is initiated.

Early Release: Sometimes, it is useful to prevent cars from leaving an intersection too early and collecting at a downstream intersection, either because there is a limited amount of space available for cars to queue, or to absolutely ensure that the downstream light is green when they arrive. This is sometimes called metering. To handle these situations, the optimizer has the configuration option of restricting early release of a tunnel phase at an intersection.

Period Length Evaluation: During local optimization, the intersection continually analyzes its queue lengths and percentage of occupancy for each phase. If the intersection determines that it has not been given enough overall time to adequately clear out it queues (drop the percentage of occupancy to a desired threshold), the intersection reports this to the facilitator (see previous section: The Global Element). This method of adjusting the period is reactionary. To be more proactive, each intersection constantly analyzes the flow rate of vehicles in each phase. If the current flow rates over the past 15 minute durations are comparable to the historically logged flow rate, then it is assumed that the current pattern of traffic can be served using a period that was adequate enough to serve the historical pattern of traffic. This process is known as period prediction and is communicated to the facilitator intersection along with the real-time load analysis of the local intersection. The facilitator intersection uses all of this information to determine if/how it will adjust the current period.

2.3 Interface Methodology—How InSync Determines/Inputs Optimized Calls to Signal Controllers

InSync is a plug-in technology that interfaces with all existing traffic signal controller/cabinet architectures. It controls traffic signals by submitting calls to the traffic controller through detector cards, just as inductive loops do. However, InSync only allows one phase pair at a time to be input to the controller by filtering, prioritizing and suppressing the demands generated by the detection of vehicles that are approaching an intersection in real time. Pedestrian calls are also filtered by InSync and are permitted at the times deemed optimal by InSync's real-time coordination. InSync's calls are passive in that InSync will yield to any higher priority calls that are directly communicated by users' choice into the controllers: i.e. preemptions or central system software priorities. In these cases, InSync will continue to serve as a detection device and then revert to its optimization mode when the controller begins to respond to its calls again. It can also be configured to toggle automatically between detection mode and optimization mode by TOD/DOW if users desire to utilize pre-determined timing plans.

2.4 Digital Signal Control Concepts—Finite Number of Signal States; No Transition

InSync does away with set cycle lengths, set splits and offsets to a fixed point in the cycle that have been traditionally considered essential for signal coordination. These concepts are germane to a linear/analog approach to signal coordination. InSync is an artificially-intelligent/digitally-based finite state changing machine. By its method of externally influencing a controller, it causes any controller to effectively function digitally. This digitization does not refer to the nature of the component parts of the controller, but rather, a "digital methodology" of how traffic signal phases are chosen. In relation to traffic movements, there is a maximum of 16 possible sequences of phase pairs (states) at any quad intersection. Since it knows the real time traffic demand, InSync is able to instantaneously select and input to the controller any user permitted phase pair associated with these 16 sequences that it deems optimal. InSync only needs to decide: 1. Optimal Sequence. 2. When to initiate a state (phase pair). 3. Duration of that state. It is not limited in its choices or their duration times by a set cycle length, split or offset. Except for minimums and maximums and 1 second passage times, all typical volume density inputs to the controller are disabled so that it runs in free mode. This permits the controller to quickly react to and change the traffic signal according to the optimized calls coming through InSync.

Another important advantage of this "state-changing" architecture and methodology of signal optimization is that the traffic flow disruption caused by the transition from one static timing plan to the next, or by preemptions, is eliminated.

2.5 Failure Mitigation

When a sensor is placed in emergency/fog mode, InSync will access 4-weeks of historic green split data for specific TOD/DOW at that particular approach. This data is normalized into a split time to put in to the controller until the sensor is functioning again. A call is issued for every phase for at least a minimum split time. This happens in the following cases: 1) A camera fails to talk to a video processing detector subsystem. 2) The video processor determines the view is not sufficiently clear. 3) The processor doesn't hear from a particular detector subsystem. A text alert is seen on the video image when a sensor is in emergency/fog mode.

Calls on all phases are automatically input when 1) InSync determines that the detected traffic is significantly lower than historical averages, which indicates a sensor failure. 2) The I/O board fails to hear from the processor for two seconds. 3) A detector card fails to hear from the I/O board for two seconds.

If communications between networked intersections fail, individual processors will continue to perform local optimization functions.