Smart Infrastructure for Future Urban Mobility

Stephen F. Smith

■ *Real-time traffic signal control presents a challenging multiagent planning problem, particularly in urban road networks where, unlike simpler arterial settings, there are competing dominant traffic flows that shift through the day. Further complicating matters, urban environments require attention to multimodal traffic flows (vehicles, pedestrians, bicyclists, buses) that move at different speeds and may be given different priorities. For the past several years, my research group has been developing and refining a real-time, adaptive traffic signal control system to address these challenges, referred to as* scalable urban traffic control (Surtrac). *Combining principles from automated planning and scheduling, multiagent systems, and traffic theory, Surtrac treats traffic signal control as a decentralized online planning process. In operation, each intersection repeatedly generates and executes (in rolling horizon fashion) signal-timing plans that optimize the movement of currently sensed approaching traffic through the intersection. Each time a new plan is produced (nominally every couple of seconds), the intersection communicates to its downstream neighbors what traffic it expects to send their way, allowing intersections to construct longer horizon plans and achieve coordinated behavior. Initial evaluation of Surtrac in the field has demonstrated significant performance improvements, and the technology is now deployed and operating in several U.S. cities. More recent work has focused on integrating real-time adaptive signal control with emerging connected vehicle technology, and exploration of the opportunities for enhanced mobility that direct vehicle (or pedestrian) to infrastructure communication can provide. Current technology development efforts center on vehicle route sharing, smart transit priority, safe intersection crossing for pedestrians with disabilities, real-time incident detection, and integrated optimization of signal control and route choice decisions. This article provides an overview of this overall research effort.*

Mobility is key to quality of life, equity of opportunity, and economic growth in urban environments. When transportation infrastructure operates efficiently and people are able to move freely, cities thrive. The major deterrent to urban mobility is traffic congestion. It is estimated that congestion costs residents of U.S. cities $160,000,000,000 in lost time and fuel costs, while pumping an additional 50,000,000,000 tons of CO_2 into the atmosphere (Texas A&M Transportation Institute 2018). To make matters worse, people are increasingly moving to cities. The number of people living in urban areas is expected to grow from fifty percent of the world's population currently to sixty-eight percent by 2050 (United Nations, Department of Economic and Social Affairs 2019). Unfortunately, the traditional traffic engineering approach to mitigating congestion, that of building more road capacity, is not typically possible in urban environments due to land-use issues, geographical constraints, and so forth that limit space for expansion. Instead, policies for reducing the number of vehicles on the road and increasing reliance on mass transit systems must be combined with mechanisms and technologies for increasing the efficiency of surface street traffic flows. In this article, we focus on this latter issue.

One key reason for poor traffic flows on urban surface streets, and hence one major cause of congestion today, is poorly timed traffic signals. By and large, traffic signal control in urban road networks has changed surprisingly little over the past fifty years. Traffic signal timing plans are generally preprogrammed in advance, based on snapshot assessments of average traffic conditions by traffic engineers. Actual traffic conditions can vary significantly from such average assessments, and, in any event, traffic patterns change over time as neighborhoods evolve. There is continual cost in keeping preprogrammed traffic signals up to date, and frequently municipalities do not have the funds to maintain them. Although more recent technological advances in sensing (for example, aboveground detection systems such as video cameras and radar) are often adopted by cities as they refresh their traffic signal infrastructure, these sensors still tend to be used in the same mundane ways that older, in-ground induction loop technology was previously used — to detect and react to vehicles that are already stopped at the intersection rather than looking ahead and anticipating the arrival of approaching vehicles. Despite the advances in intelligent systems and machine learning technologies in recent years, the traffic signal at the typical urban intersection remains an extremely unintelligent decision-making system.

For the past several years, my research group has been developing and refining an approach to real-time, adaptive traffic signal control aimed specifically at urban road networks. Real-time traffic signal control presents a challenging multiagent planning problem in this setting, where, unlike simpler suburban corridors, there are competing dominant traffic flows that shift through the day. Further complicating matters, urban environments require attention to multimodal traffic flows (vehicles, pedestrians, bicyclists, buses) that move at different speeds and may be given different priorities. Our approach to this problem is embodied in a real-time, adaptive traffic signal system called *scalable urban traffic control* (Surtrac). Combining principles from automated planning and scheduling, multiagent systems, and traffic theory, Surtrac treats traffic signal control as a decentralized online planning process. In operation, each intersection repeatedly generates and executes (in rolling horizon fashion) signal-timing plans that optimize the movement of currently sensed approaching traffic through the intersection. Each time a new plan is produced (nominally every couple of seconds), the intersection communicates to its downstream neighbors what traffic it expects to send their way, allowing intersections to construct longer horizon plans and achieve coordinated behavior. The Surtrac system has proved to be quite effective in practice, reducing travel times through controlled networks by twenty to twenty-five percent, and number of stops and wait times by thirty to forty percent. The system is now deployed and operating in several U.S. cities, including Pittsburgh, Pennsylvania; Atlanta, Georgia;

Portland, Maine; and Quincy, Massachusetts. More recent work has focused on integrating real-time adaptive signal control with emerging connected vehicle technology, and exploration of the opportunities for enhanced mobility that direct vehicle-to-infrastructure (V2I) and pedestrian-to-infrastructure (P2I) communication can provide. Current technology development efforts center on vehicle route sharing, smart transit priority, safe intersection crossing for pedestrians with disabilities, real-time incident detection, and integrated optimization of signal control and route-choice decisions. In this article, we provide an overview of our overall research effort and vision of future urban mobility, summarize the results we have obtained to date, and highlight the challenges that remain.

The remainder of the article is organized as follows. I will first introduce the traffic signal control problem and review past work toward its solution. Next, this problem is formulated as an online, collaborative planning process, and describes the algorithms underlying the Surtrac approach. I then summarize experimental results obtained with Surtrac in the field and current deployment progress, and then turn attention to more recent work aimed at boosting Surtrac performance through direct V2I and P2I communication of real-time vehicle information. The article concludes with a discussion of additional opportunities for enhancing future mobility and the technical challenges that remain.

The Traffic Signal Control Problem

Contemporary signalized intersections are controlled by a signal timing plan, which dictates in what order and for how long various movements through the intersection get the green signal. Consider the basic intersection depicted in figure 1, which identifies eight individual vehicle movements through the intersection. Typically, compatible movements are grouped into movement phases (for example, "phase 2,6 pairs east and west traffic flows"), and the intersection's timing plan is then specified as a sequence of phases. Although not necessarily the case, the ordering of phases is more often than not kept fixed from cycle to cycle (as is shown in the sample timing plan of figure 1), and we will make this assumption in the work to be presented. With this assumption, the crux of the problem of generating a signal timing plan is to determine how much time to allocate to each phase (with a zero duration indicating that the corresponding phase will be skipped). As mentioned earlier, conventional signal systems use fixed, preprogrammed timing plans, sometimes employing simple actuation to trigger or skip minor phases (for example, left turn signals) based on the presence or absence of a queue. Adaptive traffic control systems, alternatively, sense approaching traffic flows and dynamically adjust timing plans over time.

