Journal of Traffic and Transportation Management
Volume 3, No. 1 (2021) pp. 01-09

On the use of active mobile and stationary devices for detailed traffic data collection: A simulation-based evaluation

Johan Holmgren[a], Henrik Fredriksson[b] *, Mattias Dahl[b]

[a]*Department of Computer Science and Media Technology, Malmö, Sweden, 20506*
[b]*Department of Mathematics and Natural Sciences, Blekinge Institute of Technology, Karlskrona, Sweden, 37179*

Abstract

The process of collecting traffic data is a key component to evaluate the current state of a transportation network and to analyze movements of vehicles. In this paper, we argue that both active stationary and mobile measurement devices should be taken into account for high-quality traffic data with sufficient geographic coverage. Stationary devices are able to collect data over time at certain locations in the network and mobile devices are able to gather data over large geographic regions. Hence, the two types of measurement devices have complementary properties and should be used in conjunction with each other in the data collection process. To evaluate the complementary characteristics of stationary and mobile devices for traffic data collection, we present a traffic simulation model, which we use to study the share of successfully identified vehicles when using both types of devices with varying identification rate. The results from our simulation study, using freight transport in southern Sweden, shows that the share of successfully identified vehicles can be significantly improved by using both stationary and mobile measurement devices.

Keywords: *Traffic data collection; stationary devices; mobile devices; traffic simulation*

1. Introduction

It is widely known that transportation plays a fundamental role for the economic and social development [1, 2]. For example, transportation allows people to commute to and from workplaces and educational institutes and enables transportation of goods. In order to maintain an efficient transportation infrastructure and to plan for present and future needs, there is a need to continuously gather information about the current state of the transport network. Thus, an important task for traffic authorities and other responsible actors is to collect up-to-date and accurate data about the traffic in order to evaluate the performance of the network [3]. Even though large amounts of resources are spent collecting, processing and analyzing traffic data, there is typically lack of knowledge on how individual vehicles travel over larger areas. For example, there is often a need to gain deeper knowledge of the movements of trucks carrying hazardous materials and to trace vehicles traveling on networks where road user charging is applied [4, 5]. This knowledge gap mainly stems from the inability, of the dominating types of technologies that are currently in use for collecting traffic data, to identify vehicles.

One of the main objectives of collecting data is to accurately reflect the real-world state about the traffic in the network for further analysis. The data is vital for planners and traffic analysts to efficiently manage and coordinate the usage of the network and to ensure that adequate safety measures are taken. Traffic data has a wide range of applications. For instance, traffic data can be used to estimate travel-times, predict congestion, and to find accident-prone locations in the network [6-8].

The purpose of this work is to evaluate how active stationary and mobile measurement devices can be combined for traffic data collection in a road transport network. An active measurement device is equipped with technology able to detect and identify vehicles, for instance, automatic license plate number recognition. Due the complementary characteristics of stationary and mobile devices, we argue that they should be used in conjunction with each other for efficient traffic data collection. To evaluate the idea of using both types of measurement devices for traffic data collection, we developed a traffic simulation model, which we applied in order to simulate movements of freight trucks in southern Sweden. The simulation model is able to reproduce traffic flows on individual road segments, and it allows us in a cost-effective way, to investigate different scenarios with varying identification rate to gain insight into how stationary and mobile devices complement each other. In our scenario study, we study the share of the transportation fleet that was successfully detected and identified, as well as the share of freight trucks that was successfully detected and identified on the first link and the last

* Corresponding author. Tel.: +46455385548
E-mail: henrik.fredriksson@bth.se

DOI: 10.5383/JTTM.03.01.001

link along their journeys. We consider the latter information to be relevant since it can be used to estimate the travel demand between locations in the network (e.g., OD-matrix estimation). The current paper extends the paper by Holmgren et al. [9]. In particular, in the current paper we provide additional details about the mathematical model and new simulation results useful for vehicle movement analysis.

The remainder of the paper is organized as follows. In Section 2, we discuss the key properties of stationary and mobile devices for traffic data collection. Section 3 describes the mathematical model of our traffic simulation model, which we used in order to evaluate our idea on traffic data collection in the simulation study presented in Section 4. Finally, we provide some concluding remarks in Section 5.

2. Traffic data collection overview

Related work includes various studies on how to integrate traffic data from multiple measurement sources. The movements of freight trucks can, for instance, be analyzed by the means of data obtained from automatic number plate recognition (ANPR) cameras and GPS-data from the on-board units of the trucks [10]. The state of traffic (e.g., traffic density, vehicle speeds and flows) can be estimated by employing vehicle-to-infrastructure communication using connected vehicles and traffic counts from stationary devices, and from mobile devices and loop detector data using Kalman filtering [11, 12].

In the remainder of this section, we discuss some important properties of stationary and mobile measurement technologies. For example, the two properties of geographic coverage, which is higher for mobile devices and the possibility to successfully detect and identify vehicles, which we argue is better for stationary devices, makes it relevant to consider both types for detailed real-time traffic data collection [13].

When a vehicle is detected by an active measurement device, either stationary or mobile, the device should be able identify the detected vehicle, and possibly the position and time of the detection, and communicate the data to the responsible authority [14]. The identification could, for instance be made by the means of radio-frequency identification (RFID) or ANPR cameras.

Stationary devices have fixed placements, meaning that they are only able to observe those vehicles that pass the device. However, since they continuously operate in the same location, stationary devices can discover patterns of traffic flows over time and discover anomalies that deviates from prior measurements. Mobile devices, such as probe vehicles have a significantly larger *geographic coverage* than stationary devices, since they drive around in the network and collect data wherever they happen to drive [15, 16]. On the other hand, as their location is constantly changing, probe vehicles do usually not collect detailed data about the traffic at specific locations over time.

