A Real World Evaluation of the Effects
of Predictive Real-Time Traffic Signal Information

David W. Etherington
Connected Signals, Inc.
355 Goodpasture Island Road, Suite 200
Eugene, OR 97401-2119
541.654.5802 ether@connectedsignals.com

Aymeric Rousseau  Vadim Sokolov
Argonne National Laboratory
9700 S Cass Avenue
Argonne, IL 60439
630.252.7261 arousseau@anl.gov
630.252.6224 vsokolov@anl.gov

Christian Schmid
BMW North America, Inc.
2606 Bayshore Parkway
Mountain View, CA 94043
650.934.3415 Christian.CA.Schmid@bmw.de

Abstract

Research to date on the safety and efficiency implications of providing drivers with real-time, predictive information on the state of traffic signals have relied on modeling and simulation, or closed-track experiments. This paper describes a study currently underway to evaluate these questions in a real-world deployment. Working in collaboration with the cities of San Jose and Walnut Creek, CA, Argonne National Laboratories, BMW, and possibly other leading vehicle OEMs, roughly 400 drivers will be recruited. Over a period of approximately six months, their behavior in everyday urban driving will be analyzed with and without access to live traffic signal data, including time-to-green and green-wave speed recommendations. A variety of indicators of fuel efficiency, greenhouse gas production, and safety, including acceleration and braking, red-light arrivals, and red-light wait times, will be collected and analyzed to determine effects on safety and efficiency.

Keywords: V2I, Traffic Signals, Connected Vehicles, Safety, Fuel Efficiency, Traffic Data
Topic Area: Connected/Autonomous Vehicles
Pathway: Journey to Journal
Introduction

Vehicle-to-Infrastructure (V2I) systems have been the subject of intense interest in recent years, offering the promise of significant reductions in fuel consumption and greenhouse gas and other emissions, as well as safety improvements. While there have been many testbed studies that support these promises, as well as anecdotal evidence from limited deployments, the effectiveness of various V2I mechanisms in real-world remains largely unproven. Many factors that could confound the anticipated benefits, such as driver compliance, distraction issues, and the impact of other drivers’ behavior are difficult, if not impossible, to estimate. This uncertainty can make agencies reluctant to either invest in V2I systems or make their infrastructure data available to the public for fear of inadequate return on investment or, worse, paradoxical negative safety impacts (c.f., (1,2)).

The federal government has repeatedly emphasized the importance of V2I communications. However, in order to increase transportation authorities’ willingness to deploy V2I systems, it is important that real-world studies be conducted to gather data that can serve to evaluate these concerns. Such studies would, ideally, compare a statistically significant number of individual drivers’ performance with and without V2I assistance, under a variety of driving conditions, for long enough to explore “novelty” effects (e.g., whether drivers stop paying attention once they become habituated to the technology).

This paper describes a new study that will address these goals with respect to a traffic signal state V2I implementation. The study is supported in part by the U.S. Department of Energy under its “Small Business Vouchers Pilot” program, and includes participation by Connected Signals, Inc., Argonne National Laboratories, BMW, and the cities of San Jose and Walnut Creek, California. Other OEMs may be recruited during the study.

Background

Traffic-signal V2I systems offer enormous potential. The USDOT reports that, in 2014, two trillion vehicle miles were travelled on US urban roadways, two thirds of all miles driven nationwide (3), 84% in private vehicles (4). Providing real-time, predictive traffic signal information to vehicles offers the potential to reduce urban fuel consumption by 8–15% or more, according to estimates from automakers such as BMW (5,6) and Audi (7,8), as well as from the National Renewable Energy Laboratory (9). To date, however, most studies have focused on modeling and simulation (10,11) or “professional driver on closed course” studies (c.f., (6)).

Similarly, it is estimated that red-light accidents accounted for over 250,000 accidents in 1999, with over 40% resulting in injury or death (12). It seems logical that providing drivers with signal information could help reduce this toll, but there is little hard evidence. For example, a recent paper projecting the benefits of deploying V2I traffic signal information systems (13) “assumed vehicle deployment, effectiveness, and infrastructure deployment rates”. Another estimates “V2I systems potentially address about 25 percent of all crashes involving all vehicle types” (14) (emphasis added), but again lacks actual data. Experience shows that things that “obviously” should improve safety, such as red-light cameras, sometimes produce counterintuitive results (1).

