Optimal Traffic Monitoring in a New Data Age

September 2019

I-95 Corridor Coalition

Contents

Abstract 3

Acknowledgements 4

1.0 Executive Summary 5

2.0 Introduction and Background 7

3.0 Let's not Forget Where We've Come from … Inductive Loops … 8

4.0 Newly Proven Technologies 9

4.1 Commercial Probe Data 10

4.2 Vehicle Re-identification Data 13

4.3 High-Resolution Controller Data 15

4.4 Cell Phone Geo-Location Data 16

5.0 Applications, Applications, Applications 17

5.1 Travel Time on Signs – the App that Ushered in Probe Data 17

5.2 Probe Data at Work in Traffic Management Centers 17

5.3 Bluetooth - First Task of Validating Probe Data 18

5.4 Bluetooth - Assessment of Signal Performance 19

5.5 HRCD for Automated Traffic Signals Performance Measures 20

5.6 Smart Work Zones 21

5.7 NPMRDS and MAP-21 Performance Measures 21

6.0 Emerging, Leading Edge Technologies 22

6.1 Vehicle Trajectory Data 22

6.2 Estimating Volume with Probe Data 22

7.0 Conclusions and Recommendations 23

References 24

Abstract

This guidance document was created by the I-95 Corridor Coalition with funding from the Multistate Corridor Operations and Management Program (MCOMP) II Grant awarded by the Federal Highway Administration (FHWA). It is intended to raise the awareness of viable traffic data sources that have emerged over the past decade. The document is targeted to engineers, planners, and managers at state, county, metropolitan planning organizations (MPOs), and cities involved in transportation planning and operations. This new age of data and communications has opened up new avenues for traffic monitoring. The objective of this paper is to outline how to optimally leverage these new data sources in combination with traditional data resources.

While much of the news cycle about transportation relates to autonomous vehicles or traffic networking companies such as Uber and Lyft, the traffic industry continues to undergo a data and information revolution fueled by proliferation of Global Positioning System (GPS) equipment, inexpensive wireless communications, and big data processing. Third-party sources have now gained extensive knowledge of the state of traffic stream through aggregating information from fleets that self-report position and speed data frequently. This data, when appropriately integrated with traditional sources, has the ability to accurately characterize the nature of traffic flow at nearly all locations on the roadway network at all times. Data from these sources have fueled existing applications and have enabled applications that were once prohibitively expensive or impractical to implement with conventional sensors. Combining the strengths of traditional sensors with new data technologies, roadway jurisdictions can gain tremendous advantage for planning, operating, and managing traffic in this new data age.

Acknowledgements

This work was supported via the Multistate Corridor Operations and Management Program (MCOMP) II Grant awarded by the Federal Highway Administration (FHWA) to the I-95 Corridor Coalition. The Views expressed in the article do not necessarily represent the views of the FHWA or the U.S. Government.

1.0 Executive Summary

With the growing use of outsourced probe data, development of new data collection technologies and services, and prevalence of legacy sensor investments, many agencies are re-examining how best to architect their traffic monitoring and management systems. A decade ago, most agencies were data starved; any improvement in available traffic monitoring directly impacted the quality of the downstream applications. Now with multiple data sources, types and levels of fidelity, the design of a monitoring system is becoming requirements driven, and is always fiscally constrained. The goal of this document to provide guidance to coalition member agencies in architecting and planning data acquisition to meet the growing needs of operations, performance measures, and planning. The section below provides recommendations that have been suggested by the authors, and vetted by Coalition leaders and members to provide practical guidance.

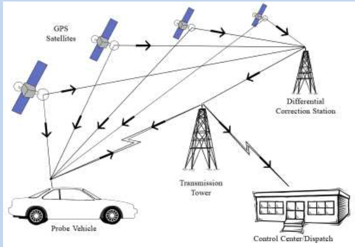
- **Commercial probe data** should be in the arsenal of any state DOT, and shared with sub-jurisdictions. The value proposition, scalability and usefulness for a variety of applications, spanning operations, performance measures and planning, are well documented. Analytics options are robust and growing, and supported by a number of industry players. Key personnel within the DOT should be well-versed in its capabilities and limitations.
- **Conventional sensor** investments are still needed and viable, and will be for the foreseeable future. Conventional sensors are justified on critical portions of the roadway where ownership and direct control of the data stream trump the value proposition of probe data. Sensors are also needed to continue to sample across a broad array of road classes and types as ground-truth sources for spot speed and counts. Sensors are needed for site-specific data, like signalized intersections using adaptive control. The data from agency-owned sensors are owned by that agency, and can be shared and used without being subject to licensing.
- **Re-identification data** (such as Bluetooth and WiFi) should be viewed as travel time sensors (as opposed to speed sensors). Such data is needed for location specific studies related to travel time or origin-destination (O-D) studies. Re-identification is typically used as ground truth for validating accuracy of sources of travel time data (such as commercial probe data) and will continue in that capacity.
- **High-resolution controller data (HRCD)** and the corresponding Automated Traffic Signal Performance Measures (ATSPMs) are in the domain of traffic signal engineers. Moving forward, any signal upgrade should include consideration for acquisition and processing of HRCD and ASTPMs. Note that this may also require additional investment in sensors at the intersection.
- Planning level studies should consider the value proposition of **Cell Phone Geo-Location data** and other emerging forms of O-D. The authors fully acknowledge that such data is difficult to ground truth or prove representativeness, but the cost, sample size, and timeliness of these new sources cannot be ignored.
- **Volume estimation** deserves special attention, as it completes the dimensionality of traffic data. The authors anticipate that volume estimates for planning (AADT), for traffic (turning movements), and for operations (vehicles/hour) will enter the mainstream outsourced traffic data market within 24 months, with early AADT products already available. As with commercial probe data, the I-95 Corridor Coalition is actively engaged to accelerate the maturity of this new data source, bringing unbiased expert resources to the discussion and research, and act as the collaborative liaison between vendors and DOTs.

Optimal Traffic Monitoring in a New Data Age – Executive Summary

Conventional Sensors

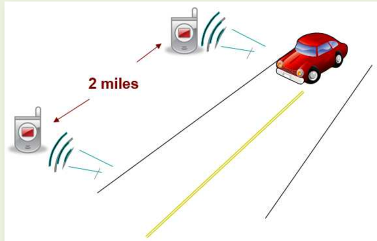
- Still needed and viable, and will be for the foreseeable future.
- Justified on critical portions of the roadway where ownership and direct control of the data stream trump the value proposition of probe data.
- Needed to continue to sample across a broad array of road classes and types as ground-truth sources for spot speed and counts.
- Data are owned by the agency and can be shared and used without being subject to licensing.