The *identification rate* of a measurement device is the ability to successfully detect and identify a vehicle, which is to a large extent affected by external factors. For instance, when a device is equipped with camera, a high identification rate requires that the device is sufficiently close to the traffic, and that there are no obstacles between the camera and the vehicle to be observed. A stationary camera-based device can be placed above the road to ensure that no obstacle cover the traffic to be detected, meaning that they are able to detect vehicles even if the traffic density is high. Probe vehicles have less possibility to control the external conditions which means they have to operate under conditions that are often less favorable than for stationary devices. For instance, other vehicles may block the view of the camera of a probe vehicle if the traffic density is high on roads with multiple lanes, meaning that the probe vehicle is unable to detect oncoming vehicles on the opposite side of the road.

We argue that the *cost* in general will be lower for mobile devices than for stationary devices. Today, modern cell phones are typically equipped with technologies suitable for vehicle detection, such as camera, Global Positioning System, and Bluetooth [17, 18]. This means that virtually any kind of vehicle could serve as probe vehicle. As stationary devices are placed along or above the road, they may require both building permits, and operating personnel [17]. Here, we would like to advocate for the potential of using already existing equipment in the traffic infrastructure for data collection. For instance, speed cameras possess the ability to identify vehicles, and could serve multiple purposes, by installing software for data collection.

Since it may be costly to build new stationary devices, we find it reasonable to assume that the *ownership* of stationary devices mainly belongs to public authorities. Hypothetically, any vehicle can operate as a probe vehicle, which also means that privately owned vehicles can be used for traffic data collection; however, the willingness of private actors to use their vehicles as probe vehicles depends to a large extent on their incentives. A possible incentive is the will to contribute with data that can be used to analyze traffic, and in the longer perspective, contribute to the development of the transport network.

The *number of devices* to use for traffic data collection depends on the budgetary constraints of building and maintaining stationary devices, as well as the number of available probe vehicles. Since we expect the cost for stationary devices to be relatively high, we believe that the possibility to build new stationary devices is limited. However, as previously discussed, utilizing the possible synergies between different ways of using technologies currently in use, may reduce the need for new stationary devices. The number of mobile devices mainly depends on the incentives of drivers to let their vehicles act as probe vehicles.

The major issues of using privately owned vehicles as probe vehicles for data collection are *privacy* and *integrity*. Which type of data that is collected, how long it is stored, and data encryption for secure transmissions are issues that should be addressed to ensure the data collection process is handled correctly. Therefore, there is a need of a standardized policy framework for the involved actors (e.g., mobile phone operators and authorities). We suggest that public cryptography, where the measurement devices make use of the public encryption key of the responsible authority, can be used for secure data transmission and encryption of the signatures of probe vehicles.

3. The simulation model

We describe a transport network by a set of nodes N and a set of links A. The nodes represent intersections and other locations in the network connecting roads. The links represent road segments in the network. For each node $n \in N$, we let O_n and I_n denote the set of outgoing and incoming links from and to n, respectively. The travel time for link a_k is denoted by τ_k, and $\varphi_k(t)$ is the flow of vehicles on that link at time t. The simulation model used in this paper utilize hourly flows over a 24-hour period, and for this end we use periodic step functions to describe $\varphi_k(t)$. The functions $\varphi_k(t)$ are constant over a one-hour interval and have an overall period of 24 hours. Since the functions $\varphi_k(t)$ are constant during one-hour intervals, we may without loss of generality assume that t is a discrete integer variable. We follow the common assumption that the number of arriving vehicles at a certain node is Poisson distributed [19].

The time from t_0 until the next time a vehicle enters a_k is a random variable X_k with the cumulative distribution function

$$F_{X_k}(t \mid t_0) = P(t_0 < X_k \le t) = \int_{t_0}^{t} f_{X_k}(\xi)\, d\xi. \tag{1}$$

The expected number of entering vehicles per unit of time t, is given by the intensity function

$$\begin{aligned}\varphi_k(t \mid t_0) &= \lim_{h\to 0} \frac{P(X_k \le t+h \mid X_k > t)}{h}\Big|_{t\ge t_0} \\ &= \lim_{h\to 0} \frac{P(t < X_k \le t+h)}{h\cdot P(X_k > t)}\Big|_{t\ge t_0} \\ &= \frac{f_{X_k}(t \mid t_0)}{1-F_{X_k}(t \mid t_0)},\end{aligned} \tag{2}$$

which yields the distribution function

$$F_{X_k}(t \mid t_0) = 1 - \exp\left(-\int_{t_0}^{t} \varphi_k(x)\, dx\right). \tag{3}$$

It follows that the expected number of vehicles passing node n during a time interval of length T is

$$\lambda = T \sum_{a_k \in O_n} \varphi_k, \tag{4}$$

meaning that the expected number of departures from node n during hour t can be calculated by

$$\delta_n(t) = \sum_{a_k \in I_n} \varphi_k(t), \tag{5}$$

and the time between departures is exponentially distributed. Since $\varphi_k(t)$ are constant during one-hour intervals, it follows that any vehicles reaching a node n_1 has a probability of

$$p = \frac{\varphi_j}{\sum_{a_k \in O_{n_1}} \varphi_k}, \tag{6}$$

to travel from n_1 to n_2, where j is such that $a_j = (n_1, n_2)$. The number of vehicles traveling directly from n_1 to n_2 follows the distribution $\mathrm{Bin}(X, p)$, where X is a random variable from the distribution $\mathrm{Po}(\lambda)$. From a practical perspective, it is reasonable to assume that the incoming flow to a node is the same as the outgoing flow, albeit with some time delays based on τ_k. In the model it is required that

$$\sum_{a_k \in I_n} \varphi_k(t - \tau_k) = \sum_{a_k \in O_n} \varphi_k(t) \tag{7}$$