Adaptive traffic signal systems have been around since the 1980s. The most successful have been systems that adapt through adjustment of timing plan
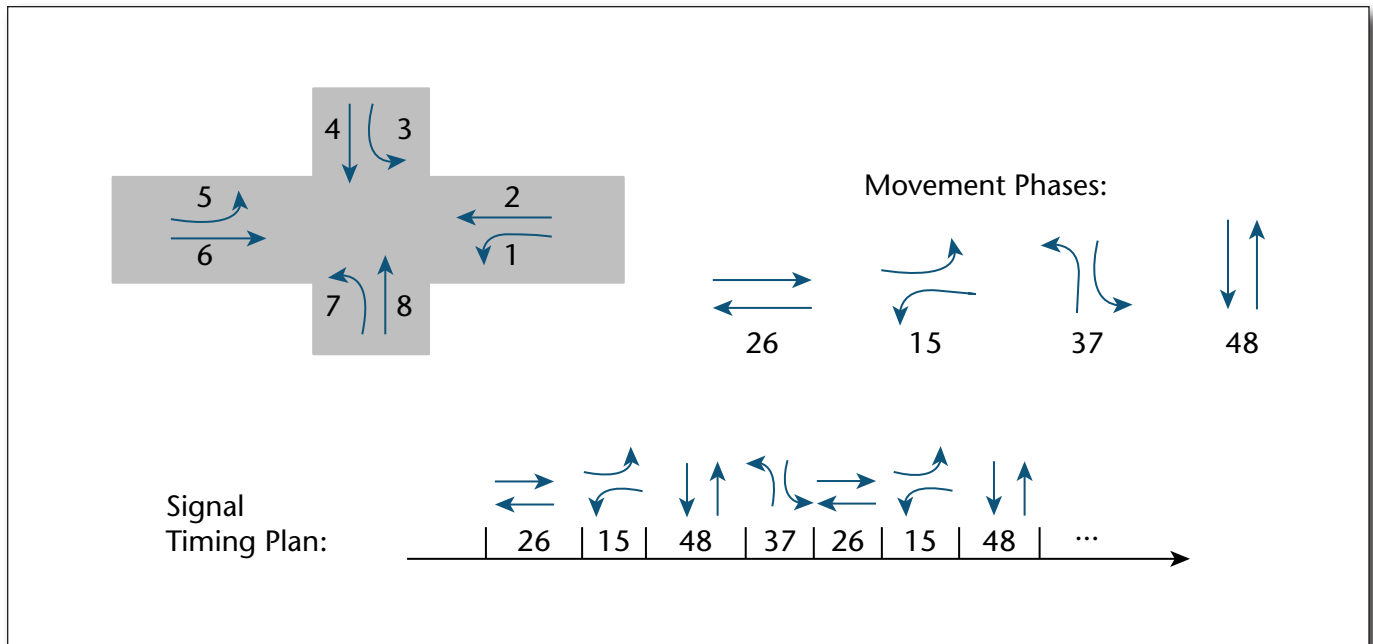
Figure 1. A Sample Signal Timing Plan.

Adapted with permission from Sen and Head (1997).

parameters over time, including the splits (or proportion of time given to each phase in the cycle), the cycle length, and the offsets between the cycle starts of neighboring intersections. These systems include the Split Cycle Offset Optimization Technique (Robertson and Bretherton 1991) and the Sydney Co-ordinated Adaptive Traffic System (Lowrie 1992), two of the earliest commercial adaptive signal systems, and Adaptive Control System-Lite (Luyanda et al. 2003), a more recent Federal Highways Administration-sponsored effort to create a lightweight adaptive signal system alternative. These systems operate by integrating traffic flow information from sensors at or before the intersection over time, and then periodically updating timing plan parameters. They tend to operate well in circumstances where traffic flows evolve smoothly and continuously, but because they tend to centralize decision-making and rely on past traffic flow information to make adjustments for future traffic, they can be limiting with respect to real-time response to deviations in actual conditions.

Another class of adaptive signal system, referred to here as *online intersection planning*, does focus more directly on the real-time traffic signal control problem and extends naturally to network-level coordination. Prototypical examples of this approach include Real-Time Hierarchical Optimized Distributed Effective System (Mirchandani and Head 2001; Shelby 2001). These systems place more emphasis on decentralized decision-making at each intersection with peer-to-peer communication, sometimes with additional coordination through hierarchical reasoning.

Historically, these approaches have been limited by the computational requirements associated with online intersection planning, which have significantly restricted either the planning horizon or the temporal precision of the timing plan. In the work presented here, we adopt this same online planning perspective, but introduce a new intersection scheduling formulation that overcomes this limitation.

Yet another approach to adaptive signal control that has emerged in recent years and has had particular success in suburban corridor settings, involves use of a shared global plan. If a single dominate traffic flow can be preidentified, then an optimized fixed plan can be developed for this flow in advance, reducing the adaptive signal control problem to that of managing the side-street traffic. The In-Sync system (Stevanovic and Zlatkovic 2011) is based on this philosophy. However, this approach is less suitable for more complex urban road networks where there are competing dominant flows.

The artificial intelligence (AI) research community has also proposed several approaches to the adaptive traffic signal control problem. Early work by Bazzan (2005, 2008) focused on multiagent systems approaches, emphasizing both distributed constraint optimization and multiagent reinforcement learning approaches. See Bazzan and Klügl (2013) for a general review of research in this area. Most recently, a multiagent reinforcement learning system called Multiagent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (El-Tantawy, Abdulhai, and Abdelgawad 2013) has

demonstrated significant improvement when applied to a simulated network of fifty-nine intersections in the city of Toronto. From a real-time adaptive perspective, however, one difficulty of reinforcement learning approaches is the time that is required to adapt. In circumstances where traffic patterns are transient (for example, the emptying of parking garages after a sporting event), the system may not be sufficiently reactive. Another stream of AI research has introduced a class of reservation-based approaches that could be applicable in a future, connected autonomous vehicle world (Dresner and Stone 2008). Under these schemes, approaching vehicles communicate with the intersection to request space and time for crossing the intersection, and the intersection reserves slots so as to guarantee safe passage. Recent work in this paradigm (Stevanovic and Mitrovic 2018) has shown the potential benefits of giving up classic assumptions about unidirectional traffic lanes. A major challenge for this approach is realizing a world where all intersection stakeholders, including pedestrians and bicyclists, are fully connected with the coordination infrastructure. The work described herein also anticipates a connected and increasingly autonomous world, but our focus is on nearer-term scenarios where connectivity and autonomous vehicles are not yet ubiquitous, and the need for physical traffic signals remains.

Other AI research has explored the benefits of new formulations of the traffic signal control problem. One effort by Guilliard et al. (2016) has developed a novel mixed-integer linear program formulation that significantly extends the scalability and practicality of classic models. Another has examined the effectiveness of Planning Domain Definition Language-based modeling and planning of traffic signals in response to detected congestion conditions (Vallati et al. 2016). These works, however, are more strategic and do not address the real-time control problem of interest here.

Finally, there are an increasing number of big-data approaches to the real-time traffic signal control problem. In Baluja, Covell, and Sukthankar (2017), for example, microauctions are combined with machine learning techniques to produce signal timing plans that capture real-time conditions. Alternatively, Didi, the Chinese ride hailing company, is working to use the data it is collecting from its drivers to regularly compute and install fixed timing plans that reflect recent traffic flow data. From a real-time control perspective, however, these approaches continue to be limited by their reliance on models of past observed traffic data to control current traffic flows.

Surtrac: Traffic Control as Decentralized, Online Planning

The Surtrac system that we have developed (Xie, Smith, and Barlow 2012; Xie et al. 2012; Smith et al. 2013) is designed specifically for complex urban road networks, where there are multiple, competing dominant flows that change through the day. To emphasize responsiveness to real-time events, it takes a decentralized, online planning approach. Each intersection continually senses its approaching traffic and optimizes its own local traffic flow. Intersections communicate flow information with their neighbors to achieve coordinated behavior at the network level.

Figure 2 illustrates the basic concept of operations. We put a computer running the Surtrac system at every intersection. At the beginning of each planning cycle, each intersection pulls a snapshot of its approaching traffic flows from its local sensors (video cameras, radar, etc.) and develops a predictive model of when traffic in various approaching directions is expected to arrive at the intersection. Based on this predictive model, the system generates, in real time, a signal timing plan that optimizes the movement of approaching traffic through the intersection. Once the timing plan is generated, the system begins to execute it, sending a command to the controller (the hardware device that actually controls the signals) to either extend the current phase or switch to the next phase. The system also communicates the traffic it expects to be sending to its downstream neighbors. Downstream intersections are doing the same thing, generating their own local timing plan, but now, in addition to the approaching traffic that they can see through their local sensors, they have an expectation of what traffic is coming down the pike behind them, which allows them to generate a longer horizon plan. Each intersection asynchronously initiates a new planning cycle every second or two.

There are several basic advantages to this real-time approach to traffic signal control. First and foremost, traffic signals are optimized for the actual traffic on the road at any point in time. Second, the approach is designed for optimization of grids and other complex urban road networks, with suburban corridors being handled as a special case. Third, the predictive model that is generated can be weighted according to traffic mode (for example, passenger vehicle, bus, pedestrian, bicyclist), and, therefore, Surtrac-generated timing plans can reflect multimodal optimization criteria. Finally, because the system is decentralized, there is no centralized planning bottleneck, and the system is inherently extensible to city scale. Moreover, the system's decentralized framework also promotes incremental deployment, which allows cities to spread their infrastructure investments over time.