Commercial Probe Data



- Useful for any state DOT and sub-jurisdictions.
- High value proposition, scalability and usefulness for a variety of applications from planning to operations.
- Analytics options are robust and growing, and supported by a number of industry players.
- Key personnel within the DOT should be well-versed in its capabilities and limitations.
- **Useful Applications: Travel Time on Signs, Signal Performance Studies, Smart Work Zones**

Re-identification Data



- Bluetooth and WiFi
- Should be viewed as travel time sensors (as opposed to speed sensors). Such data is needed for travel time or O-D studies.
- Re-identification is typically used as ground truth for validating accuracy of sources of travel time data (such as commercial probe data).
- **Useful Applications: Travel Time on Signs, Travel Time Validation, Signal Performance Studies, Origin-Destination Studies, Smart Work Zones**

HRCD (High-Resolution Controller Data)

- HRCD and the corresponding Automated Traffic Signal Performance Measures (ATSPMs) are in the domain of traffic signal engineers.
- Signal upgrades should include consideration for acquisition and processing of HRCD and ASTPMs.
- **Useful Applications: Signal Performance Studies (ATSPMs)**

Emerging, Leading Edge Technologies

Trajectory Data

- Waypoint data every 1 second
- OD studies, arterial analysis, freight studies
- Market-ready by 2021

Estimated Volumes from Probe Data

- AADTs, turning movements, vehicle/hour
- I-95CC Validation
- Market-ready by 2021

September 2019

2.0 Introduction and Background

With the prevalence of data services offered by the private sector, many roadway agencies are re-examining how best to architect their traffic data collection plans, be it for monitoring, management or planning efforts. This document is targeted toward a broad audience that includes technical disciplines such as engineers, planners, modelers, information technology (IT) professionals, and data analysts, as well as a non-technical audience that grapple with policy and funding in the transportation space, such as managers, appointees, and policy personnel. As such, the vocabulary and level of technical details is targeted to provide a broad understanding of the traffic data revolution that we are in the midst of, and how to best leverage it. Inasmuch as possible, technical details are provided in referenced material for those wishing additional information.

The new data landscape means that agencies can rethink and retune their data collection strategies. The high cost and maintenance of installed roadside sensors are now being weighed against the value of the data that can be delivered by commercial companies. Whereas fixed sensors provide precise data with known accuracy and are fully owned and controlled by the agency, data offered by private companies gracefully scales to the entire network and measures some attributes that are difficult to capture with sensors. It is also easier for agencies to change data vendors or switch out data layers than replacing permanent fixed sensors.

Mapping existing business processes, as well as new capabilities into this new reality of traffic information choices is the subject of this guide sponsored by the I-95 Corridor Coalition (or simply the Coalition henceforth). It is intended both to educate and inform, and is written for the benefit of the Coalition's members as well as other like agencies and institutions. This guide is organized as follows:

- *Let's not Forget Where We've Come from … Inductive Loops …*
  The contrast of today's information rich outsourced traffic data is best viewed in contrast with the capabilities, costs, and constraints of the yester year. This section takes a look at sensor technology, specifically the venerable 'loop detector', providing a summary of its capability and limits as well as how it was and continues to be leveraged in transportation practice.
- *Newly Proven Technologies …*
  This section offers an overview of four traffic data technologies that have enhanced traffic data information over the past decade. These four technologies include:
    - Commercial Probe Traffic Data Services
    - Vehicle Re-identification Technology
    - High-Resolution Controller Data
    - Cell-tower Geo-Location Data
- *Applications, Applications, Applications …*
  This section looks at half a dozen applications that have utilized the four technologies described above for the benefits of the road jurisdiction and its customers. These can be considered case studies for the audience, to cross-reference the technologies to applications and the agencies that have blazed the trail.
- *Emerging, Leading Edge Technologies …*
  New technologies are constantly being introduced to provide additional traffic data. A couple of the emerging (though not yet proven) technologies are introduced here. Although these emerging traffic technologies have not hit mainstream, they promise even increased efficiencies. Agencies are encouraged to keep an eye open as these data sources mature.

- *Conclusions and Recommendations …*
  Final recommendations are provided to agencies with respect to architecting a robust traffic data system that leverages the best of both traditional sensor systems, as well as takes advantage of the scalability and cost efficiency of new data sources.

It is not the intent of the guide to be a definitive academic reference, but rather to bring an easy-to-understand layman's perspective that provides historical context while accurately documenting the progression and validity of newly emerged (and emerging) traffic data sources. Reference material and sources are cited for follow up for engineers and technicians, while only summaries and overviews are provided herein. In all this document points toward valid uses, and appropriate mix of traditional sensors with new traffic data technologies.

3.0 Let's not Forget Where We've Come from … Inductive Loops …

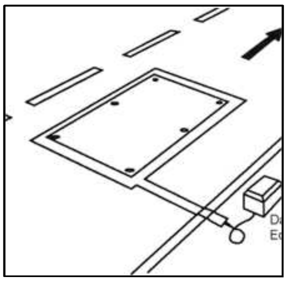


Figure 1 - An Inductive Loop

Today's emerging traffic information industry is best viewed in contrast with the capabilities, costs, and constraints of the yester year. This section takes a look at sensor technology, specifically the venerable 'loop detector', providing a summary of its capability and limits, as well as how it was, and continues to be leveraged in transportation practice.

An inductive loop is a coil of conductive wire placed in the pavement. A vehicle, composed heavily of magnetic material such as iron and steel, passing over an inductive loop creates an electrical signal in the loop. The electrical signal can be / is used to:

- Determine traffic volume - This is the basis of most data collection for the Highway Performance Monitoring System (HPMS), the federal program that collects data for various uses, the most notable of which is allocation of the federal gas tax fund.
- Indicate vehicle presence, or that of a queue of vehicles at a traffic signal – a typically 50' loop detects if a vehicle is waiting to be serviced. The absence of a vehicle signal indicates that the queue has been emptied, allowing the signal to progress to servicing other approaches (if so programmed). In other applications, an inductive loop in this configuration triggers ramp metering.
- Assess the speed of traffic - two small loops places in sequence allows for the accurate calculation of speed (spot speed) at the sensor.
- Gauge the density of traffic – The percentage of time the loop detects a vehicle (referred to as activated) versus the time in which it is dormant reflects the density of traffic on the roadway. Technically this is referred to as 'occupancy' ratio, but not to be confused with vehicle occupancy which is the number of people in the car.
- Classify types of vehicles – axles tend to present stronger signals that other parts of the vehicle allowing indicative loops to distinguish large trucks from light-duty vehicles based on the pattern of electrical pulses representing the number and spacing of axles.

Note that although loop detectors are considered the most accurate (ground truth or reference data), they are subject to known errors. If a vehicle contains little to no iron or steel such as motorcycles or bicycles, there is a chance the vehicle will not be detected, or be mis-classified. Extremely dense traffic congestion may create situations where individual vehicles cannot be distinguished, thereby compromising the traffic count. More importantly, loop detectors are invasive, requiring alteration of the pavement, and subject to costly maintenance for both the loop and its processing equipment. Note that many other sensors technologies have entered the market based on RADAR, acoustics, imaging, vibration and other mechanisms, but all mimic the functionality of a loop detector is some respect.