holds for all $n \in N$, and all $t \in \mathbb{R}$. In equation (7), the left-hand side is the flow going into node n at time t, and the right-hand side is the flow going out from node n at time t. From a logistic point of view it is a challenging task to measure the traffic flows of a transportation network over multiple points simultaneously over a 24-hour period. Hence, traffic flow data over a 24-hour period may contain deviations such that nodes have larger inflow than outflow or vice-versa. This may occur when for instance the measurements are made by temporary pneumatic tubes that are moved around in the network according to some schedule, or when there is no measurement performed and the flow data is instead estimated. This calls for calibration of the data to satisfy equation (7). Calibration of the data can be performed by introducing loops at each node in the network. The loops are links of the form $a_k = (n_i, n_i)$, which act as "parking". As soon as a vehicle reaches a node, the vehicle will temporarily leave the simulated system with a particular probability and be "parked" for a certain amount of time. By introducing loops at each node, we can always find flows $\varphi_k(t)$ on the loops such that the system is balanced for each one-hour interval and node [20]. The loops also allow the model to handle the overall number of vehicles in the network, which vary hourly over the day. For simplicity we assign all loops a travel time $\tau_k = 1$ (the same length as the constant interval of the step function). The vehicle flow data (from the databases TFK and TIKK), provided by the Swedish Transport Administration, used in our simulation study did not satisfy equation (7) and required calibration as previously described.

4. Simulation study

The purpose of our simulation study is to evaluate and illustrate the idea of using stationary and mobile devices for collecting individual-level data about vehicles. Based on real traffic intensities on each road segment over a day, the aim of the simulation model is to reproduce the hourly traffic flows on each link. The main idea is to simulate the movements of vehicles over time, and to log each successful observation made by a measurement device. We assume that a vehicle can be identified by a stationary device when the vehicle passes the location of the device, and by a mobile device when the vehicle and a probe vehicle meet on the opposite links of a road.

We study the share of the simulated traffic for a selection of routes, which is successfully observed and identified by either a stationary or mobile device, for each of the 24 hours of the day with varying identification rate. This allows us to verify our idea of the complementary characteristics of stationary and mobile measurement devices. We also study the share of vehicles that is expected to be successfully observed and identified on both the first link and the last link on a route, as this provides deeper knowledge on how vehicles are expected to move over larger areas in the network.

4.1. Scenario description

We focus on freight transport since trucks may have fewer concerns regarding privacy issues and may therefore be more prone to operate as probe vehicles. The network used in our simulation study consists of 33 nodes and 102 bi-directed links, and it is a representation of a road traffic network in southern Sweden (see Fig. 1).

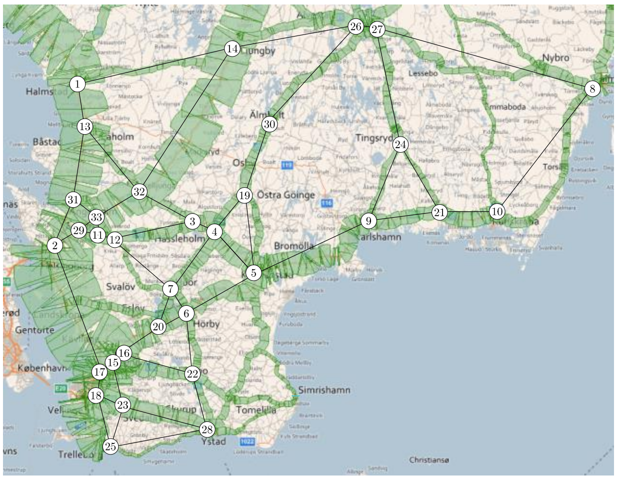

Fig. 1. The traffic network used in our simulation study on top of the traffic flow map. The traffic flow map was provided by the Swedish Transport Administration.

The selection of the road transport network to consider in our simulation study is based on the traffic flow map in Fig. 1 and the underlying truck flow data (average daily flows). In the simulation study, we let the speed cameras located in the network act as stationary devices, and freight trucks are used as probe vehicles. We assume that both the stationary and mobile measurement devices are active, meaning that they are capable of recording and communicating the observed vehicles' identities (e.g., license plate number), position, travel direction, and the time for the observation.

The choices of which nodes and links to include in our selected network was made to capture the major traffic flows in the considered region. The set of nodes contains locations where routes are expected to intersect, and where large traffic flows propagates in the network. This argues why some nodes are located very close to each other. For each of the 24 hours of a day, we extracted the corresponding distribution of the traffic volumes for each link and identified the locations of the speed cameras along the links. As far as we know, the sole purpose of speed cameras in Sweden is to observe and identify vehicles violating the speed limits. However, as discussed in Section 2, we assume that the functionality of speed cameras can be extended to also operate as devices for traffic data collection.

The considered region is rather densely populated with a high amount of traffic and we mainly included roads with high traffic volumes. This means that the density of speed cameras in our study is higher than in the whole of Sweden. For instance, there are large areas in Sweden, particularly in the northern part, where the traffic volumes are significantly lower, and the density of speed cameras are also lower. We mainly included roads where traffic volumes are high, which are also those roads that typically contain speed cameras. In total, the selected roads contain 121 speed cameras.

4.2. Route selection and link categorization

The number of routes in a transportation network can be very large and it may be a cumbersome job to analyze each one of them. Also, routes may share similar characteristics, which means that it may be sufficient to study a selection of routes with different properties to get an overview of the entire network. The aim of our route selection process is to identify a "minimum" set of routes to be included in the analysis of the output of the simulation study. We claim that a successful identification of a vehicle primarily depends on the number of devices it is expected to encounter on the link, which is 1) the number of speed cameras located along the link, and 2) the number of probe vehicles that are simultaneously driving on the opposite link. There are other external factors such as environment, current lightning and weather conditions that affect whether a vehicle is successfully identified. However, these external factors are not considered explicitly since we mainly believe that the number of measurement devices a vehicle is expected to encounter on its journey has a more significant impact on the possibility to be identified.