Schedule-Driven Traffic Control

Two key technical ideas underlie the Surtrac approach to real-time traffic control. The first is a novel formulation of the intersection control problem as a special kind of single machine scheduling problem. An aggregate representation of approaching traffic is used to identify the input jobs that must be sequenced through the intersection to produce the timing plan, providing an abstraction that permits efficient solution while retaining the nonlinear nature of traffic flows. The second is the idea of communicating planned outflows to downstream neighbors,
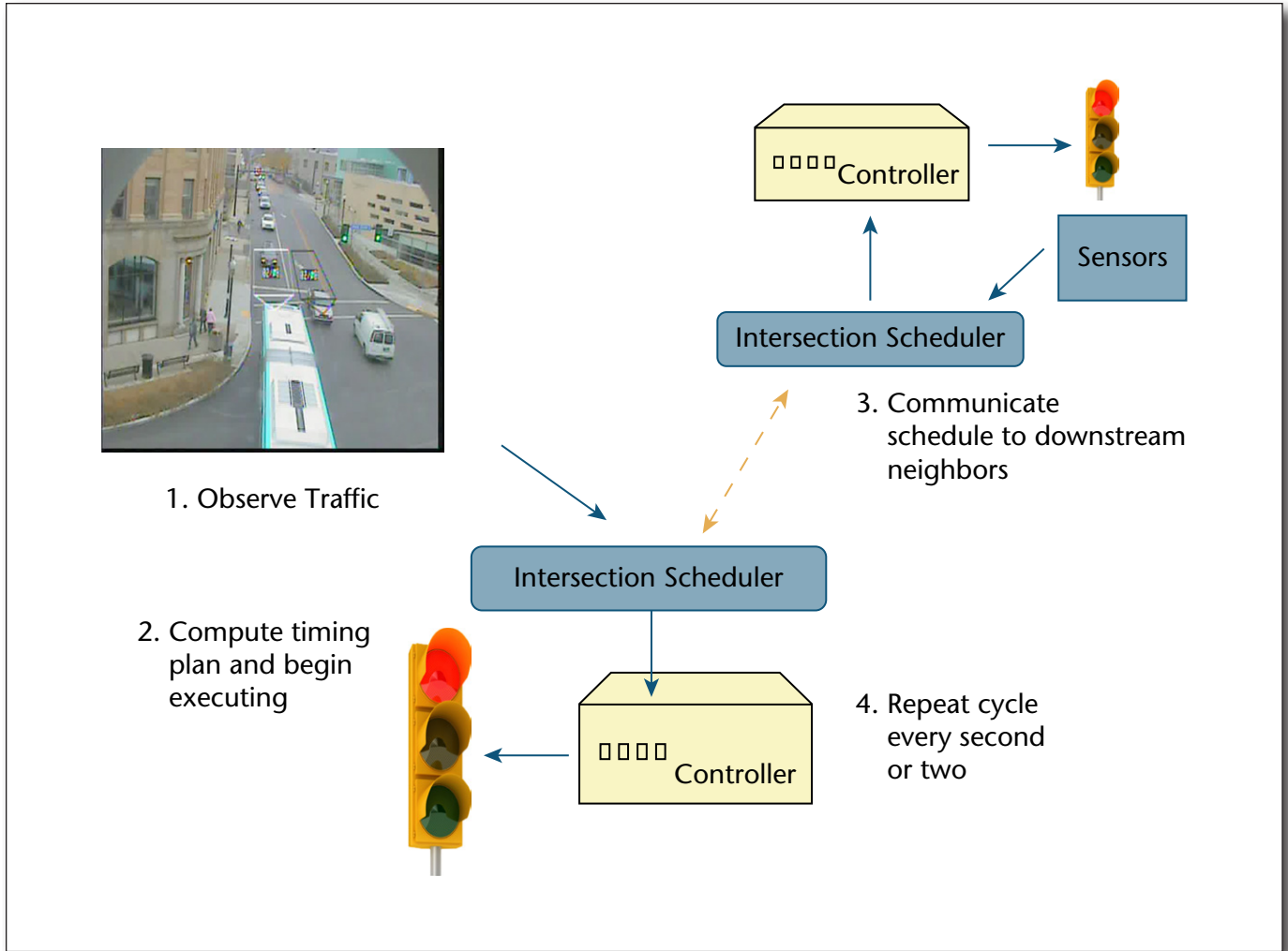
Figure 2. Surtrac Concept of Operation.

to provide visibility of likely future input jobs and enable generation of longer horizon timing plans. Although intersections only communicate with their direct neighbors, this information can flow multiple hops over repeated planning cycles.

Figure 3 illustrates the abstract representation of approaching traffic used by Surtrac's schedule-driven approach in more detail. Sensor data are used to generate an ordered sequence of pairs < arr_j, dep_j > for each approaching road segment, indicating the projected arrival and departure times at the intersection for each currently sensed vehicle (shown in figure 3a). These input data streams are then aggregated into larger vehicle clusters (platoons and queues) based on proximity (as shown in figure 3b and c). More precisely, each sequence of vehicle arrival and departure time pairs is transformed into a sequence of triples < arr_j, dep_j, cap_j >, where arr_j is the arrival time of the first vehicle in cluster j, dep_j is the departure time of the last vehicle in cluster j, and cap_j is the number of vehicles in cluster j.

Given this predictive model of current traffic flows, the problem of generating an intersection timing plan becomes one of merging all traffic cluster sequences approaching the intersection from different directions into a single sequence, subject to temporal constraints relating to fairness (that is, minimum and maximum bounds on phase durations) and safety (that is, fixed yellow and all red periods between each phase). Consider the simple example in figure 4 where the problem involves an intersection with two competing signal phases. Treating this problem as a single machine scheduling problem with non-pre-emptable jobs (that is, indivisible clusters), figure 4 shows a feasible solution, from which a specific intersection timing plan can be extracted. The basic objective is to minimize cumulative delay of all input jobs. More precisely, the intersection scheduling problem is formulated as

$$Minimize \sum_{c \in C} \left(exit_c - dep_c \right)$$

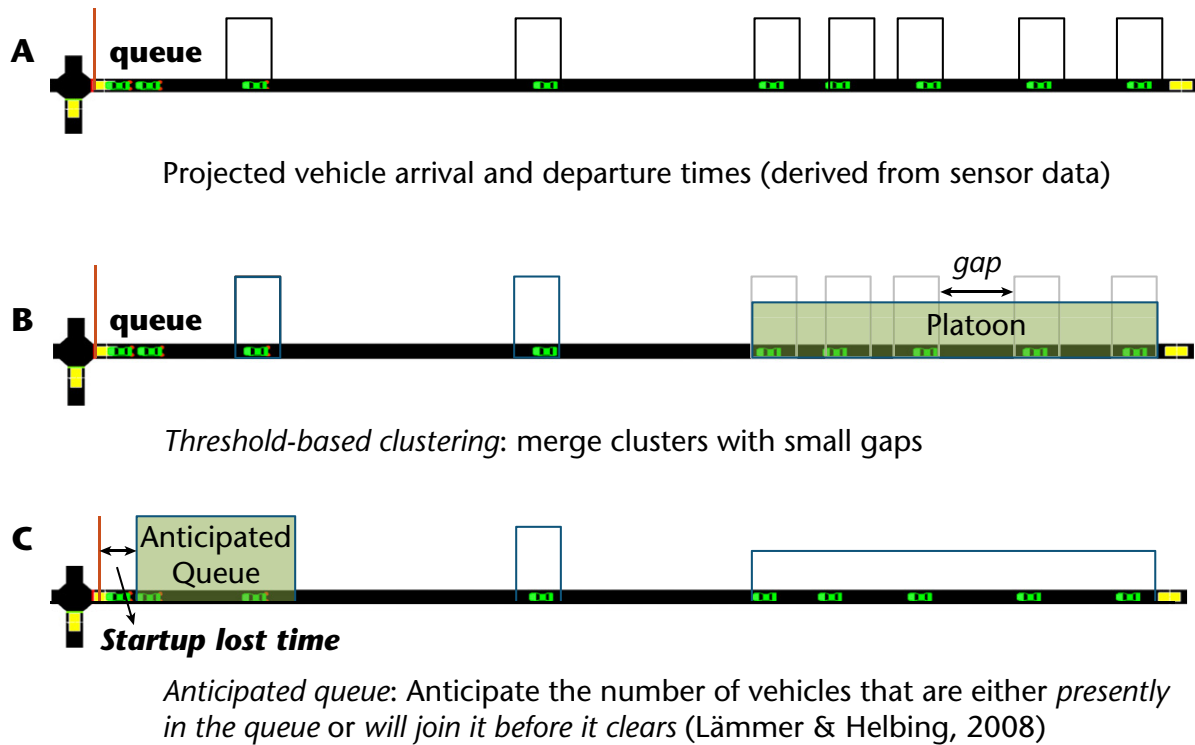*Figure 3. Abstract Representation of Traffic Flow as Vehicle Clusters.*