To help achieve the anticipated benefits, the federal government is promoting a DSRC (Digital Short-Range Communications) approach to providing vehicles with traffic signal
information using specialized roadside and in-vehicle equipment. While DSRC offers many benefits, including nearly instantaneous relay of information, this approach requires a significant investment in new infrastructure and will take decades to be deployed widely enough to have a significant impact on US fuel consumption, emissions, and safety.

Connected Signals has developed and demonstrated a complementary approach to relaying signal information to vehicles that exploits existing connections between traffic signals and municipal traffic management systems (TMSs), and existing connections between TMSs and the Internet, to access signal data. Cellular technology is used to communicate with vehicles. This approach avoids the need to deploy special-purpose hardware at each intersection and in each vehicle. A number of pilot deployments have been completed in cities in the US, New Zealand, and Australia. Given that a large fraction of urban traffic signals are connected to TMSs, and that vehicles increasingly have built-in cellular connectivity, this approach offers the prospect of being able to connect many signals to many vehicles almost immediately at very low cost. Signal information can be accessed through Connected Signals’ EnLighten® smartphone app, as well as directly through integrated systems that have been developed with a number of major vehicle OEMs, including BMW.

For this approach to be successful, three major technical challenges need to be overcome. First, a way to securely relay signal information to vehicles must be identified. Second, the information must be made predictive, since vehicles must act on future-state information to achieve fuel savings. Third, steps must be taken to ensure that information is accurate and actionable when it arrives at the vehicle.

**Security:** To ensure system security, TMSs communicate with cloud-based servers using a lightweight, connectionless data transmission model called the Uniform Datagram Protocol (UDP) (15). Because UDP does not involve handshakes or bidirectional transmission, it lets transportation authorities provide access to their signal data in real time without fear that their systems will be compromised by “upstream” traffic. Cryptographic authentication can also be used to preclude malicious insertion of data. Connected Signals currently supports digital signing of data from most TMSs. A number of existing municipal signal streams currently incorporate digital signatures, and this is expected to become more widespread. Encryption, as well as signing of data sent to end users are also possible, although current deployments do not yet exploit these capabilities.

**Prediction:** Once the data reaches the cloud, it is used to provide current signal state information to vehicles and stored for use in building predictive models. These models are constructed using sophisticated statistical and machine-learning techniques. These techniques incorporate a variety of factors (such as current signal state, vehicle and pedestrian calls, and many others).

**Actionability:** Cloud-based servers use vehicle position and speed information, current signal states, and the predictive signal models to provide data to vehicles using the cellular network. Applications in the vehicle (or on smart phones in the vehicle) can exploit this information to provide various services such as signal countdowns and engine start-stop timing, green wave speed recommendations, or red signal violation warnings. The predictive nature of the information makes it possible to overcome the small latency inherent in the various communications, allowing the information received at the vehicle to be actionable in real time.
Figure 1 shows a high-level view of the overall system architecture. The cloud services elements can be replicated as required to provide redundancy, geographic distribution, and load balancing.

Figure 1: Architectural Overview

One advantage of this approach is that it is compatible with DSRC-based V2I systems. Agencies can rapidly and cheaply connect vehicles to their signals that have reachback to their TMS, focusing their limited new-hardware deployment budgets on DSRC installations where those provide unique advantages (either at signals that cannot be readily connected to the TMS or where other considerations favor DSRC). In-vehicle systems that exploit signal information can easily be configured to utilize DSRC-provisioned information when it is available.

The Study

Participants

The study currently underway is a partnership between Connected Signals, Argonne National Laboratories, BMW, and the cities of San Jose and Walnut Creek, CA. Approximately 400 drivers who live or work in the study areas are being recruited to participate in the study, which will collect data for approximately six months.
San Jose and Walnut Creek are providing and supporting the real-time signal feeds that will support the study. BMW and any other participating OEMs will help Connected Signals recruit drivers. They will also help develop the instrumentation needed to collect the data that will be used to assess the impacts of access to real-time signal data. The Systems Modeling & Control Group at Argonne’s Center for Transportation Research (ANL) will take the lead on experimental design and data analysis, ensuring that the data collected supports accurate and statistically significant determination of the impact of signal data. The group provides innovative solutions to the design of efficient and sustainable transportation systems. Their researchers have developed a set of tools that allow tackling the challenges posed by next-generation transportation systems. These tools include an agent-based modeling framework for transportation system, data analysis methodologies for traffic and vehicle data, as well vehicle powertrain modeling simulations for assessing energy impacts. Transportation energy modeling tools developed at Argonne have become industry standard and extensively used by private companies as well as local and federal governments for policy making.