Since a route is a sequence of links connected by nodes or in some cases a single link, we argue that it is relevant and necessary to consider the number of links of a route and the characteristics of the links to define the main characteristics of a route. As the network in our study contains 33 nodes and 102 links, there are 102 routes of length one, 326 routes (i.e., valid link sequences) of length two, and 1050 routes of length three. We considered routes of length one, two and three, which we categorized based on partitions of the links according to the two above mentioned criteria.

As seen in Table 1, the speed cameras are not evenly distributed in the network as the number of cameras on the links varies between zero and eleven. There are 45 links with one or more cameras and 57 links without cameras. For the links equipped with at least one speed camera, the mean and median numbers of cameras are 2.69 and 2 respectively. Thus, we partitioned the links into two groups, links without speed camera, and links with at least one speed camera. The two group of links were further partitioned into subgroups, determined by the estimated number of expected meetings along the link. Here, it should be emphasized that when estimating the number of expected meetings, for simplification, we assumed that the number of departures on a link is uniformly distributed over each of the 24 hours of the day. This assumption is exclusive to the process of categorizing the routes, and not considered in the simulation model where the number of departures is assumed to be Poisson distributed. We further assumed that the estimated daily truck traffic volume of a link is identical to the truck traffic volume of its opposite link, i.e., we assume $\int_0^{24} \varphi_k(t)\,dt = \int_0^{24} \varphi_l(t)\,dt$ holds for opposite links $a_k = (n_i, n_j)$ and $a_l = (n_j, n_i)$.

For each link in the network, the traffic volumes vary over the day; in particular, the traffic volumes during night hours are always lower than the traffic volumes in the daytime. The peak hours when the traffic volumes are high, vary for the links, but in general, the hour distributions of the traffic flows are not essentially too different for the different links. We leave to future work to consider the daily variations. Still, we argue that we get a valid indication of how the number of meetings is expected to vary for the different links even if we do not consider the variations of the traffic volumes over the different hours of the day. Suppose that a_k and a_l are opposite links. At the time a truck enters a_l, the expected number of trucks that is at the same time driving on the opposite link a_k is $\frac{\tau_k}{24}\int_0^{24} \varphi_k(t)\,dt$. The expected number of trucks that enter a_k while the truck traverses a_l is $\frac{\tau_k}{24}\int_0^{24} \varphi_k(t)\,dt$. It follows that the number of trucks a vehicle is expected to meet when traveling on a_l is $2 \cdot \frac{\tau_k}{24}\int_0^{24} \varphi_k(t)\,dt$. Furthermore, during the link categorization process, we assumed that the expected traveling speed for all trucks on all links in the considered scenario is identical. Since the allowed speed limit for a heavy truck (vehicle transporting goods with a capacity above 3.5 tons) in Sweden is 80 km/h (or 90 km/h when traveling on motorways), we used an estimated average traveling speed of 68 km/h for all trucks over the links.

The number of vehicles that a truck traveling on a particular link is expected to meet, varies from 0.054 to 222, where the mean value is 25.5, and the median in 15.7. In addition, the cumulative distribution over the measure (see Fig. 2) clearly shows that the smaller values of expected meetings dominate.

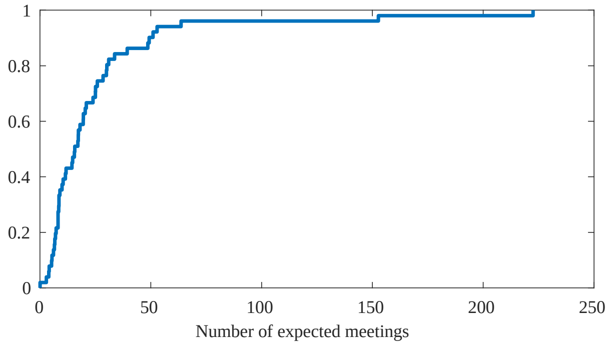

Fig. 2. Empirical cumulative distribution of the number of expected meetings.

For example, 98% of the values are below 200, 96.1% of the values are below 100, and 90.2% of the values are below 50. It can be also seen in the histogram over the expected number of meetings (see Fig. 3) that the smaller values are dominating.

As previously discussed, the links were partitioned into groups based on the number of expected meetings. We used different value ranges of the expected number of meetings depending on how many links a route contains.

Table 1. Links (node pairs), travel times, link lengths, and number of speed cameras along the links. Each pair of nodes in the table represents two links in opposite directions. For instance, the node pair 3, 32 represents the two links (3, 32) and (32, 3), where one speed camera is located along the link (3, 32) and two speed cameras along the opposite direction (32, 3), denoted 1/2 in the "Cameras" column.

Link (node pair)	Travel time (min)	Length (km)	Cameras (amount)	Link (node pair)	Travel time (min)	Length (km)	Cameras (amount)	Link (node pair)	Travel time (min)	Length (km)	Cameras (amount)
1, 13	11	20.0		7, 12	35	39.8		16, 20	13	22.0	
1, 14	47	66.3	2/2	7, 20	15	16.0	2/2	16, 22	23	31.6	3/4
2, 17	35	62.6	1/0	8, 10	59	80.2	11/11	17, 18	6	9.9	
2, 29	2	4.2	0/1	8, 27	71	101.0	2/2	18, 23	7	11.8	
2, 31	11	20.5		9, 21	18	28.7	0/1	18, 25	19	29.9	
3, 4	6	9.7		9, 24	24	34.8		19, 30	22	32.3	
3, 12	26	33.7	0/1	10, 21	17	25.0	0/1	21, 24	25	31.4	0/1
3, 32	20	26.7	1/2	11, 12	3	4.3		22, 28	20	25.3	2/1
4, 5	17	25.9	1/1	11, 29	7	10.9		23, 25	16	18.1	4/3
4, 7	23	31.0	3/3	11, 33	7	8.8	1/1	23, 28	24	36.1	
4, 19	13	18.8		13, 31	16	29.1		24, 27	37	48.9	3/3
5, 6	30	41.0	3/2	13, 32	31	39.6		25, 28	47	45.6	
5, 9	38	55.3	2/2	14, 26	38	58.3	1/1	26, 27	5	7.0	
5, 19	27	33.3		14, 32	38	71.2		29, 33	5	9.0	
6, 7	8	10.8	2/2	15, 15	2	3.1		26, 30	38	54.2	6/6
6, 20	8	12.0		15, 17	4	6.4		31, 33	12	14.3	2/3
6, 22	21	27.0	2/2	15, 23	17	20.7	4/5	32, 33	12	21.7	