Whether an actual inductive loop, or a sensor that emulates the inductive loops in some fashion, all share common characteristics:

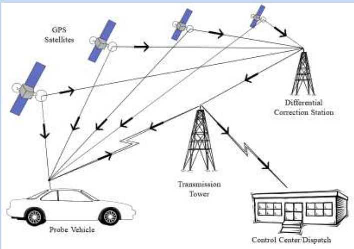
- All sensors measure traffic at a specific point. Sensors typically measure the volume of traffic and the speed of traffic as they traverse a specific point. Inferring travel time from sensor-based traffic data is problematic. On freeways, the speed observed at a sensor is considered representative of the speed along the corridor. On freeways this is generally a safe assumption. Sensors placed approximately ¼ mile apart are considered sufficient to accurately infer travel time, even during congested conditions when speed can fluctuate substantially. On signal-controlled arterials with frequent stopping and queuing, the authors are not aware of any convention with which to use sensor-based speed data to accurately calculate a corridor travel time.
- All sensors capture data for 100% of the vehicles that traverse that point in contrast to newer technologies that only gather data from a sample of the vehicle stream.
- Sensors are capital and maintenance intensive. Roadway agencies must purchase, install, maintain, as well as integrate into a data communications network to harvest roadway data.
- Data from sensors are owned by the roadway jurisdiction. There are typically no licensing agreements, or data use agreements to sign from traditionally installed sensors.
- Sensors cannot infer origin or destination of vehicles, or track the vehicle through the network. Data is limited to only the location of the sensor, and linking of vehicles from sensor to sensor is practically not possible.

Many of these characteristics are in contrast to the new technologies proliferating as will be discussed in the next section.

4.0 Newly Proven Technologies

While much of the news cycle in transportation have been given over to the race to fully automate vehicles, the traffic data industry has been undergoing a relatively silent revolution. Third-party sources provide data derived from what is referred to as *probes*, vehicles or drivers that periodically report their position and speed. This phenomenon took root over a decade ago when GPS subsystems began to be inexpensive enough to be embedded in vehicles and mobile phones, and digital communications (2G and higher cell phone infrastructure) proliferated across the US. When introduced in the mid-2000s, probe data companies relied heavily on data from GPS equipment embedded in the freight transport vehicles (and monitored by fleet telematics companies) as well as data from portable navigation devices with data connectivity. With the smart phone revolution, this was augmented with data harvested from smart phones that periodically report position and speed. This section also covers Bluetooth re-identification data, high-resolution controller data, and cell-phone geo-location data.

Commercial Probe Data



- Useful for any state DOT and sub-jurisdictions.
- High value proposition, scalability and usefulness for a variety of applications from planning to operations.
- Analytics options are robust and growing, and supported by a number of industry players.
- Key personnel within the DOT should be well-versed in its capabilities and limitations.
- **Useful Applications: Travel Time on Signs, Signal Performance Studies, Smart Work Zones**

4.1 Commercial Probe Data

Today it is easy to imagine your Android or iPhone providing information about roadway conditions based on an aggregation of measured GPS locations and speed from thousands of users in a metropolitan area, as is common in many navigation apps. Rather than a stationary sensor that detects individual vehicles at single point, myriads of vehicles and smart phone users scattered randomly across the roadway network report their individual location and speed. While a sensor has high precision at a single point, probe data can accurately characterize the nature of traffic flow across the entire roadway network and at all times, provided that there are sufficient number of probe vehicles. This ability to scale to the entire network without having to purchase, install, power, maintain, and communicate with a network of sensors creates a value proposition which departments of transportation (DOTs) are increasingly utilizing. These ever-improving traffic data services have enabled many applications that were once extremely costly, or completely infeasible with traditional sensors.

Probe data relies on a mixture of devices self-reporting the location and speed (not just smart phones) in order to gather enough information to understand the traffic picture across the network at all times. Some of the data sources are presented here.

- **(1990s) Long distance freight hauled by large trucks** were the first vehicles to begin to self-report their position and speed periodically. During the 1990s, satellite communications combined with GPS technology enabled the first generation of fleet telematics that track and report the movement and delivery of high-value freight shipments. These early systems were only cost competitive for high-value freight movements, and as such were primarily used only on interstate shipping. Such systems reported only once per hour due to the high-cost of satellite communications. It was soon recognized that by observing the time and location of a large number of self-reporting freight vehicles, the traffic speed and congestion on interstate highways could be roughly inferred.
- **(Late 1990s – 2000)** As **cellular based data communications** were introduced late 1990s and proliferated into 2000s, and as the cost of GPS equipment declined, it became less costly to implement fleet telematics, allowing such services to spread to regional carriers, and lower value freight shippers. As more fleets became equipped with telematics, the ability to infer traffic conditions from such data expanded beyond interstate highways.

- **(Mid-2000s)** As **fleet telematics** switched from costly satellite communications to more affordable cellular networks, vehicles could report more frequently. Instead of systems which reported once per hour, systems emerged that report once every 15, 10, 5, or even 1 minute. As more fleets invested in this technology, the first commercial traffic data services based on aggregated fleet telematics data began to appear in the mid-2000s.
- **(Mid-2000s)** Concurrently, widespread availability of **GPS navigation devices** brought about a corresponding revolution in digital mapping. GPS navigation devices called Personal Navigation Devices (PNDs) also featured embedded cellular data communications in higher-end models, allowing the user to communicate their location and speed as well as obtain travel data such as construction and congestion bottlenecks. Consumer electronic devices created another data collection channel for probe data in addition to fleet telematics from which to more accurately infer traffic conditions.
- **(2010) Car manufacturers** soon came on board in the 2010s, and fully connected built-in navigation options on luxury models appeared. The first generation of vehicle navigation systems relied on stored electronic maps and traffic information broadcast on FM sidebands. These later evolved into two-way communications integrated via cellular data services such that they reported their position and speed to the OEMs, creating additional streams of vehicle probe data to reflect roadway conditions.
- **(Mid-2000s – 2019) Smart phones** came into the picture in the mid-2000s, and grew quickly such that in 2019 (as of the writing of this report) 81% of adults own a smart phone (Pew, 2019). Each smart phone has a built-in GPS sub-system, and ability to communicate data at large bandwidths to a variety of application providers. As smart phones proliferated, the traffic data that could be harnessed grew exponentially. The majority of drivers report location and speed, some up to once per second if running the appropriate application software, to a central location. Cell phone manufactures, cell phone app providers, cell phone operating systems, and cellular service providers all have become players and conduits of massively large data streams from which roadway traffic conditions and travel patterns can be inferred.
- Moving forward, a large proportion of vehicles coming off the assembly line are networked to their manufacturer to monitor the vehicle performance, monitor security, respond to emergency situations, and to push out new software upgrades. Such highly networked vehicles promise to provide not only location and speed of the vehicle, but also report parameters such as temperature, precipitation, humidity and other factors used in the data systems of modern engines.