BMW has already partnered with Connected Signals to integrate predictive signal data streams into their infotainment units (16). The functionality is available for the majority of their vehicles produced in 2010 and later. The initial fielded version supports signal countdowns and “green-light-assist” information that tells drivers whether they will make an upcoming signal at their current speed. A sample screenshot is shown in Figure 2. The display indicates that the light is currently green and will remain so for 28 seconds. The green arrow shows that, at the current speed, the car will arrive at the intersection during the green phase. The application is available in selected cities, including San Jose and Walnut Creek, CA. BMW’s implementation and Connected Signals’ EnLighten app will serve as the starting points for the signal information distribution and data collection needs of this study. Special versions of each will be produced to accommodate the study’s needs.

![Figure 2: BMW EnLighten display: next light will be green, based on current speed](image)

**Framework**

The study will involve recruiting roughly 400 drivers. With the drivers’ consent, data on vehicle position and speed, acceleration, braking, and (where possible) fuel consumption will be collected. This data will be transmitted to servers in the cloud, where it will be merged with cotemporaneous signal-state information for later, offline, analysis.
Drivers will be recruited who spend significant portions of their driving time both in and out of the covered areas. This will allow their behavior to be compared longitudinally, making it possible to detect habituation effects and eliminate biases that might occur in simple sequential “without data/with data” trials. ANL will help tailor the analysis to match like data to ensure that, for example, freeway data does not pollute the out-of-coverage datasets.

Data collection will run for approximately six months, to ensure that a meaningful amount of data is collected for each driver, and that each driver experiences a variety of driving conditions.

When in covered areas, drivers will be provided with predictive signal information telling them, when possible, whether they will make or miss the next signal at their current speed, a recommended speed to make the next signal, and countdowns for red signal durations when they are stopped. For safety reasons, speed recommendations will be limited by the current speed limit, and red-light countdowns will stop at 5 seconds before the signal changes to force drivers to rely on the physical traffic signal. At that time, a chime will also sound to alert drivers to return their focus to driving in case they may have become distracted while waiting for the signal to change.

**Experimental Design**

The experimental design for the study is intended to maximize our ability to determine the effects of signal data availability, given the constraints of what can readily be obtained from a collection of privately owned vehicles and a self-nominated group of participants.

A number of steps will be taken to minimize—as much as possible—the effects of such factors as driver and vehicle variability, habituation, and differing driving conditions in and out of signal coverage. First of all, during an initial period, drivers will not be provided with signal information for a sufficient number of trips to establish a baseline. During that time, information will be collected on drives and correlated with real-time signal state information. This will allow determination of how drivers respond to the signal state information they get in the normal way (looking at the lights) without additional predictive or guidance information.

Secondly, throughout the study data will be collected from drives both inside and outside the signal-coverage area. Since the locations (but not the states) of signals outside of coverage are known, this will help distinguish between changes in drivers’ behavior that result from access to signal data and changes that result from other factors such as weather or traffic conditions. While this is not a perfect comparison, it should provide reasonable indicators of the significance of the observed results. ANL will also exploit their unique skills and tools to analyze driver behavior from GPS traces drawn from both the Transportation Secure Data Center (TSDC) (17), which contains second-by-second data for millions of miles of driving, to help understand observed variability and control for extraneous factors that may occur in the out-of-coverage data.

Finally, each driver will be assigned a unique ID that will be used to associate all their drives. This will allow changes in driver behavior to be analyzed longitudinally over the course of the study, including between control and signal-informed driving conditions. The unique IDs will be created so they cannot be inverted to identify particular drivers to ensure the privacy of drivers in the study, and all data will be anonymized using these IDs as it is received.
Although the characteristics of the study’s drivers and their route selections cannot be controlled for, the ability to compare individual drivers longitudinally, both with and without signal data, over an extended period should make it possible to minimize the influence of such factors. ANL will use standard statistical techniques to determine the necessary sample sizes (number of trips) for each participant to ensure statistical significance of the results. Variability in driver behavior may mean that more data is required for some drivers than others.

For those vehicles with integrated signal information capabilities, the necessary information will be captured directly from the participating vehicles. For vehicles without direct integration, a special-purpose version of EnLighten will be developed to acquire the necessary data from the smartphones’ sensors. For each trip, time-series data will be collected on:

- Vehicle position, heading, and speed,
- Number and duration of stops,
- Acceleration and deceleration profiles,
- Energy consumption (if available),
- Availability, timing, and content of provided signal information, and
- Actual signal state (if known).