Single links that was considered as a route was partitioned by the value ranges [0.0539, 8.54], [8.55, 20.9], and [21.0, 222] into six categories (each value range gave two route categories: one with and one without speed camera). The links that are a part of routes of length two or three, was partitioned by the value ranges [0.0539, 15.7] and [15.8, 222] into four categories. The link categories are shown in Table 2.

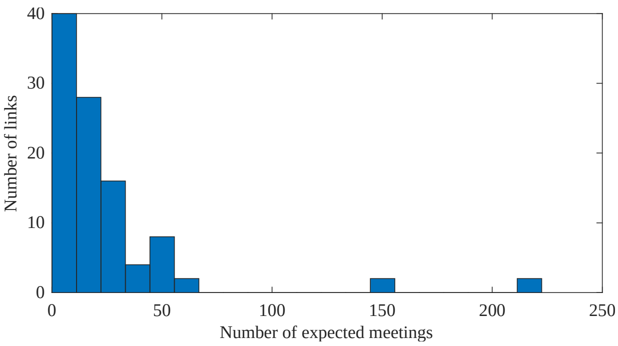

Fig. 3. Histogram over the number of expected meetings.

Table. 2. Link categories (1-6) for routes with one link, and link categories (7-10) for routes with two or three links.

Link category	Speed camera	Expected meetings (range)
1		[0.0539, 8.54]
2		[8.55, 20.9]
3		[21.0, 222]
4	✓	[0.0539, 8.54]
5	✓	[8.55, 20.9]
6	✓	[21.0, 222]
7		[0.0539, 15.7]
8		[15.8, 222]
9	✓	[0.0539, 15.7]
10	✓	[15.8, 222]

In order to obtain a manageable number of routes with different characteristics to be included in the output analysis, we identified and partitioned routes with similar characteristics into different categories, based on speed camera availability and expected number of meetings on the links. From the 10 different link partitions, we considered 56 different route categories, from which we randomly selected one route from each category to be included in our output analysis.

4.2. Average share of successfully identified vehicles

In our simulation study we used 13200 trucks, which were initially scattered uniformly over the nodes in the network. To ensure that the simulation was near equilibrium (near the true traffic flows) we used a 24-hour warm-up period. The simulation was run for a simulated 10-day period, where each successful observation and identification by a measurement device was logged. In our study we let the identification rate for the respective measurement devices be identical over all the links. For instance, if we assume that the identification rate is 10% for stationary devices, the identification rate is 10% on all the links, and if we assume that the identification rate is 20% for mobile devices, the identification rate is 20% on all links. We also include functionality to control the share of vehicles acting as probe vehicles in the identification rate. Suppose the share of vehicles that operates as probe vehicles is x, and the share of encountered vehicles that a probe vehicle is expected to successfully identify be y, then the probability of a successful identification for mobile devices, according to our definition, is $x \cdot y$. We argue that $x \cdot y$ should be proportional to the total number of identifications made by a probe vehicle (i.e., no additional information is gained by knowing both x and y).
We examined the share of the vehicle fleet that was successfully identified at least once on the considered routes for each hour of a day. The results are presented in Fig. 4 and Table 3, where we considered varying identification rate, and assumed identical identification rate for all stationary and mobile devices. The results confirm that it is possible to achieve a high share of successfully identified vehicles, even for a low percentage of the transportation fleet acting as probe vehicles, and for low identification rate for stationary devices.
In the diagrams in Fig. 5 and Fig. 6, we illustrate how the share of successfully identified vehicles is expected to vary over a 24-hour period and the average share of identified vehicles for varying identification rate for a link without speed cameras with expected number of meetings in the range [0.0539, 8.54], respectively. As shown in Fig. 5 it is challenging to successfully identify vehicles during the night when the number of expected meetings is low, and there is no stationary device located along the link.

Table 3. The average share of identified vehicles at least once on all routes with varying identification rate.

Time period	Identification rate									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0-1	0.243	0.384	0.491	0.57	0.634	0.675	0.713	0.737	0.757	0.777
1-2	0.259	0.4	0.5	0.586	0.646	0.696	0.729	0.758	0.781	0.798
2-3	0.288	0.446	0.549	0.625	0.688	0.726	0.755	0.778	0.799	0.813
3-4	0.35	0.524	0.638	0.712	0.76	0.796	0.833	0.865	0.881	0.898
4-5	0.482	0.692	0.803	0.862	0.906	0.931	0.945	0.956	0.966	0.973
5-6	0.72	0.874	0.928	0.951	0.969	0.983	0.99	0.993	0.994	0.996
6-7	0.706	0.885	0.934	0.957	0.975	0.978	0.983	0.988	0.992	0.997
7-8	0.729	0.889	0.936	0.963	0.981	0.987	0.992	0.997	0.997	0.999
8-9	0.721	0.882	0.944	0.965	0.979	0.987	0.99	0.994	0.996	0.999
9-10	0.728	0.901	0.944	0.965	0.979	0.984	0.99	0.994	0.998	0.999
10-11	0.738	0.879	0.935	0.958	0.973	0.983	0.99	0.995	0.996	0.999
11-12	0.744	0.889	0.945	0.968	0.977	0.987	0.991	0.994	0.996	0.999
12-13	0.723	0.864	0.925	0.951	0.969	0.978	0.983	0.989	0.991	0.992
13-14	0.732	0.872	0.931	0.961	0.971	0.98	0.986	0.99	0.994	0.997
14-15	0.731	0.889	0.946	0.965	0.975	0.982	0.988	0.989	0.993	0.998
15-16	0.7	0.873	0.934	0.96	0.976	0.981	0.984	0.987	0.99	0.993
16-17	0.639	0.806	0.883	0.926	0.947	0.965	0.974	0.979	0.985	0.99
17-18	0.577	0.757	0.838	0.886	0.911	0.937	0.949	0.958	0.967	0.972
18-19	0.506	0.692	0.794	0.856	0.899	0.926	0.943	0.955	0.963	0.968
19-20	0.464	0.652	0.757	0.822	0.857	0.885	0.908	0.922	0.937	0.944
20-21	0.406	0.582	0.687	0.766	0.817	0.854	0.875	0.898	0.911	0.928
21-22	0.356	0.541	0.655	0.727	0.785	0.827	0.855	0.88	0.9	0.918
22-23	0.302	0.477	0.597	0.666	0.721	0.758	0.795	0.825	0.844	0.855
23-24	0.259	0.401	0.505	0.583	0.643	0.682	0.709	0.728	0.743	0.755