The previous paragraphs provided an overview of how commercial probe data evolved. Currently, three primary vendors (INRIX Inc., HERE, and TomTom) supply probe traffic data to government jurisdictions in the US based on these methods. All three are approved and qualified vendors in the I-95 Traffic Data Market Place that was initiated in 2006, otherwise known as the *Vehicle Probe Project* (*VPP) Marketplace*. The Vehicle Probe Project remains a vibrant and comprehensive asset to the roadway jurisdictions in the Coalition to acquire, at competitive rates, the probe data services supplied by these companies both for their state's use, as well as adjoining states and sub-jurisdictions within the state.

At the heart of commercial probe data services is the reporting of speed (or conversely travel time) on roadway segments. Common to all three vendors is the offering of a traffic data feed, technically known as an Application Programming Interface (API), from which the speed of a roadway segment can be obtained. Typically speed data is most complete for the higher roadway classes: freeway, arterials and major collectors. As the technology has progressed, data providers have been expanding their services to encompass more roadways – now even down to smaller arterials and local roads. In addition to an API for speed and travel time data, websites and associated analytic services are also available. Speed and travel time data can be viewed in real time as a thematic map, as well as historical data be accessed for research, performance metrics, and analysis purposes.

Commercial vehicle probe traffic data differs between vendors in a number of ways. A discussion of the primary differences is described below:

- Each vendor estimates traffic conditions from a unique set or collection of base probe vehicles. Some are more heavily weighted by fleet telematics, or by a group of vehicle manufacturers with built in mapping and traffic data, or by smart phone applications.
    - Freight data (associated with fleet telematics) tends to be a smaller proportion of the overall traffic stream, but is very consistent over a 24-hour period (since freight moves at all hours of the day).
    - Availability of data derived chiefly from consumers tends to inform traffic conditions well during peak travel hours, but can be much less consistent during off-peak and overnight hours.

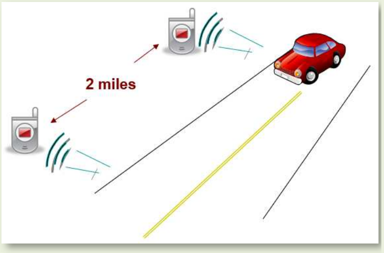
  There is no ideal source or ideal mix of data. Each vendor uses a mix of data from various sources, some more dependent on one source than another.

- A commonality of all vendors is that they report the speed on a roadway segment. The most common industry standard way of reporting segment data is the use of Traffic Message Channel (TMC) codes. The TMC segments grew out of the traffic reporting industry as a way to standardize roadway locations, making use of limited data capacity on radio side-bands. The I-95 Corridor Coalition released a report that details the history and usage of TMCs, which can be found here: https://i95coalition.org/wp-content/uploads/2015/02/TMC_White_Paper-Final.pdf?x70560. TMC segmentation data is typically from intersection to intersection and covers most major roadways and highways. TMCs were adopted by the National Performance Measures Research Data Set (NPMRDS) procured by the US Department of Transportation for the MAP-21 performance measures.

- In addition to supporting TMCs, whose segments tend to be long as it is primarily intersection to intersection, each vendor offers data on more highly granular networks whose segments lengths are shorter, allowing the agency to identify where queues and congestion begin and end. However, each segmentation method is proprietary and it is the responsibility of the agency to transform the data into the agency's map, network, or linear referencing system as needed, a process called conflation.

- Lastly, how do agencies know the data is accurate? The I-95 Corridor Coalition has proactively provided quality control of commercial vehicle probe data within its *Vehicle Probe Project*, and continues to do so today. In short, commercial vehicle probe data has been shown to be very accurate for freeways and other roadways without signal control. Probe data accuracy for signal-controlled roadways is not as accurate, due to queuing at traffic light, but has improved significantly in recent years, detecting a high majority of traffic slow-downs. I-95 Corridor Coalition summaries of probe data fidelity are available on the Coalition's VPP Marketplace page under the Data Validation tab located here: https://i95coalition.org/projects/vpp-marketplace/

Re-identification Data



- Bluetooth and WiFi
- Should be viewed as travel time sensors (as opposed to speed sensors). Such data is needed for travel time or O-D studies.
- Re-identification is typically used as ground truth for validating accuracy of sources of travel time data (such as commercial probe data).
- **Useful Applications: Travel Time on Signs, Travel Time Validation, Signal Performance Studies, Origin-Destination Studies, Smart Work Zones**

4.2 Vehicle Re-identification Data

Vehicle re-identification data is often referred to as Bluetooth traffic data, because it was first implemented using the Bluetooth communication protocol. It has since been expanded to also include the WiFi communication protocol, and is now referred to generically as re-identification data. The operating principle is best understood from an analogy with either toll-tags or license plates. To measure the travel time between two locations, a vehicle can be observed at an upstream and downstream location – noting the exact time of observation at both locations. The difference in travel time divided by the distance between the two locations provides a speed estimate for traffic. This method of collecting travel times for vehicles can be accomplished using license plate numbers, or electronically with toll-tag identifiers. Since 2007, it can now be accomplished using digital communications identifiers (namely Bluetooth and WiFi) at highway speeds. Using the digital identifiers (sometimes referred to as MAC addresses) in Bluetooth and WiFi transmissions, travel times can cost-effectively be measured on any roadway. The Bluetooth and WiFi wireless communication protocols require devices to broadcast their presence periodically – figuratively shouting 'I am here and this is my address' – so that other devices in the vicinity are aware of their presence, and can direct communications to them using the device's MAC address. Using MAC address matching is directly analogous to matching license plates, or toll tag identifiers.

Although there are multiple Bluetooth and WiFi re-identification vendors, and each implements the technology in a slightly different way, they all share some commonalities:

- Travel times are directly measured from a portion of the traffic stream. Early Bluetooth re-identification was found to sample approximately 5% of the traffic stream (1 in 20 vehicles). Modern implementations with Bluetooth and/or WiFi have been reported as high as 25% - though with great variation. The percentage of traffic observed with either Bluetooth or WiFi is dependent not only on the technology, but also the environment. Detection percentage is dependent on speed of traffic, the location of the radio equipment relative to the roadway, and the receive and transmit power of the digital radios.

- The amount of usable information derived from re-identification data is also a function of the complexity of the corridor or network. If there are on-ramps, off-ramps, intersections, side-roads, or curb-cuts between the upstream and downstream locations, a vehicle may be observed at one location but not at the other. In other words, the sampling percentage is only applicable to through traffic.
- As a general rule, a Bluetooth or WiFi re-identification sensor can only isolate a vehicle location to within approximately 300 feet, which is the approximate distance that Bluetooth and WiFi are regulated within. The physics of detection, including sensor scan times and position uncertainty, are beyond the scope of this paper. A generally accepted rule of thumb is to space re-identification detectors at least one minute of travel time apart to gain acceptable accuracy. A spacing of 2-5 minutes travel time separation is considered as ideal. There is really no upper limit to placement of Bluetooth sensors based on the physics of detection. The complexity of the network, need for minimum latency (a travel time datum is only received after the vehicle passes the downstream sensor), and needs of the application generally govern maximum spacing.
- Privacy questions often arise with use of MAC re-identification. MAC identifiers are unique like license plates, but are arbitrary. Unlike license plates, there exists no master database to track a MAC identifier to the device owner. Also, some devices periodically rotate MAC identifiers to prevent use of the identifiers as a tracking mechanism.
- Because Bluetooth and WiFi re-identification directly samples travel time, this process works equally as well on signalized arterials as it does on freeways. As such, application of re-identification data for signal performance were some of the initial applications of the technology.