This information will be associated with the vehicle type and driver ID. All data will be sent to the cloud both to facilitate provision of signal-state predictions to the vehicle and for recording for subsequent analysis. Data will be quality checked and batched for relay to ANL, where it will be analyzed.

Since baseline and longitudinal data without signal provisioning will be collected in addition to data with signal information, it should be possible to reliably estimate the effects on energy consumption and safety and estimate the impact of signal time information.

**Data analysis**

As mentioned above, the analysis phase of the study will be led by ANL, who have an extensive set of analysis and simulation tools and data that can be brought to bear. These will be important to help ensure statistical significance of the study results and to maximize the conclusions that can be drawn.

For example, the voluntary nature of driver participation offers limited control over the process of selecting the participants for the study, and limits the ability to observe any information about a participant, such as age or years of driving experience. Absent any randomization of the participant selection, or ability to control for drivers’ characteristics, some confounding effects can be expected in the energy and safety outcomes (e.g., people who volunteer to participate in the study might already be conscious about their driving style or drive on routes for which they can predict the traffic signal changes and adapt accordingly). To help ameliorate these effects, the analysis will compare driving patterns of the study drivers with a much larger sample collected by other studies (17) that includes second-by-second vehicle GPS-track database representing multiple millions of miles of real-world driving data. Such a comparison would support identification of potential biases in the sample and categorization of study drivers so that confounding effects can be eliminated.

The goals of this study are to determine the impacts of traffic-signal V2I on fuel consumption, pollution, and safety. Because no mechanisms are available to directly collect safety data,
other data that can be collected (acceleration/deceleration, path, speed profile, and light state) will be used as surrogates (18,19,20).

**Expected Nature of Results**

The main results expected will show how the availability of signal information affects:

1. Driver speed behavior approaching red/green lights,
2. Arrivals at red lights,
3. Sudden deceleration behavior, and
4. Wait times at red lights.

**Approach speed:** Anecdotal information from informal A/B studies at Connected Signals and BMW suggest that, contrary to what some might expect, drivers are less likely to speed up approaching green signals when they are told whether or not they will make the light. It should be clear from the data that will be collected whether this is indeed the case.

**Red-light arrivals:** Data show that arriving at a red signal is inherently dangerous—a certain number of drivers fail to stop and collide either with a vehicle that has stopped or one with the right of way (12). The data that will be collected will show how often drivers arrive at (or go through) red signals, supporting an analysis of whether the provision of signal information (including speed to green) reduces this.

**Sudden deceleration:** When drivers decelerate suddenly, it may result from a natural response to behavior of drivers ahead, a change in a traffic signal, or an accident. Depending on the rate of deceleration, it may be possible to distinguish accidents from the other cases. Similarly given that we know the state and location of traffic signals ahead of the driver, we should be able to distinguish cases where drivers have had to stop for a red signal. In any event, since sudden deceleration is typically less safe than more gradual deceleration, a significant variance between signal-aware and non-signal-aware driving can be taken as providing evidence of safety benefits or risks from having the V2I data available.

**Wait times:** Finally, the data will show whether the amount of time spent stopped at red lights changes significantly. In conjunction with TSDC data, this information can be used to infer changes in fuel consumption (for vehicles without automatic engine stop/start features, at least).

**Conclusions**

Decisions by transportation authorities on whether to deploy signal-information V2I systems (and by OEMs on whether to include features based on them in their vehicle offerings) are currently made based on arguments about what might be possible and studies undertaken in controlled environments that may not accurately reflect the real world and its complexities. Intuitively, this type of data offers significant potential environmental and safety benefits, but there is little actual real-world data. The study this paper describes offers a way to provide objective data about the actual benefits that can be achieved.
Acknowledgment and Disclaimer

The authors would like to thank Matt Ginsberg and Tim Stirling of Connected Signals for valuable comments, criticisms, and suggestions on an earlier draft of this paper. Of course, responsibility for remaining deficiencies belongs to the authors.

This material is based upon work supported in part by the U.S. Department of Energy’s Small Business Vouchers Pilot program. Neither the United States Government nor any agency or employee thereof, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process or service does not necessarily constitute or imply its endorsement, recommendation or favoring by the United States Government or any agency thereof. The views and opinion of authors expressed herein do not necessarily reflect those of the United States Government or any agency thereof.

Bibliography


2. Laura Wells. Personal communication. 2015.