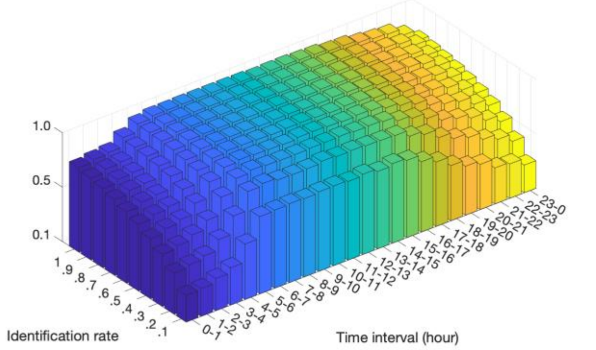

Fig. 4. Share of vehicles that were successfully identified at least once per hour, where the identification rate is varied in the same way for stationary and mobile devices. The share of successfully identified vehicles is the average taken over all simulated routes in the network for each hour and each of the considered identification rates.

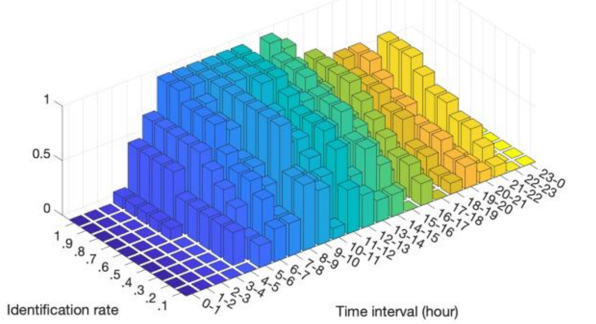

Fig. 5. Share of successfully identified vehicles over a 24-hour period on a link without camera and expected number of meetings in the range [0.0539, 8.54].

Fig. 6 shows the average share of successfully identified vehicles where the identification rate for stationary and mobile device varies separately. We note that even for a high identification rate, it is expected that more than 40% of the trucks will not be identified. This can be explained by that when a truck drives on a link without a speed camera, there must be at the same time at least one probe vehicle that drives on the link in the opposite direction in order to be able to identify the truck. For instance, this situation may not occur during night hours when the number of meetings is low.

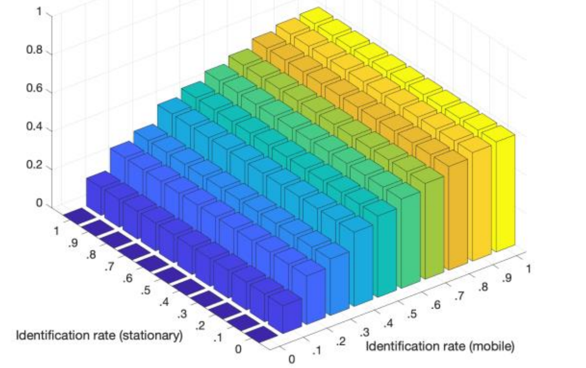

Fig. 6. Average share of successfully identified vehicles over a 24-hour period with varying identification rate on a link without camera and expected number of meetings in the range [0.0539, 8.54].

For comparison, in Fig. 7 and Fig. 8 we illustrate the share of successfully identified vehicles on a link, again with the expected number of meetings in the range [0.0539, 8.54], but the link also has two speed cameras. Fig. 7 and Fig. 8 indicates that for the links with a low number of expected meetings, the share of successfully identified vehicles can be significantly increased by using (especially during nighttime), in this case two stationary devices. On links with a reasonably high number of expected meetings, the share of identified vehicles is expected to be rather high, even for lower shares of vehicles acting as probe vehicles or lower identification rate.

4.3. Successful identification on the first and the last link

In order to further study the movements of trucks in the network, we also considered the share of vehicles that were successfully identified on both the first link and the last link along routes of length three. Fig. 9 shows the share of successfully identified vehicles on the first link and the last link on a route consisting of links without speed cameras and expected number of meetings in the range [0.0539, 15.7]. The highest achieved share of identified vehicles is 0.683, which may argue that it may be insufficient to only use probe vehicles to determine the travel demand on a route with a low number of expected meetings.