Because re-identification technology directly samples travel time, one of the first uses of Bluetooth re-identification technology was to provide the reference (ground-truth) data to validate commercial probe data for the Vehicle Probe Project, and continues to be used in that capacity today. More information on validation processes can be found under the Data Validation tab of the I-95 Corridor Coalition VPP Marketplace site here: https://i95coalition.org/projects/vpp-marketplace/

A roadway jurisdiction can either directly purchase the equipment or contract for services. Correct design and placement of sensors is critical to obtain quality data. All vendors provide software for processing (filtering and matching) re-identification data.

Bluetooth can also be used for Origin-Destination (O-D) studies, but requires additional technical expertise. When used for O-D studies, a stable and calibrated re-identification sample rate is critical for O-D accuracy.

HRCD (High-Resolution Controller Data)

- HRCD and the corresponding Automated Traffic Signal Performance Measures (ATSPMs) are in the domain of traffic signal engineers.
- Signal upgrades should include consideration for acquisition and processing of HRCD and ASTPMs.
- **Useful Applications: Signal Performance Studies (ATSPMs)**

4.3 High-Resolution Controller Data

High-Resolution Controller Data (HRCD) is the timestamped data collected from the traffic controller that includes all vehicle detection data, service requests and phase change events recorded to 1/10 of a second. Such data has existed for decades in modern (digital) signal control cabinets, but until recently, has been too costly to transmit, store, and process at a central location. Enabled by an enumeration standard put forth by researchers at Purdue University, HRCD data has quickly been adopted by the majority of traffic signal control equipment manufacturers. As with other technologies, the proliferation of higher bandwidth telecommunications (either wireless, cellular, landline, or fiber connections) has enabled this data collection. Increased data and processing power of computers have allowed for central storage and analytics.

HRCD cannot be discussed without the corresponding advances in performance analytics techniques that it has enabled. These new performance measures, referred to as Automated Traffic Signal Performance Measures (ATSPMs) are metrics that operate on harvested HRCD to calculate and visualize the performance of a signalized intersection at a level of detail and accuracy previously not attainable.

HRDC requires upgrades to the controller equipment to collect and transmit the necessary data to a central processing location. The extent of data that is available is predicated on the sensors deployed at the signal. If the signal is simply pretimed, with no approach or stop bar sensors, only phase change data will be available. The power of HRDC is realized when adequate sensing is available at the intersection and that data is analyzed in parallel to yield highly effective performance metrics. These metrics include coordination diagrams, approach delay, pedestrian metrics, traffic counts, arrival approach volumes and arrival on red diagram as well as pre-emption calls. An in-depth, comprehensive discussion of the HRCD combined with ATSPMs is available from Purdue University entitled 'Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach' - https://docs.lib.purdue.edu/jtrpaffdocs/3/

In addition to ASTPMS enabled by HRCD, vehicle travel time distributions are becoming an increasingly utilized tool to understand the flow of complex traffic situations such as signalized corridors, as well as highly congested freeways. Enabled by Bluetooth re-identification data, and since extended to other technologies that can directly sample vehicle travel time, travel time distributions are leveraged through use of Histograms and Cumulative Frequency Diagrams. Visualizations depict the distribution of travel times along a corridor reflecting both average travel times as well as the reliability of the corridor during a selected time interval.

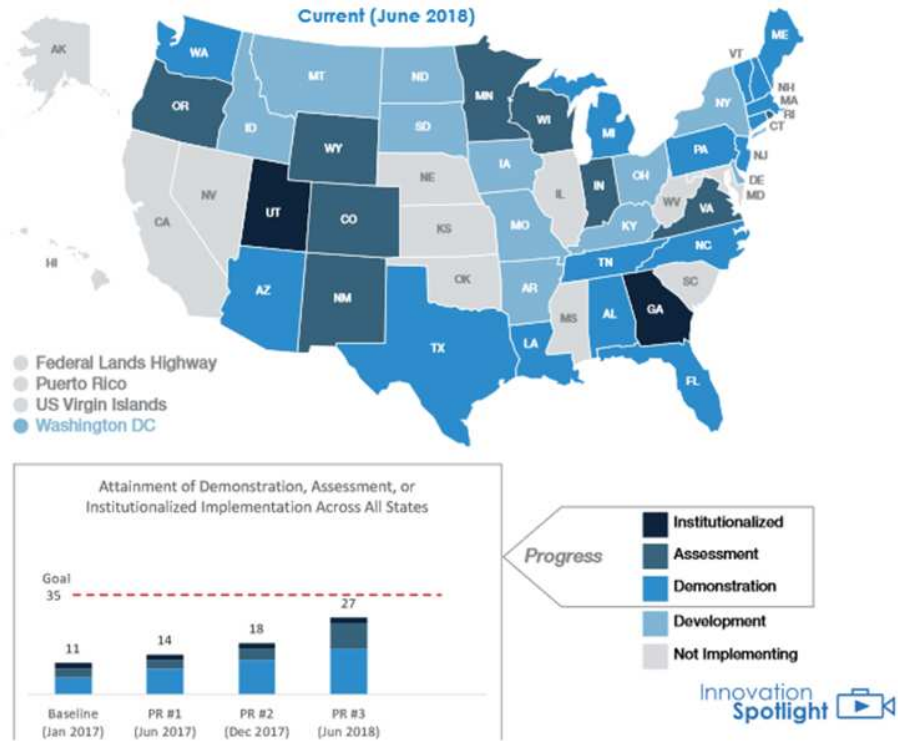


Figure 2 – National Progress on HRCD and ATSPM institutionalization (USDOT, 2018)

4.4 Cell Phone Geo-Location Data

Cell phone geo-location data (or CPGLD), derives travel patterns by observing cell tower traffic. Unlike commercial probe data which obtains information from self-reporting-GPS enabled devices, CPGLD gains information by observing the pattern created as a mobile cell phone (or smart phone) user traverses the system of cell towers. As the cell phone is handed off from tower to tower, a rough approximation of the trajectory and location of each cell phone can be inferred. Using a rule of thumb that a cell phone tower services an area roughly two miles in radius, such data can create a trajectory trail, as well as enable observation of origin and destinations patterns appropriate to planning level purposes, such as transportation analysis zones (TAZs). Also, as cell tower traffic typically increases with population density, requiring greater cell tower density (and thus higher spatial resolution), CPGLD tend to provide data of sufficient spatial resolution to support regional and state planning functions.