such that for each phase j, $G^{min}_j \leq duration_j \leq G^{max}_j$ and for any two consecutive phases i and j in the sequence, $end_i + AllRed \leq start_j$, where C is the set of clusters in the schedule, $exit_c$ is the time that cluster c exits the intersection according to the schedule, G^{min}_j and G^{max}_j are phase minimum and maximum times, and *AllRed* is the safety time required between phases.

Surtrac solves this intersection scheduling problem via forward dynamic programming (DP) search. At each stage of the search, a new job (cluster) is added to the set of partial solution sequences, and checks are performed to eliminate dominated solutions (that is, partial solutions in the same phase, with the same included clusters and higher accumulated delay). Two versions of the algorithm have been formulated and tested, one that guarantees optimality (relative to the abstracted representation of traffic flows), and one that is more efficient by greedily retaining only one minimum delay solution at each stage of the search, and thus is not guaranteed to find the optimal.[1] Experimentally, these procedures have been found to provide orders-of-magnitude speedup over previous solutions to the online intersection control (Xie et al. 2012). Although some of this improvement is due to increases in computational power, the larger boost is due to the single machine scheduling problem formulation. Prior work has emphasized time-based problem formulations, which result in a search process that expands solutions tick by tick. This introduces a problematic dependency on the granularity of a tick. If the tick size is too small (for example, one second), then the practical planning horizon is too short to be useful. Alternatively, if the tick size is large (for example, say one minute), then the actual nonlinear flow of approaching traffic is significantly obscured. In the case of Surtrac's formulation, problem complexity is instead a function of the number of clusters and number of phases into the future to be scheduled, which drastically reduces computational requirements and enables rapid generation of long horizon plans. Please refer to the article by Xie et al. (2012) for full details of the DP search procedure. In the article by Xie et al. (2014), a weighted formulation is introduced for purposes of assigning different levels of importance to different travel modes (for example, giving pedestrians higher priority than passenger vehicles in a downtown cultural district) if travel mode data can be captured by the type of sensors that are being used.
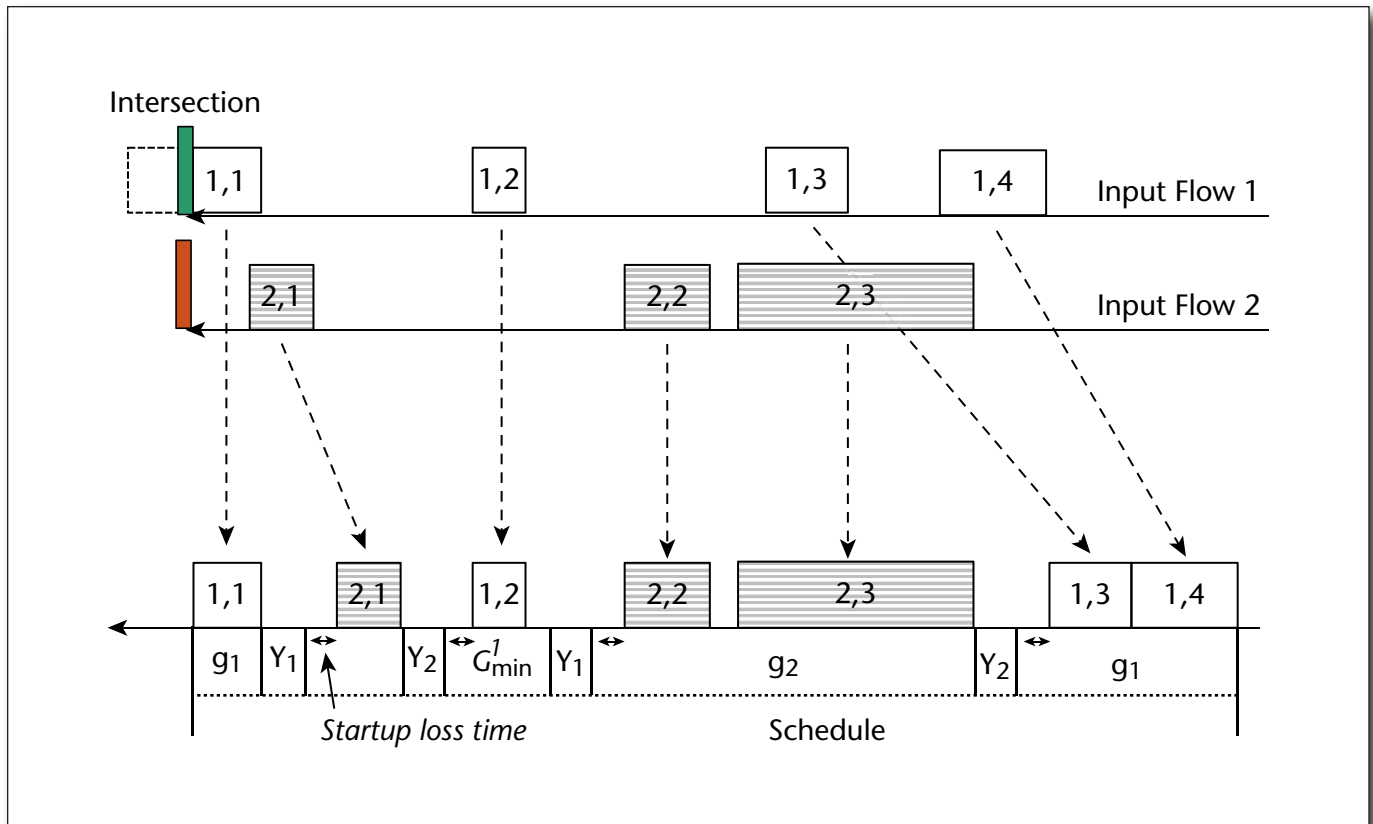
Figure 4. An Example Scheduling Problem and Solution.

Adapted with permission from Xie, Smith, and Barlow (2012).

Another key idea underlying Surtrac's schedule-driven approach is a protocol for communication of planned outflows to downstream neighbors. Once an intersection computes a new intersection timing plan, it utilizes accumulated historical information on turning proportions to estimate the clusters that will be proceeding downstream in different directions, and communicates these cluster sequences forward. Recipient Surtrac processes at downstream intersections append this information to the appropriate vehicle cluster sequences obtained through local sensor data to formulate a longer horizon planning problem. In essence, Surtrac makes the optimistic assumption that traffic flow will proceed as planned and that if it does, neighboring intersections will be perfectly coordinated. Of course, execution does not always unfold as planned, and Surtrac hedges against this possibility by layering checks for specific situations of poor coordination. For example, if *spillback* (a circumstance where vehicle queues exceed the capacity of the approaching road segment and extend back through the upstream intersection) is detected during development of an intersection schedule, the intersection scheduler will instead take a locally suboptimal decision to mitigate this upstream blockage. Full details of this approach to multiagent coordination

and decision-making can be found in the article by Xie, Smith, and Barlow (2012).

Experience in the Field

Figure 5 depicts an example of the hardware setup of the Surtrac system in the field. Surtrac runs on a separate processor (shown on the bottom shelf of the traffic cabinet), and in this case communicates through a serial port connection to the hardware controller (sitting on the top shelf in the cabinet). At the intersection depicted in figure 5, detection is via video cameras (shown mounted at the top of the light pole). To communicate with Surtrac processes at adjacent intersections, point-to-point IP radios are used (mounted on the mast arm in the intersection that is shown). In many urban settings, intersections are alternatively interconnected with fiber-optic cable. The intersection shown in figure 5 is also equipped with a dedicated short-range communication (DSRC) radio, which allows direct communication with similarly equipped vehicles (or pedestrians). We will describe more recent work with this connected vehicle technology later in the article.