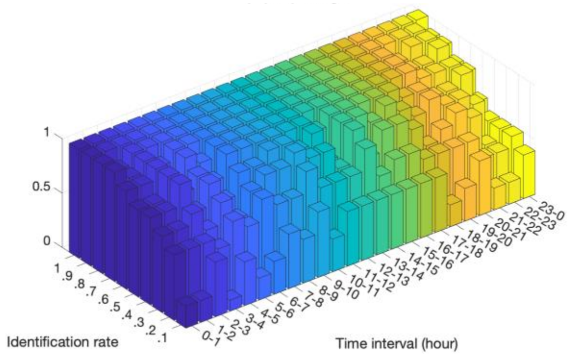

Fig. 7. **Share of successfully identified vehicles over a 24-hour period on a link with two cameras and expected number of meetings in the range [0.0539, 8.54].**

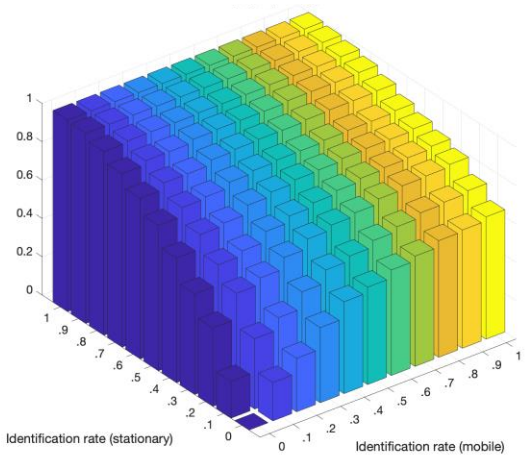

Fig. 8. **Average share of successfully identified vehicles over a 24-hour period with varying identification rate on a link with two cameras and expected number of meetings in the range [0.0539, 8.54].**

In contrast, Fig. 10 shows the share of successfully identified vehicles on the first and the last link with the same range of expected meetings, where the first and last link also are equipped with speed cameras. Even though the expected number of meetings is low, the share of identified vehicles is considerably higher (mean value over all identification rates is 0.859 with standard deviation 0.287). Fig. 11 shows the share of successfully identified vehicles on the first link and the last link, where neither the first link nor the last link has speed camera, and the number of expected meetings is in the range [15.8, 222]. The mean value over all identification rates is 0.742 with standard deviation 0.238. For comparison, in Fig. 12 we show the share of successfully identified vehicles on the first and last link along a route, where the first and last link have a number of expected meetings in the range [15.8, 222], and where speed cameras also are located along the links. From the underlying results of Fig. 12, we conclude that a very low identification rate is required to achieve a high share of successfully identified vehicles when speed cameras are available, and the number of expected meetings is in the range [15.8, 222] (mean value over all identification rates is 0.870 with standard deviation 0.195). For instance, the simulation results show that measurement devices with identification rate of 0.4 or higher, is expected to successfully identify at least 90% of the vehicles on both the first link and on the last link.

5. Conclusions

In this paper we have evaluated the idea of using active stationary and mobile measurement devices for traffic data collection. Based on the results, we suggest using both types of technologies in order to achieve a high share of successfully identified vehicles, due to their complementary characteristics.

In order to validate our idea, we developed a traffic simulation model which utilize flow data with 24-hour resolution. In our simulation study, we applied freight transport data in a region in southern Sweden. The aim of the study was to study the share of successfully identified vehicles with varying identification rate. The results of our study, where speed cameras and freight trucks where used as stationary and mobile devices, shows that a reasonably high share of successfully identified vehicles can be achieved even for low identification rate. Furthermore, the results also indicate the applicability of using stationary and mobile devices to study how vehicles move around in the network. Future work includes studying the integrity aspects of the suggested idea, in particular data encryption and how a public cryptography system can be implemented for secure data transmission.

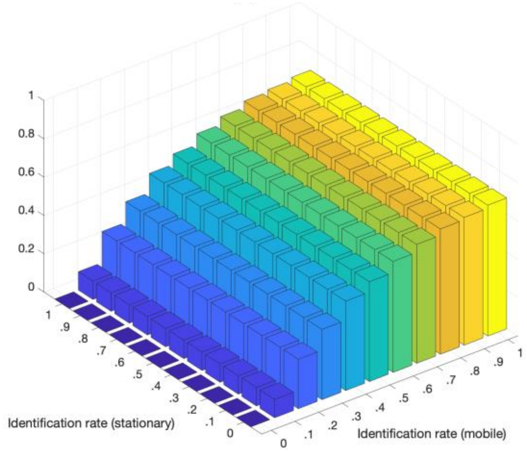

Fig. 9. **Share of vehicles that were successfully identified on both the first and the last link along a route with varying identification rate without cameras on the links and with expected number of meetings in the range [0.0539, 15.7].**

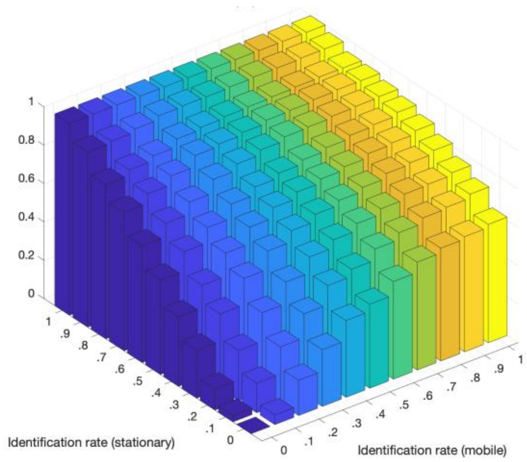

Fig. 10. **Share of vehicles that were successfully identified on both the first and the last link along a route with varying identification rate with cameras located on the links and with expected number meetings in the range [0.0539, 15.7].**

Acknowledgments

This work was partly funded by the Regional Council of Blekinge, the Swedish Agency for Innovation Systems, the Swedish Transport Administration, and the Swedish Transport Agency.