The size of the sample is dependent on the carrier. Previously, the representativeness of CPGLD data was brought into question. With cell phone and smart phone adoption rates soaring to 96% and 81% of the adult population, respectively, this is becoming less of an issue (Pew, 2019). However, the authors are not aware of any literature that directly addresses the ‘representative sample’ issue. The enormous quantity of information available from even a 10% to 20% sample far surpasses what could be done with manual survey even a decade ago.

State DOTs are increasingly using CPGLD data for planning purposes, providing substantial cost savings over traditional survey methods, as well as sample size improvements. Also, time savings cannot be ignored. Analysis requests from legislators that once required a year or more to plan, deploy and analyze can now be accomplished in a matter of weeks.

5.0 Applications, Applications, Applications

This section explores success stories of how the data technologies described in Section 4.0 were harnessed for significant benefit. These applications' stories are by no means exhaustive, but rather are chosen to highlight the breadth of usefulness for which the new traffic data technologies have been leveraged. These highlights are kept as brief as possible to illustrate value, as well as provide context for use.

5.1 Travel Time on Signs – the App that Ushered in Probe Data

The I-95 Corridor Coalition's Vehicle Probe Project, initiated in 2006 was a multi-state initiative to develop a traffic monitoring system based on new methods of collecting traffic data at scale, which resulted in one of the initial procurements of Commercial Probe Data. The first major application of that data came from the State of Maryland's Highway Administration (SHA). Posting anticipated travel time on signs was pushing state of the art for highway operations at the time. Sensors were typically deployed every ¼ to ½ mile, and the information brought into a central server, filtered, processed, and resulting estimated travel times posted to signs. Maryland SHA had plans for a major rollout of the technology that would take multiple years, and an investment of millions of dollars in permanent sensors. When the initial Vehicle Probe Project became operational in 2008, and the validation showed acceptable accuracy on freeways, the Maryland SHA capitalized on the project. They saved millions of dollars and many years by rolling out the program in a much shorter time frame (deployed by 2010) using probe data instead of a vast network of sensors. The resulting system is still in use today. Its success provided confidence in the fidelity of the data for other business purposes within the Maryland SHA.

Figure 3 – Travel time on sign posted on I-495 Outer Loop in Washington, D.C. (Google Maps, 2019)

5.2 Probe Data at Work in Traffic Management Centers

Probe traffic data in various varieties soon found its way into highway operations center, being used as a tool to capture network level traffic flow alongside site specific CCTV feeds. The image below is now typical in many Traffic Management Centers (TMCs) on the east coast and throughout the nation. A network view of the traffic (see map with red/yellow/green color coding in figure) at the network level informed by probe vehicle data provides the overall common operational views, while a myriad of CCTV cameras provides detailed views of incidents and slowdowns.

In fact, a network view of this manner within Maryland soon after the probe data network went live informed of a slowdown in the western part of the state on a minor highway. The operator, dubious that a slowdown of any nature could occur in the state's western rural regions where traffic is relatively light, investigated the incident and discovered that the probe data was reflecting slowdowns associated with a construction zone. This event ultimately added to the confidence in the probe data system.

Figure 4 – An example of probe data in a TMC - the Florida District Four Traffic Control Room – traffic speeds from probe data displayed alongside CCTV video of trouble spots (Smith, 2019)

5.3 Bluetooth - First Task of Validating Probe Data

One of the first large scale applications of Bluetooth re-identification data was to provide benchmark travel time data to validate Commercial Probe Data as part of the I-95 Vehicle Probe Project (VPP). As shared in Section 4.1, prevailing practice for travel time estimates was data from speed sensors placed periodically along the roadway. Although such data was available in a few locations along the I-95 Corridor, it was not prevalent, and would have significantly restricted the locations where the VPP could be tested. Also, travel time from periodic speed sensors provided only an estimation of travel time. As traffic slowed to a crawl, significant errors could occur. For these reasons, the validation program turned to Bluetooth re-identification technology that could be temporarily deployed anywhere within the VPP coverage region, and collect data directly measuring ground truth travel time, highly accurate even during significant congestion. A sample result of the ensuring results comparing the VPP commercial Probe Data with Bluetooth re-identification data is shown in the figure below. The cloud of blue data represents ground truth travel times (presented as travel speeds for the segment) recorded by Bluetooth equipment. It is overlaid on top of the red points, that is travel time (again displayed as average speed) reported by probe data vendors, showing good agreement. Validation activities comparing Bluetooth travel times to that reported in probe data feeds from vendors has been ongoing since 2007, and continues to this day in the I-95 Corridor Coalition. It recently has been expanded to cover National Performance Measures Research Data Set (NPMRDS) data accuracy as well.

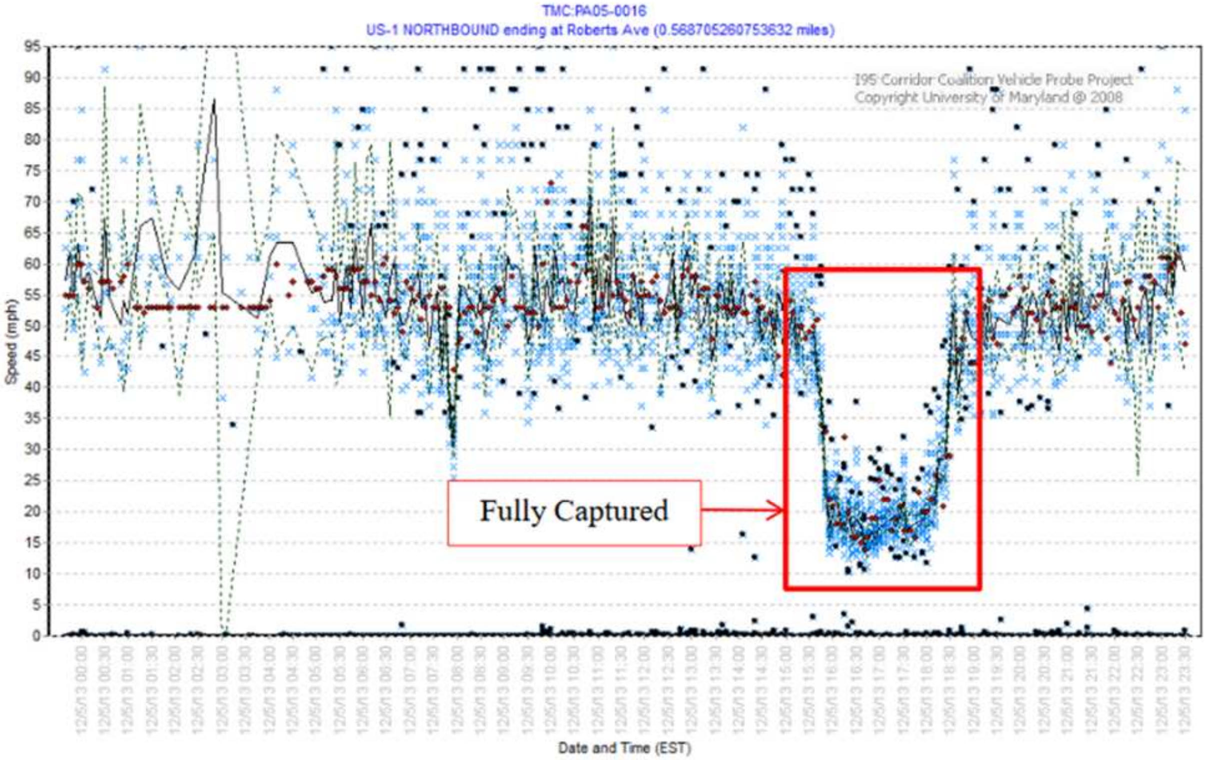


Figure 5 – Graphing probe data validation (VPP Validation Program, 2015)