Surtrac was first pilot-tested in the field in June 2012 on a nine-intersection network in the East Liberty area of Pittsburgh. Performance was assessed by doing

Figure 5. Hardware Setup.

physical drive-through experiments on the twelve most heavily traveled routes through the network, using a global positioning system tracking device to collect itinerary data. Routes were driven with the lights controlled by pre-existing timing plans and by Surtrac at multiple times of the day, and over multiple days. As can be seen in figure 6, different levels of improvement were observed over different periods of the day, but overall, averaging the results weighted by traffic volumes, travel times were reduced by over twenty-five percent, number of stops were reduced by over thirty percent, and amount of wait time was reduced by forty percent. Although emissions data were not collected, a standard fuel consumption model was applied and showed an estimated reduction of emissions of about twenty percent. Further details of the pilot test can be found in Smith et al. (2012, 2013).
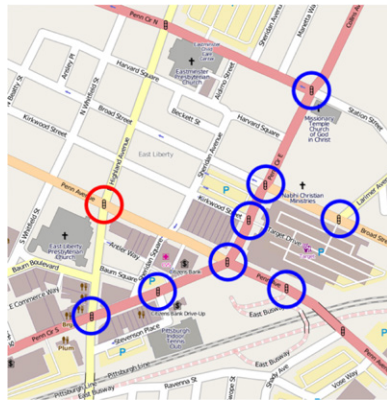
On the strength of these results, we were able to obtain additional funding to expand the network and in November 2013, the size of the Surtrac-controlled network was doubled to eighteen intersections (Barlow, Smith, Xie, and Rubinstein 2014). The same comparative evaluation was performed, this time with the original nine intersections left running with Surtrac on, in both conditions. As can be seen in figure 6, we observed essentially the same levels of performance improvement. Since that time, the network has been expanded a couple of more times, and currently Surtrac controls an interconnected network of fifty intersections, shown in figure 7. The City of Pittsburgh has recently initiated a Smart Spines project that includes a plan to add 150

additional intersections to the deployment over the next few years. In 2015, Rapid Flow Technologies, Inc. was founded to commercialize the Surtrac technology and Surtrac is now deployed in several cities in North America.

Extended Schedule-Driven Traffic Control Models

Work in recent years has focused on extension and refinement of the basic schedule-driven traffic control paradigm. One thread of research (Hu and Smith 2017a; 2017b) has examined a problem that arises in circumstances when the volume of traffic on the road network approaches saturation. Consider an extreme situation where an intersection's approaching road segments are all filled to near capacity. In this case, the aggregate predictive model produced to formulate the input scheduling problem is likely to contain a single cluster in each competing direction. In essence, the problem is no longer a sequencing problem but a problem of deciding when to split clusters (or of managing queues). To address this issue, Hu and Smith (2017a, 2017b) have proposed a hybrid online planning procedure that combines schedule-driven control with the backpressure algorithm, a state-of-the-art approach to maximizing throughput from queuing theory. Using queue length, both locally and from downstream neighbors, as a basis for weighting clusters associated with a given signal phase, a *softmax* function is specified that increasingly biases intersection scheduling decisions toward managing queues as traffic congestion increases, and extends the schedule-driven approach to incorporate the basic guarantee of stability provided by queuing theory. In simulation, the hybrid approach was shown to outperform the initial baseline Surtrac scheduling procedure and the backpressure algorithm alone by fifty percent and fifteen percent, respectively, in scenarios involving heavy traffic volumes, while providing no performance advantage over baseline Surtrac in scenarios with low and medium traffic volumes. Please refer to the articles by Hu and Smith (2017a; 2017b) for details.

A second research thrust has explored the benefits of using a higher-fidelity predictive model of traffic flows. The aggregate representation described earlier and used in the initial Surtrac intersection scheduling procedure makes a number of simplifying assumptions. For example, to characterize approaching traffic flows for typical two-way traffic phases (for example, east–west, north–south), cluster sequences from both constituent directions are merged into one composite cluster sequence before generating the timing plan (essentially by folding one sequence over on top of the other and aggregating based on temporal proximity regardless of direction). Similarly, only a single cluster sequence is associated with a given direction, regardless of the number of traffic lanes present. Both of these assumptions simplify the search, but also eliminate opportunities to minimize delay. In Goldstein and Smith (2018), a

Penn Circle Test Site (Jun 2012):

% Improv.	Travel Time	# of Stops	Wait Time	Emissions
AM rush	30%	29%	48%	24%
Mid Day	33%	53%	50%	29%
PM rush	23%	9%	36%	18%
Evening	18%	35%	28%	14%
Overall	26%	31%	41%	21%

Bakery Square Expansion (Nov 2013):

% Improv.	Travel Time	# of Stops	Wait Time	Emissions
AM rush	17%	34%	33%	16%
Mid Day	21%	37%	38%	18%
PM rush	29%	45%	46%	25%
Overall	24%	40%	42%	21%

Figure 6. Initial Results in the Field.

more expressive real-time intersection scheduling approach (called Expressive Real-Time Intersection Scheduling) is proposed, where clusters moving in different directions within a phase are considered independently during schedule construction, and cluster sequences are attributed to individual lanes in cases of multilane approaches. To manage the larger search space implied by this higher-fidelity predictive model of traffic flows, an A* search procedure is substituted for the DP search, and a novel admissible heuristic is defined to manage search efficiency. Comparative simulation analysis with the baseline Surtrac intersection scheduling procedure shows improvements in cumulative delay of up to twenty percent in low-volume traffic control scenarios, tapering to 0.5 percent to five percent in high-volume settings. Full details can be found in the article by Goldstein and Smith (2018).

A third direction of recent research has considered the impact of exchanging additional information between neighboring intersections. The original Surtrac network coordination model (Xie, Smith, and Barlow 2012) concentrated on communication of expected outflow information to downstream intersections. However, timing plans that are developed without regard to downstream conditions are susceptible to myopic decision-making and can exacerbate downstream congestion problems. An extension of schedule-driven coordination protocols to include bidirectional information flow is developed in Hu and Smith (2019). Under this scheme, an estimated delay for each vehicle is computed as an indicator of the level of congestion constraining a particular traffic flow through an intersection, and this information is communicated to upstream intersections at the conclusion of every planning cycle. The estimated delay of clusters associated with the phase corresponding to this traffic flow at the upstream intersection is then extended to incorporate downstream estimated delay so as to bias signal phase change decisions toward traffic flows that are more likely to move unimpeded through downstream intersections. Experimental analysis has shown variants of this extended protocol to significantly reduce network wide delay in high-volume situations and to generally outperform the original unidirectional scheme (Hu and Smith 2019).

