Abstract—Since vehicle to infrastructure communication will become state of the art in a few years, data from mobility and environment sensors in the vehicle could produce valuable information for traffic management and vehicle safety applications. The massive deployment of such data sources could however lead to high data traffic that would unnecessarily strain both the communication network, processing and storage resources.

In this paper we present Controlled Probing, a system solution for floating car data collection which allows to select exactly the location, the time period and the type of sensor data to be collected. We address the system performance with respect to the accuracy and timeliness of the received probes.

I. INTRODUCTION

Floating Car Data (FCD) collection is today a mature and commonly used method for collecting traffic related information such as vehicle speed, temperature, etc. directly from vehicles on the road.

The first systems were built by equipping dedicated fleets (taxis, trucks etc.) with dedicated add-on boxes, capable of collecting and uploading FCD via GSM ([1], see field trials in Stockholm [2], Berlin and Vienna [3], China [4]).

Today, a vast range of systems use the driver’s mobile phone and its GPS in order to create detailed mobility traces and upload them to a server, where traffic information is derived and distributed back to the users. The business models vary from advertisement based (Google Maps, Waze), to subscription based services; Coyote and Tom-Tom HD Live Traffic [5]. These approaches are proprietary and each service has to build up their own user-base for it to work.

Finally, cooperative ITS approaches which are under researched will make use of a road operator driven Roadside Unit (RSU) infrastructures, using internationally standardized, IEEE 802.11p based wireless communication. In combination with an On-board Unit (OBU) with direct access to the vehicle’s CAN bus, it can collect data from all sensors from the vehicle and deliver it to an RSU when in range. This approach generally allows for collecting data from numerous vehicles, without the need of user interaction.

Common for the above described systems is that they are pre-configured to collect their position and any other type of sensor data periodically and send this information to a processing central server. The sampling interval is predefined to e.g., 30 seconds or every 50 meters in order to reduce communication costs and the amount of processing and storage resources. However, different applications need different sampling rates. Therefore, most of the current approaches are not flexible, that is, they do not allow for a targeted and controlled collection of FCD, which can be adjusted or even completely switched off, as needed.

In this work we present Controlled Probing (CP), a system for flexible targeting of data collection to specific road areas and FCD which utilizes IEEE 802.11p based communication between roadside infrastructure and OBUs.

The main contributions of this work are the architecture and the interaction protocol of CP. This is supported by an investigation of how OBU penetration rate and sensor sampling rate affect the quality of the estimated value, thereby arguing that high resolution of a particular area is possible by using CP.

Addressing the specific collection of vehicle speed probes, we compare the delay from measurement until the delivery of aggregated values to a central application for both CP and for the FCD collection using a dedicated vehicle fleet and wireless cellular communication.

A. Related Work

The challenge of reducing the amount of collected FCD has already been identified in numerous publications. The authors of [7] propose to transmit the probe only if it exceeds a predefined threshold. [8] improves this approach by adding a randomization function which thereby avoids outlier bias. These approaches reduce the amount of data being communicated, but can only be used to decide whether to send probe vehicle data or not.

In [9] the authors suggest using eXtended FCD (xFCD) which contain events rather than raw sensor data. This approach only allows for collection of atypical behaviour.

A general architecture for probe collection management have been introduced in ISO TS-25114 [6], in which two message formats are defined; the probe data management (PDM) and the probe vehicle data (PVD) messages. However, the specific role of the infrastructure nodes in the command distribution, probe collection and processing are developed for first time in this work.

In [10], [9] and [11] the authors argue that penetration rates between 1% and 5%, are enough to provide useful information on the urban traffic state. These numbers depend on the road type and the sampling interval, on the delivery time and interval, etc. The studies are restricted to velocity information, however, we focus on the penetration and sampling rates which provide the minimum error.

The rest of the paper is organized as follows: In Section II the system and its interactions are described, Section III
describes the aggregation algorithm used for the evaluation, Section IV presents experiments and their evaluation, finally Section V concludes and identifies further research tasks.

II. SYSTEM DESCRIPTION

To describe the operation of CP, we consider the setup in Figure 1 in which vehicles equipped with OBUs traverse the area of interest. Each collection job is defined and managed by the Traffic Control Center (TCC), which can start, deactivate and reactivate jobs. With the job, a set of upstream RSUs is calculated to which the command is disseminated.

The AOI is defined by the start point and the end point, each consisting of a longitude and latitude pair. The RSUs situated upstream of the AOI periodically broadcast the Probing Command to all passing vehicles. This set of RSUs is calculated by searching the road graph. After traversing the AOI, a vehicle can deliver the collected probes at any subsequent, downstream RSU.