5.4 Bluetooth - Assessment of Signal Performance

As Bluetooth directly measures travel time, it provided insight into arterial traffic flow that was difficult to measure without significant investment in floating cars. Floating car studies involve hiring drivers to periodically traverse a corridor recording their location and time. Originally performed with stopwatches, GPS equipment has made them easier to perform, but remained costly due to labor. Bluetooth sensors placed upstream and downstream from a signal could provide similar data, but with minimal cost and disruption. One of the first applications was north of Baltimore on MD 24, also called Vietnam Veterans Memorial Highway. The Maryland State Highway Administration implemented signal retiming in the corridor in late February of 2009. Bluetooth equipment placed along the corridor captured before and after vehicle travel times.

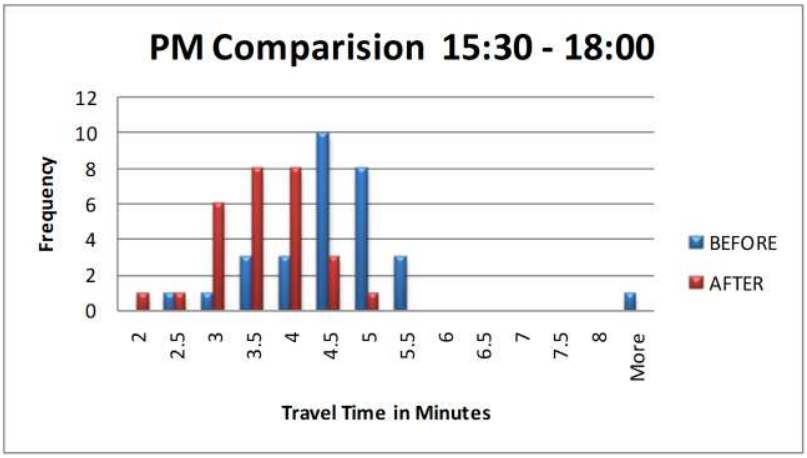


Figure 6 – a histogram of travel times experienced by vehicles before (in blue) and after (in red) (Young, 2012)

Now such studies for before and after are commonplace using temporary deployment of re-identification equipment. In some places Bluetooth equipment is permanently deployed, providing detailed travel times in signalized corridors.

5.5 HRCD for Automated Traffic Signals Performance Measures

Utah DOT was one of the first jurisdictions to invest in the use of high-resolution controller data for automated signal performance measures. Beginning in 2013, Utah DOT began to invest resources into developing a system for automatically downloading the data and providing the performance measures using a web-based system which is in full use today. The software for the website is open-source code, and has served to enable several other implementations of the technology. A sample of the Purdue Coordination drawn from the active Utah DOT web analysis system is shown below.

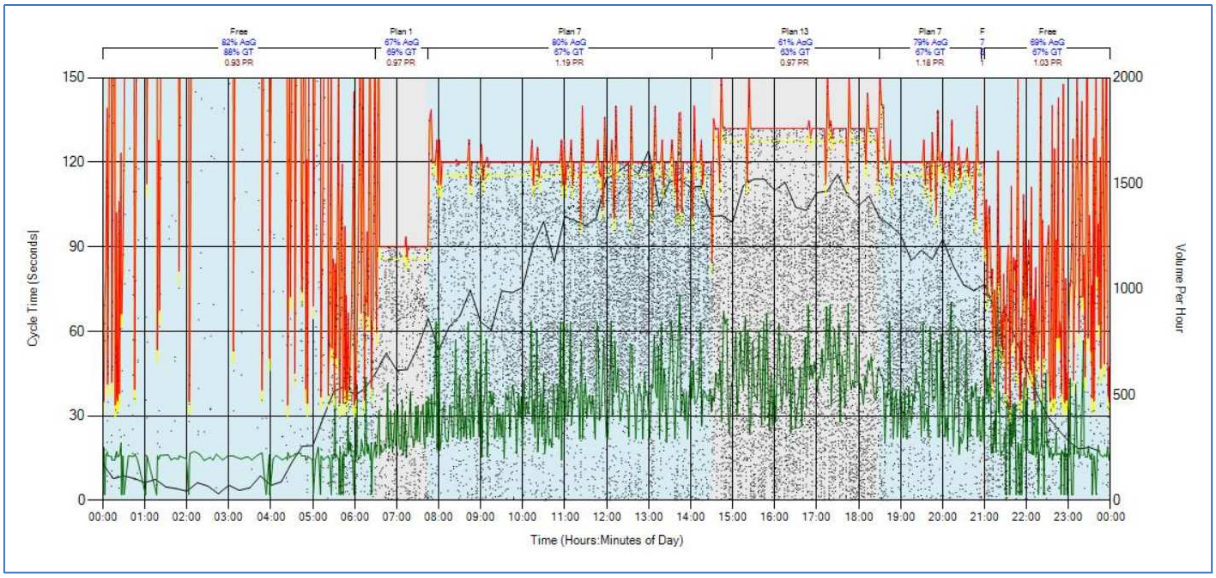


Figure 7 – An example of a Purdue Coordination Diagram generated with HRCD (Utah DOT, 2019)

Since 2013, approximately half the states have some type of active ASTPM project underway.

5.6 Smart Work Zones

Smart work zone technology relies on leveraging speed, volume, or presence data and portable changeable message signs (PCMS) to improve work zone safety and mitigate traffic impacts. Smart work zones are flexible because they can rely on a variety of technology – from more traditional microwave and side-fire radar to newer, more cost-effective techniques with probe data. Smart work zones are still an emerging practice, but several agencies across the U.S. have implemented programs to great effect. Here are some of the ways data (merged between outsource probe and traditional sensors) can be used to improve work zones.