Finally, consideration has been given to the additional optimization possibilities that exist if the traffic signal control system is capable of communicating
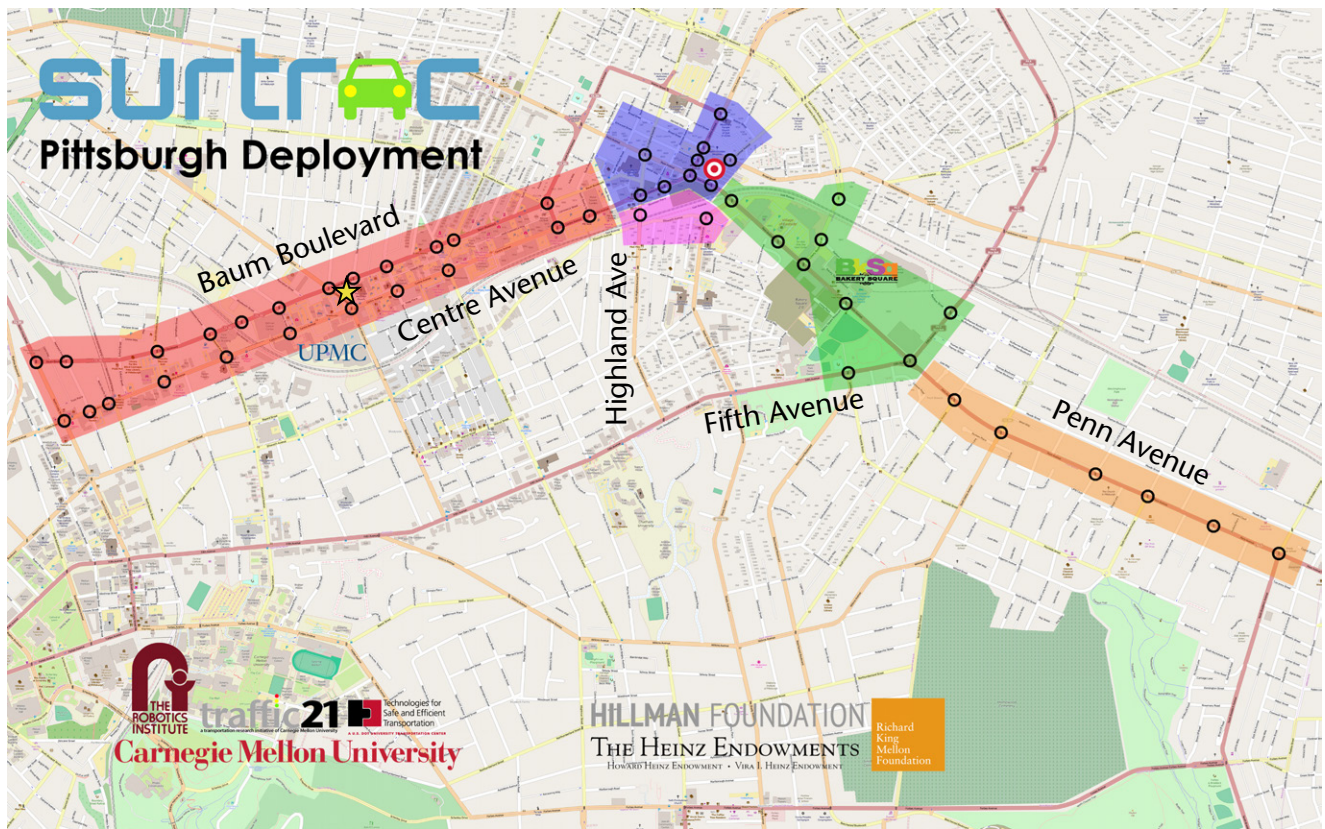
Figure 7. Current Pittsburgh Surtrac Deployment.

with vehicles and adjusting their approach speed in combination with generation of signal timing plans. Previous research has examined the environmental benefits of so-called eco-driving, that is, adjusting vehicle speed to reduce the number of stops in urban networks based on knowledge of traffic signal timing plans (for example, Xia et al. 2013). Other work has considered the problem of minimizing delay through an isolated intersection by simultaneously determining vehicle approach speeds and signal phase change times (Liang, Guler, and Gayah 2019). In Hu, Smith, and Goldstein (2019), the benefits of dictating vehicle approach speeds in conjunction with dynamically generated signal timing plans is examined and shown to outperform the basic Surtrac scheme (Hu and Smith 2017b).

Integration with Connected Vehicle Technology

Another major focus of our research over the past few years has been integration of real-time adaptive signal control with emerging connected vehicle technology. As shown earlier in figure 5, a portion of the Pittsburgh Surtrac deployment is also equipped with DSRC radios that allow direct communication between vehicles equipped with DSRC on-board units and the infrastructure. The principal motivation for such vehicle-to-vehicle and V2I communication is safety-related. Each equipped vehicle continuously broadcasts a basic safety message, which indicates its current location, heading, and speed. If two vehicles notice that they are too close to one another, then the vehicle can alert the driver or perhaps take evasive action autonomously. Similarly, an equipped intersection continuously broadcasts a signal phase and timing message, which allows vehicles to anticipate signal phase changes and adjust vehicle speed accordingly. But V2I communication also has implications for mobility. In the future when the majority of vehicles (and pedestrians) on the road are connected, the effect on vehicle sensing and adaptive signal control will be transformational. The location and types of vehicles will be known continuously, as

opposed to the use of contemporary sensing technologies, which detect vehicles only as they pass over specific spatial locations. Unfortunately, this future is still likely decades away. Our work has focused alternatively on opportunities for enhancing mobility in the shorter term, when the percentage of equipped vehicles on the road is small.

Route Sharing

One such mobility benefit that we are pursuing follows from an equipped vehicle's willingness to share its route with the infrastructure. Specifically, if a vehicle communicates its route (for example, as provided by apps such as Google Maps or Waze) to the infrastructure, then we have been able to show that Surtrac can move that vehicle through the signal network it is controlling substantially faster than it would otherwise (on average, twenty to twenty-five percent faster; Hawkes 2016). Moreover, unequipped vehicles are not really adversely affected and in some cases, benefit serendipitously. Finally, simulation results show that as the level of penetration of equipped vehicles increases, the overall performance of the network rises (Hawkes 2016). The reason for these mobility benefits is straightforward — with the receipt of this vehicle route information, uncertainty in the signal system's predictive model is reduced. For example, the system does not have to guess whether the vehicle will turn left or go straight at the next intersection. The vehicle has told it which way it will go. Consequently, the traffic signal system can do a better job of optimizing relevant signal timing plans. Technically, this is accomplished by weighting the clusters that contain each route-sharing vehicle in the system's predictive model at downstream intersections along its itinerary to reflect their greater certainty (Hawkes 2016).

Capabilities like this ability to use route information to expedite vehicles can provide incentive for travelers to be earlier adopters of connected vehicle technology. Consider a first-/last-mile freight company. This company knows its routes through the city, so if its willing to equip its fleet, a real-time adaptive signal system that supports route sharing can immediately improve its operations. Likewise for ride-hailing and package-delivery applications. We are currently working with partners to develop a pilot demonstration in the field.

Smart Optimization of Bus Movements

We are also working with the Port Authority of Allegheny County to equip the buses that move through the Pittsburgh Surtrac deployment with DSRC radios, and use V2I communication as a basis for smart transit optimization. Whereas commercial signal priority systems generally give unconditional priority to an approaching bus at the expense of all other traffic at the intersection, our goal is to achieve a more balanced and equitable approach to expediting bus movements through the intersection.

As a starting point, our focus has been on use of real-time bus information to more accurately predict an approaching bus's arrival time at the intersection, and integrating this more-accurate predictive model with route sharing (because we have access to bus routes). One complicating aspect, and the biggest source of uncertainty in predicting bus arrival times at the intersection, is the presence of near-side bus stops, which are common in Pittsburgh. This has led first to an analysis of trends in the Port Authority historical bus dwell-time data for different bus stops (Isukapati et al. 2017), and more recently, to the development of a reliable online technique for predicting bus dwell times from recent samples using a hierarchical Bayes net approach (Isakupati et al. 2019). We are currently integrating this dwell-time model into Surtrac's real-time planning loop.

Safe Intersection Crossing

A third application of connectivity that we have been pursuing is support for safe intersection crossing. We have developed a smartphone app called *PedPal* that allows pedestrians with disabilities to directly communicate with signalized intersections and actively influence traffic control decisions. Most basically, PedPal knows the speed at which its user travels and, using geometric data about the intersection received from the infrastructure, can communicate how much crossing time is required in addition to crossing intent. The traffic signal system, in return, will guarantee that sufficient time is allocated for crossing when the user gets the crossing signal. Because a smartphone has localization capabilities, PedPal also monitors its user's crossing progress. In the event that the user veers outside of the crosswalk, PedPal alerts the user, and because PedPal is connected to a real-time adaptive traffic signal control system (that is, Surtrac), the crossing time can be dynamically extended if slower than expected progress is detected. Finally, just like the vehicle route-sharing concept discussed herein, PedPal can import route information from pedestrian navigation apps, use it to anticipate the user's arrival time at the next intersection, and factor that information into the timing plan that Surtrac generates to streamline the time required to cross.

PedPal provides multiple interaction modalities to users with different types of disabilities, including a visual interface, voiceover capability, and haptic signaling. User interaction can be personalized to meet each user's needs and preferences.