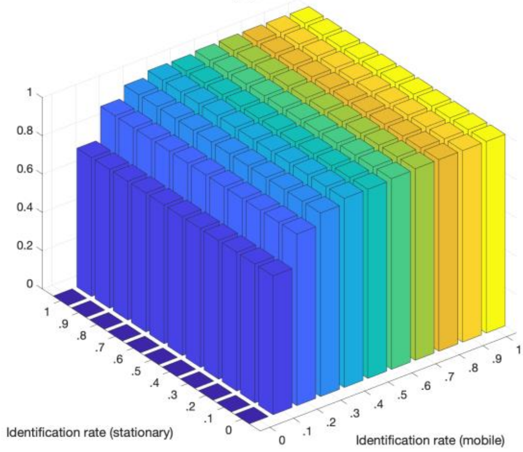

Fig. 11. Share of vehicles that were successfully identified on both the first and the last link along a route with varying identification rate without cameras on the links and with expected number of meetings in the range [15.8, 222].

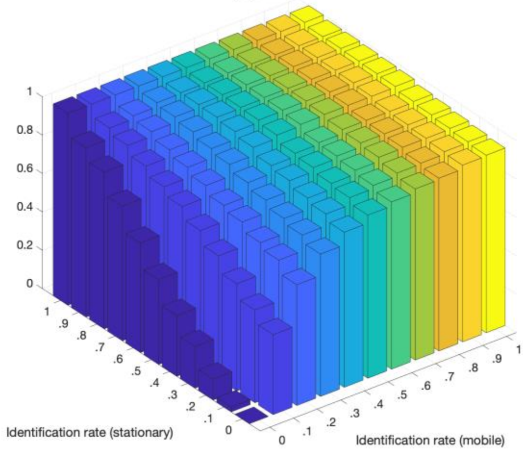

Fig. 12. Share of vehicles that were successfully identified on both the first and the last link along a route with varying identification rate with cameras located on the links and with expected number of meetings in the range [15.8, 222].

References

[1] Josa, I., & Aguado, A. Infrastructures and society: From a literature review to a conceptual framework. Journal of cleaner production 2019;238;117741. https://doi.org/10.1016/j.jclepro.2019.117741

[2] Litman, T. Evaluating transportation economic development impacts. Victoria Transport Policy Institute, 2010

[3] Leduc, G. Road traffic data: Collection methods and applications. Working Papers on Energy, Transport and Climate Change 2008;1(55);1-55.

[4] Noureddine, M., & Ristic, M. Route planning for hazardous materials transportation: Multicriteria decision making approach. Decision Making: Applications in Management and Engineering 2019;2(1);66-85. https://doi.org/10.31181/dmame1901066n

[5] Ochieng, W. Y., Quddus, M. A., North, R. J., & Noland, R. B. Technologies to measure indicators for road user charging. In Proceedings of the Institution of Civil Engineers-Transport 2010;163(2);63-72 https://doi.org/10.1680/tran.2010.163.2.63

[6] Jenelius, E., & Koutsopoulos, H. N. Travel time estimation for urban road networks using low frequency probe vehicle data. Transportation Research Part B: Methodological 2013;53;64-81. https://doi.org/10.1016/j.trb.2013.03.008

[7] Guo, S., Lin, Y., Li, S., Chen, Z., & Wan, H. Deep spatial–temporal 3D convolutional neural networks for traffic data forecasting. IEEE Transactions on Intelligent Transportation Systems 2019;20(10);3913-3926. https://doi.org/10.1109/TITS.2019.2906365

[8] Montella, A. A comparative analysis of hotspot identification methods. Accident Analysis & Prevention 2010;42(2);571-581. https://doi.org/10.1016/j.aap.2009.09.025

[9] Holmgren, J., Fredriksson, H., & Dahl, M. Traffic data collection using active mobile and stationary devices. Procedia Computer Science 2020;177;49-56. https://doi.org/10.1016/j.procs.2020.10.010

[10] Hadavi, S., Rai, H. B., Verlinde, S., Huang, H., Macharis, C., & Guns, T. Analyzing passenger and freight vehicle movements from automatic-Number plate recognition camera data. European Transport Research Review, 2020;12;1-17. https://doi.org/10.1186/s12544-020-00405-x

[11] Grumert, E. F., & Tapani, A. Traffic state estimation using connected vehicles and stationary detectors. Journal of advanced transportation, 2018. https://doi.org/10.1155/2018/4106086

[12] Herrera, J. C., & Bayen, A. M. Traffic flow reconstruction using mobile sensors and loop detector data, University of California Transportation Center, 2007.

[13] Meng, F., Wong, S. C., Wong, W., & Li, Y. C. Estimation of scaling factors for traffic counts based on stationary and mobile sources of data. International journal of intelligent transportation systems research 2017;15(3);180-191. https://doi.org/10.1007/s13177-016-0131-1

[14] Klein, L. A., Mills, M. K., & Gibson, D. R. Traffic detector handbook: Volume I (No. FHWA-HRT-06-108). Turner-Fairbank Highway Research Center, 2006.

[15] Tarnoff, P. J., Bullock, D. M., Young, S. E., Wasson, J., Ganig, N., & Sturdevant, J. R. Continuing evolution of travel time data information collection and processing (No. 09-2030), 2009.

[16] Seo, T., & Kusakabe, T. Probe vehicle-based traffic state estimation method with spacing information and conservation law. Transportation Research Part C: Emerging Technologies 2015;59;391-403. https://doi.org/10.1016/j.trc.2015.05.019

[17] Herrera, J. C., Work, D. B., Herring, R., Ban, X. J., Jacobson, Q., & Bayen, A. M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. Transportation Research Part C: Emerging Technologies 2010;18(4);568-583. https://doi.org/10.1016/j.trc.2009.10.006

[18] Nanthawichit, C., Nakatsuji, T., & Suzuki, H. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway.

Transportation research record 2003;1855(1);49-59. https://doi.org/10.3141/1855-06

[19] Mannering, F. L., & Washburn, S. S. Principles of highway engineering and traffic analysis. John Wiley & Sons, 2013

[20] Dahl, M., Holmgren, J., Fredriksson. H., & Laksman, E. Significant Route Identification using Daily 24-Hour Traffic Flows. Proceedings of 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). Rhodes, 2020 https://doi.org/10.1109/ITSC45102.2020.9294400