- Travel Time or Travel Delay Information – Travel Time systems are designed to inform motorists of traffic conditions via estimated travel times through the work zone or to other exits or landmarks.
- Queue Detection **-** Queue Detection systems identify congestion or slowed traffic in a work zone approach and alert travelers farther upstream.
- Dynamic Lane Merge **-** Dynamic Merging systems are systems designed to adapt merging conditions to current traffic demand – early merges for low volume/high speed and late merges for high volume/low speed.
- Work Vehicle Entry/Exit Warnings: Work Vehicle Entry and Exit Warning systems provide a warning to roadway users when a construction vehicle is slowing down to enter a work zone or accelerating from a work zone into travel lanes

5.7 NPMRDS and MAP-21 Performance Measures

The National Performance Measurement Research Data Set (NPMRDS) is an archived probe data set with free access for all public agencies and their consultants, paid for by the FHWA. It provides speed and travel time on TMC segments that traverse the National Highway System (NHS). Unlike many other commercial probe data products, it only reports actual speeds from observed probes and it does not impute speed based or historical data during periods of low demand. NPMRDS can be used in a similar fashion to other probe data products, but it is unique in that it is the primary method of annual reporting for Moving Ahead for Progress in the 21st Century Act (MAP-21) submissions. Specific performance measures, detailed below, are required to be generated by each state for their NHS roadway each year.

- Peak hour excessive delay (PHED)
- Level of Travel Time Reliability (LOTTR) on Interstates
- LOTTR Non-Interstates
- Truck Travel Time Reliability (TTTR)

Emerging, Leading Edge Technologies

Trajectory Data
- Waypoint data every 1 second
- OD studies, arterial analysis, freight studies
- Market-ready by 2021

Estimated Volumes from Probe Data
- AADTs, turning movements, vehicle/hour
- I-95CC Validation
- Market-ready by 2021

6.0 Emerging, Leading Edge Technologies

Technologies mentioned in this section are referred to as emerging, meaning that they show promise of good value to agencies, but are not fully vetted, or have not fully matured or reached a level of a consistent expectancy of accuracy and value. As such, any agency that procures or purchases such data services inherently takes on increased risk. That said, the authors include the following as areas to watch in the next 24 to 36 months as increased data resources are brought to bear in the traffic data market.

6.1 Vehicle Trajectory Data

Vehicle trajectory data, sometimes referred to as 'Trip Data', are samples of the waypoints of a trip. Trajectory data was made possible through wide-spread use of GPS equipment which typically log vehicle position (in the form of latitude and longitude) once per second. As the reporting frequency of commercial probe data increased, particularly for probes reporting at internals of one minute or less, the position data chained together in a sequence (often referred to as a 'bread-crumb-trail') became available to researchers at scale. Trajectory data has long been available for research through dedicated GPS data collection activities, but at substantial cost. Commercial probe data (and their respective data sources) now obtain trajectory data samples at various reporting frequencies ranging from 1 second and upward. Such data has been utilized in a number of first-of-its-kind applications in the last 24 months. Examples of applications include:

- Highly granular origin-destination studies
- Arterial signal assessment
- Select link analysis
- Freight shed analysis

In each of these early applications, all of which are in some form of R&D, the scale, accuracy, or value proposition over that which the industry currently considers 'state-of-the-art' has been greatly surpassed. The authors expect that applications of trajectory data will mature quickly in the ensuing 24 months.

6.2 Estimating Volume with Probe Data

Closely associated with trajectory data applications is the prospect of having accurate volume estimates. Commercial probe data has proven to scale network wide, providing valid speed and travel time data. As the sample sizes of commercial probe data for vendors continues to grow, they are achieving statistical significance to also estimate volume on a roadway – be it AADT, turning movements at intersections, or real-time volume counts. The Coalition sponsored a research project that provided proof-of-concept, and a follow-on funded research initiative is underway to produce prototype commercial systems. Vendors are already offering representative AADT volumes, and the USDOT is undergoing a pooled fund study to better understand accuracy, standards, quality, and application potential of such methods. Keep an eye on this quickly developing space.

7.0 Conclusions and Recommendations

The strategies included in this report represent our current understanding of how to optimally monitor traffic. Traditional traffic sensors operating from the inductive loop principle will continue to be used for sight specific, fully agency owned and controlled applications. However, new data technologies have emerged that allow traffic monitoring value propositions not achievable with sensor-based approaches. These advanced include:

- Commercial probe data provides a highly scalable, cost effective, 24x7x365 solution to monitor traffic speed and travel time across all but the lowest functional class roadways
- Re-identification provides a means to directly and accurate measure travel time, as opposed to inferring travel time from periodic speed sensors, which is subject to large errors during peak congestion.
- High speed data communications combined with modern data storage and computational efficiencies has enabled High Resolution Controller Data, which in turn enabled more effective signal performance assessment methodology (ATSPMs) which is quickly gaining traction across the US.
- Cell Phone Geo-location Data derived from the cell tower handoff patterns provides origin-destination travel patterns at resolution, scale, and sample size that cannot be ignored for transportation planning purposes.

While there will likely be a place for traditional sensors for years to come, probe data and other detection strategies enabled by new technology will continue to evolve and grow into more complete, and cost-effective strategies to fully observe the state of traffic on our roadways. The data collection technologies, methods and applications presented within this document are intended to provide the reader with practical guidance to harness emerging traffic data technologies for the benefit of their agencies.

Examining the current and evolving state of innovation around the industry, future traffic initiatives look slated to revolve around collecting very rich data sets, and will leverage machine learning and AI to combine and synthesize new performance measures as well as highly detailed and accurate real-time monitoring. Investing in a comprehensive traffic monitoring data program at your agency, including partnering with entities that have strong big-data experience, is recommended. As more mature connected and autonomous vehicle technologies arrive on the market, entirely new data sources will become available for agencies that report not only location and speed, but also parameters such as temperature, precipitation, roadway traction, humidity, road roughness, and many other parameters.

References

U.S. Department of Transportation - FHWA. Every Day Counts: An Innovation Partnership with States – EDC-4 Progress Report #2 January-June 2018. U.S. Department of Transportation - FHWA.

Google Maps, 2019. Travel time on sign posted on I-495 Outer Loop in Washington, D.C. http://maps.google.com. Accessed July 22, 2019.

Pew Research Center. Mobile Fact Sheet. Demographics of Mobile Device Ownership and Adoption in the United States | Pew Research Center. 2019. https://www.pewinternet.org/fact-sheet/mobile/ Accessed July 22, 2019.

Smith, Daniel. Dynamic Video Wall and RITIS for Improved Operational Awareness. Florida DOT and I-95 Corridor Coalition. 2019.

Utah DOT. UDOT Automated Traffic Signal Performance Measures – Automated Traffic Signal Performance Measures. Utah DOT. http://udottraffic.utah.gov/atspm/. Accessed July 18, 2019.

VPP Validation Program. I-95 Corridor Coalition Vehicle Probe Project Validation of Arterial Probe Data. I-95 Corridor Coalition. 2015.

Young, Stanley. Bluetooth Traffic Detectors for Use as Permanently Installed Travel Time Instruments. Publication MD-12-SP909B4D. Maryland DOT/SHA, 2012.