Figure 8 illustrates the basic operation of the PedPal app. Upon opening the app by tapping on the PedPal icon or asking Siri to open the app (figure 8a), PedPal begins listening for messages from nearby intersections. The "No nearby intersections detected" message (figure 8b) indicates that the device is not currently in range. When the app is brought within range of an approaching intersection's message broadcasting range, the app uses this information to display possible crossing options to the user, along with context about when the crossing direction is
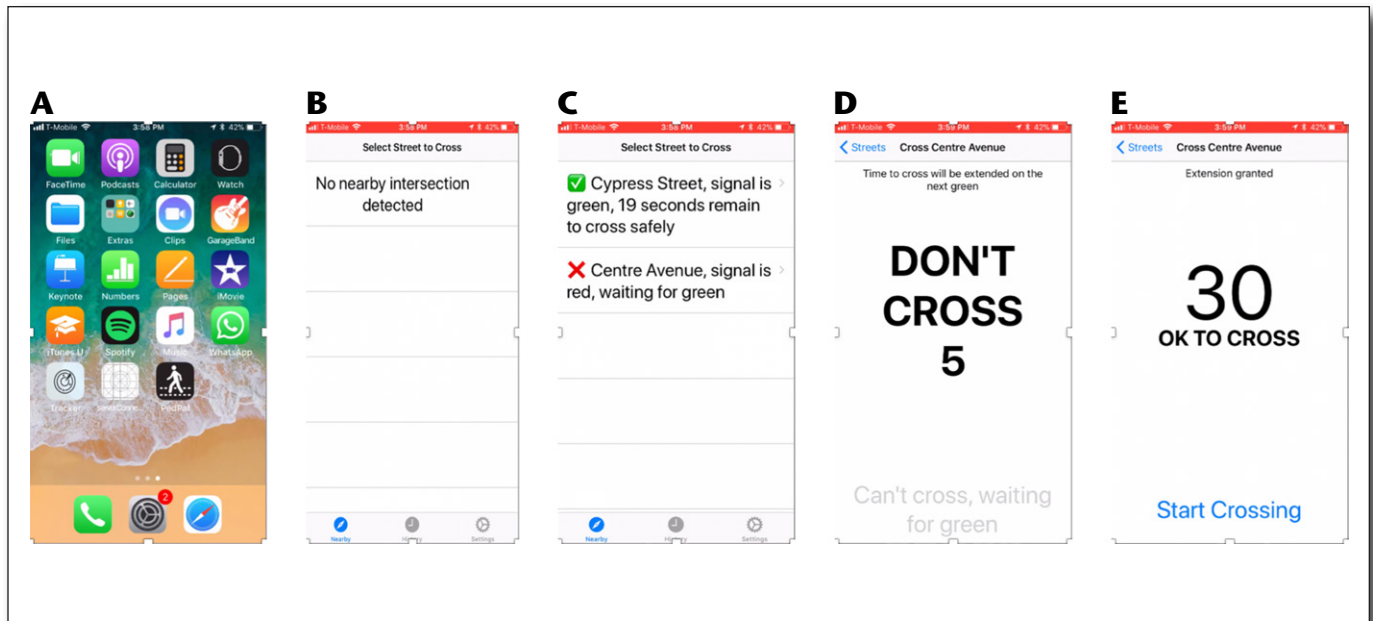
Figure 8. Basic Operation of the PedPal Smartphone App for Safe Intersection Crossing.

going to change (figure 8c). Upon arrival at the intersection, the user indicates crossing intent by tapping on one of the presented options (selected either visually or via audio voiceover). When the selection is made, the app automatically communicates both the crossing option selected and the time required, and then the display moves to another screen that provides guidance relevant to the selected crossing phase (figure 8d). In the example shown in figure 8d, the user has selected a future crossing phase, so the app warns the user not to cross. If voiceover is enabled, this message is also conveyed through audio. After delivering this message, the app goes silent but then starts to count-down once the crossing time is imminent to alert the user to get ready to go. When the crossing signal switches to the user's direction, the app announces that it is OK to cross (figure 8e). At this point, the app shows the crossing time that has been allocated and begins a countdown of the time remaining until the next signal change. In the example shown in figure 8e, the user's request for a thirty-second crossing time has been granted by Surtrac and installed as the new minimum crossing time for this cycle.

In initial user field tests, the PedPal app used a DSRC sleeve coupled to a smartphone to support connectivity to the infrastructure. More recently, a second cellular V2I communication option has been added to provide a simpler hardware solution. Please see Smith et al. (2019) for further details of this work.

The Bigger Picture

The work summarized in this article aims fundamentally at increasing the mobility of people in increasingly congested urban environments through smarter traffic signal control technology. It envisions increasingly more connected and more automated transportation systems in the future, and as connected vehicle (and connected pedestrian) technologies proliferate, we anticipate that the traffic signal network will become the gateway to travelers for real-time traffic information. There are tremendous opportunities for congestion mitigation through development of new technologies that enable online analysis of network performance, real-time incident detection, dynamic rerouting, and greater traveler safety. As autonomous vehicles emerge on the scene they will depend just as heavily (if not more) on these same capabilities. It is this vision of the future that continues to drive our research in this area. We strongly believe that smart transportation infrastructure is the key to future urban mobility.

Acknowledgments

Hu-Chieh Hsu, Joshua Lerner, Arvindh Mahendran, Pablo Nanez, Hana Rudova, Pranav Shah, Joseph Zhou, Eli Bronstein, and Conor Igoe. This article is an expanded version of the 15th Robert S. Engelmore Memorial Lecture, which I presented at the 2018 Conference of the Association for the Advancement of Artificial Intelligence, New Orleans, Louisiana.

Note

1. Given an arbitrary set of input clusters, it is not always possible to satisfy maximum time constraints on phase duration, and the just-mentioned DP search does not consider this constraint. Instead, these constraints are checked once a solution is produced, and if violated, the first offending cluster is split and the DP procedure is reapplied. See Xie et al. 2012 for details.

References

Baluja, S.; Covell, M.; and Sukthankar, R. 2017. Traffic Lights with Auction-Based Controllers: Algorithms and Real-World Data. Google Technical Report. Google, Mountain View, CA.

Barlow, G. J.; Smith, S. F.; Xie, X.-F.; and Rubinstein, Z. B. 2014. Real-Time Traffic Control for Urban Environments: Expanding the Surtrac Testbed Network. Paper presented at the 2014 World Congress on Intelligent Transportation Systems, Detroit, MI, September 2014.

Bazzan, L. 2005. A Distributed Approach for Coordination of Traffic Signal Agents. *Autonomous Agents and Multi-Agent Systems* 10(1): 131–64. doi.org/10.1007/s10458-004-6975-9.

Bazzan, L. 2009. Opportunities for Multiagent Systems and Multiagent Reinforcement Learning in Traffic Control. *Autonomous Agents and Multi-Agent Systems* 18(3): 342–75. doi.org/10.1007/s10458-008-9062-9.

Bazzan, L., and Klügl, F. 2014. A Review on Agent-Based Technology for Traffic and Transportation. *The Knowledge Engineering Review* 29(3): 375–403. doi.org/10.1017/S0269888913000118.

Dresner, K., and Stone, P. 2008. A Multiagent Approach to Autonomous Intersection Management. *Journal of Artificial Intelligence Research* 31: 591–656. doi.org/10.1613/jair.2502.

El-Tantawy, S.; Abdulhai, B.; and Abdelgawad, H. 2013. Multiagent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (Marlin-ATSC): Methodology and Large-Scale Application on Downtown Toronto. *IEEE Transactions on Intelligent Transportation Systems* 14(3): 1140–50. doi.org/10.1109/TITS.2013.2255286.

Goldstein, R., and Smith, S. F. 2018. Expressive Real-Time Intersection Scheduling. In Proceedings of the 32nd *Association for the Advancement of Artificial Intelligence* Conference on Artificial Intelligence (AAAI 2018). Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Guilliard, I.; Safier, S.; Trevizan, F. W.; and Williams, B. C. 2016. Nonhomogeneous Time Mixed Integer Linear Programming Formulation for Traffic Signal Control. *Transportation Research Record: Journal of the Transportation Research Board* 2595(1): 128–38. doi.org/10.3141/2595-14.

Hawkes, A. 2016. Traffic Control with Connected Vehicle Routes in Surtrac. Unpublished Masters Thesis, Carnegie Mellon University Robotics Institute, Pittsburgh, PA.

Hu, H.-C., and Smith, S. F. 2017a. Coping with Large Traffic Volumes in Schedule-Driven Traffic Signal Control. In Proceedings *of the* 27th International Conference on Automated Planning and Scheduling. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Hu, H.-C., and Smith, S. F. 2017b.Softpressure: A Schedule-Driven Backpressure Algorithm for Coping with Network Congestion. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. Red Hook, NY: Curran Associates. doi.org/10.24963/ijcai.2017/604.

Hu, H.-C., and Smith, S. F. 2019. Bi-Directional Information Exchange in Decentralized Schedule-Driven Traffic Control. In Proceedings *of the* 29th of the International Conference on Automated Planning and Scheduling. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Hu, H.-C.; Smith, S. F.; and Goldstein, R. 2019. Cooperative Schedule-Driven Intersection Control with Connected and Autonomous Vehicles. In Proceedings Institute of Electrical and Electronics Engineers (IEEE)/RSJ International Conference on Intelligent Robots and Systems (IROS 2019), Macau, China, November 4–8, 2019.

Isakupati, I.; Igoe, C.; Bronson, E.; and Smith, S. F. 2019. Generating Highly Predictive Probabilistic Models of Task Durations. Technical Report CMU-RI-TR-19-84. Carnegie Mellon University Robotics Institute, Pittsburgh, PA.

Isukapati, I.; Rudová, H.; Barlow, G. J.; and Smith, S. F. 2017. Analysis of Trends in Data on Transit Bus Dwell Times. *Transportation Research Record: Journal of the Transportation Research Board* 2619(1): 64–74. doi.org/10.3141/2619-07.

Lämmer, S. and Helbing, D. 2008 Self-Control of Traffic Lights and Vehicle Flows in Urban Road Networks. Journal of Statistical Mechanics: Theory and Experiment. P04019.

Liang, X.; Guler, S.; and Gayah, V. 2019. Joint Optimization of Signal Phasing and Timing and Vehicle Speed Guidance in a Connected and Autonomous Vehicle Environment. *Transportation Research Record: Journal of the Transportation Research Board* 2673(4): 70–83. doi.org/10.1177/0361198119841285.

Lowrie, P. 1992. SCATS — Sydney Co-ordinated Adaptive Traffic System — a Traffic Responsive Method of Controlling Urban Traffic. Sydney, NSW: Roads and Traffic Authority.

Luyanda, F.; Gettman, D.; Head, L.; Shelby, S.; Bullock, D.; and Mirchandani, P. 2003. ACS-Lite Algorithmic Architecture: Applying Adaptive Control System Technology to Closed-Loop Traffic Signal Control Systems. *Transportation Research Record: Journal of the Transportation Research Board* 1856: 175–84.

Mirchandani, P., and Head, K.L. 2001. A Real-Time Traffic Signal Control System: Architecture, Algorithms, and Analysis. *Transportation Research Part C, Emerging Technologies* 9(6): 415–32.

Robertson, D., and Bretherton, R. 1991. Optimizing Networks of Traffic Signals in Real Time — the SCOOT Method. *IEEE Transactions on Vehicular Technology* 40(1): 11–5. doi.org/10.1109/25.69966.

Sen, S., and Head, K. L. 1997. Controlled Optimization of Phases at an Intersection. *Transportation Science* 31(1): 5–17. doi.org/10.1287/trsc.31.1.5.

Shelby, S. G. 2001. Design and Evaluation of Real-Time Adaptive Traffic Signal Control Algorithms. PhD Thesis, Department of Civil Engineering, University of Arizona, Tucson, AZ.

Smith, S. F.; Barlow, G. J.; Xie, X.-F.; and Rubinstein, Z. B. 2012. Real-Time Adaptive Traffic Signal Control for Urban Road Networks: The East Liberty Pilot Test. Technical Report

About the Robert S. Engelmore Award

The Robert S. Engelmore Award is sponsored by AAAI's Innovative Applications of Artificial Intelligence conference and *AI Magazine.* The award and lecture was established in 2003 to honor Dr. Engelmore's extraordinary service to AAAI, *AI Magazine,* and the AI applications community, and his contributions to applied artificial intelligence.

The 2018 award was presented to Stephen F. Smith (Carnegie Mellon University) for sustained research excellence in constraint-based planning and scheduling technologies, deployment of those technologies to a range of significant real-world problems, and extensive service to the AI community that includes significant outreach to related technical fields. Smith's lecture, which provided the basis for this article in *AI Magazine,* was first presented on Sunday afternoon, February 4 in the Grand Ballroom of the New Orleans Hilton in New Orleans, Louisiana USA.

CMU-RI-TR-12-20. Carnegie Mellon University Robotics Institute, Pittsburgh, PA.

Smith, S. F.; Barlow, G. J.; Xie, X.-F.; and Rubinstein, Z. B. 2013. Smart Urban Signal Networks: Initial Application of the Surtrac Adaptive Traffic Signal Control System. In *Proceedings of the 23rd International Conference on Automated Planning and Scheduling.* Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Smith, S. F.; Rubinstein, Z. B.; Braham, S.; Dias, M. B.; and Marousek, J. 2019. Safe Intersection Crossing Through Pedestrian-to-Infrastructure Communication. Technical Report CMU-RI-TR-19-83. Carnegie Mellon University Robotics Institute, Pittsburgh, PA.

Stevanovic, A., and Mitrovic, N. 2018. Combined Alternate-Direction Lane Assignment and Reservation-Based Intersection Control. In Proceedings of the 2018 IEEE International Conference on Intelligent Transportation Systems (ITSC). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi.org/10.1109/ITSC.2018.8569755.

Stevanovic, A., and Zlatkovic, M. 2011. Evaluation of InSync Adaptive Traffic Signal Control in Microsimulation Environment. Paper presented at the 92nd Transportation Research Board (TRB) Annual Meeting, Washington, DC.

Texas A&M Transportation Institute. 2018. Texas A&M Transportation Institute (TTI) Urban Mobility Report 2018. Technical Report. Texas A&M Transportation Institute, College Station, TX.

United Nations, Department of Economic and Social Affairs. 2019. *World Urbanization Prospects: The 2018 Revision. United Nations, Department of Economic and Social Affairs (UN-DESA), Population Division ST/ESA/SER.A/420.* New York: United Nations.

Vallati, M.; Magazzeni, D.; De Schutter, B.; Chrpa, L.; and McCluskey, T. L. 2016. Efficient Macroscopic Urban Traffic Models for Reducing Congestion: A PDDL+ Planning Approach. In Proceedings of the Thirtieth Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI-16). Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Xia, H.; Wu, G.; Boriboonsomsin, K.; and Barth, M. J. 2013. Development and Evaluation of an Enhanced Eco-Approach Traffic Signal Application for Connected Vehicles. In Proceedings of the 16th International *Institute of Electrical and Electronics Engineers* Conference on Intelligent Transportation Systems (ITSC 2013). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi.org/10.1109/ITSC.2013.6728248.

Xie, X.-F.; Smith, S. F.; and Barlow, G. J. 2012. Schedule-Driven Coordination for Real-Time Traffic Network Control. In Proceedings 22nd International Conference on Automated Planning and Scheduling, Palo Alto CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Xie, X.-F.; Smith, S. F.; Chen, T.; and Barlow, G. J. 2014. Real-Time Traffic Control for Sustainable Urban Living. In *Proceedings of the* 17th International IEEE Conference on Intelligent Transportation Systems. Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi.org/10.1109/ITSC.2014.6957964.

Xie, X.-F.; Smith, S. F.; Lu, L.; and Barlow, G. J. 2012. Schedule-Driven Intersection Control. *Transportation Research Part C, Emerging Technologies* 24: 168–89. doi.org/10.1016/j.trc.2012.03.004.

Stephen F. Smith is a research professor at the Robotics Institute, Carnegie Mellon University, where he heads the Intelligent Coordination and Logistics Laboratory. He is also chief scientist of Rapid Flow Technologies Inc., a company he cofounded to commercialize the Surtrac technology. Smith's research focuses broadly on the theory and practice of next-generation technologies for planning, scheduling, and coordination. His current research interests include execution-driven planning and scheduling systems, multirobot path finding, agent-based models for distributed task and resource allocation, and intelligent transportation systems technologies.