The network of RSUs considered in this work has properties that ease the distributed protocol processing.[12] It has the architecture of an overlay network that allows a geolocated message, containing for example the coordinates of a location A in its header, to be routed to the nearest node to that location (location based routing). In addition, a spaces-based middleware software layer allows applications to communicate by exchanging asynchronous messages network-wide via special data containers.[13] This feature is used to collect probes delivered at different nodes in the RSU network.

The central data structure related to the task of collecting data from an AOI is the "job". Jobs are created, activated and deactivated by a central FCD application in the TCC. In the communication with the vehicle, the proposed CP protocol consists of the Probing Command and Probe Vehicle Data message types.

Similarly to [14] a job consists of:
- JobID
- Start collecting location
- End collecting location
- List of normative data elements; temperature, speed, etc.
- Sampling rule; time, distance or a combination of both
- Sampling interval (seconds)
- Sampling distance (meters)

A Probing Command message has the following structure:
- Job data structure
- Validity; max. delay between collection and delivery

The fact that a job is geo-located, determines in the roadside overlay network of RSUs the node responsible for the processing of probes related to this job. Several jobs may be assigned to the same processing node, as they differ by their JobID.

A Probe Vehicle Data message has the following structure:
- JobID
- List of data elements

1 Consists of RSUs on any road leading to the Area Of Interest (AOI).

CP collection job can be summarized as follows:
1) A CP job is created and the set of upstream RSUs nodes is computed, then a Probing Command message related the job is sent to these.
2) Upstream RSUs broadcast the Probing Command messages periodically over the wireless interface.
3) Passing OBUs receive the command message.
4) OBUs perform measurements accordingly.
5) Each collecting vehicle sends the Probe Vehicle Data to any of the downstream RSUs, which forward them to the designated processing RSU.
6) The TCC receives aggregated data from all active jobs.

III. AGGREGATION OF FLOATING CAR DATA

Aggregation is the process of fusion of measurements from different vehicles correlated to a defined spatial-temporal window such that the confidence interval is improved.

In the rest of the paper we will focus on vehicle speed measurements. The main challenges for speed estimation are a) the high variability of speed values in place and time, especially on roads with more than one lane, and b) achieving low data latency required in traffic control applications.

In the following, two methods are presented for the aggregation of FCD measurements.

The first method defines a distance-time aggregation rectangular box consisting of all measurements that take place during the period \([t, t+\Delta]\) and on a certain road sub-segment \([l, l+d]\) within the road segment defined by the job. The \(\Delta\) and \(d\) parameters depend on the sensor sampling time interval, nominal vehicle density of that road, number of lanes, free flow speed, etc. They depend also on the data type measured; for instance speed has other statistical properties than air temperature. We assume that the data is delivered in bulks, grouped by each vehicle and may therefore arrive out of order, depending on the vehicle's travel time.

The algorithm, named aBOX, maintains a list of chronological buffers, one buffer for each interval such as \([t, t+\Delta], [t+\Delta, t+2\Delta],\ldots\), etc. in which probes are sorted. The boxes have also a spatial size, they are for instance 50 meters apart. We define that a buffer is ready for aggregation of its content when a) at least \(N\) messages are available or b) the delay exceeds and pre-define value. After the aggregation the next buffer is considered and so on, keeping the chronological order. Probes received after the timeout are discarded.

The second aggregation method called \(k\)-newest is relatively simple; for each aggregation interval, use the \(k\) messages which have the most recent probe timestamp for a certain road segment \([l, l+d]\), to aggregate the speed. No timers are needed, as we always have \(k\) messages, that however can be quite old, in case of no traffic. In this way, the TCC can be notified periodically with the newest data; however this data is not necessarily accurate and this inaccuracy should be embedded in that data. A further refinement is to consider measurements which are at most \(n\) minutes old.
The $k$-newest messages approach is useful when dealing with periodic updates, as it does not wait for a particular number of probes for a specific time window before sending the update. However, since $k$ is static, it does not utilize all the data which is available and would have a bias towards fast moving vehicles, since they provide "newer" data.

IV. PERFORMANCE EXPERIMENTS

In this section we evaluate the impact of a) the penetration rate of FCD capable vehicles and b) the sampling rate on the aggregated data. The data is evaluated with respect to following metrics; a) the expected delay from the time of measurement until the delivery of aggregated values and b) the accuracy of collected and aggregated data as a function of sampling rate and the FCD-enabled vehicle penetration rate.

A. Evaluation Scenario

The data which provides the basis for the evaluation consists of individual probe data samples, each containing the location as $x$, $y$ coordinates, the speed, a sample creation timestamp and a probe delivery timestamp. The difference between the two timestamps is equivalent to the travel time between the location of the probe and the delivery at the RSU. Each vehicle is configured to sample its speed everywhere and every second. The reason for this is that we are interested in the results, independent of the location of the area of interest.

The traffic scenario for the evaluation consists of a motorway segment, a constant number of vehicles and their routes, but two sets of RSU setups. The first consists of three different RSU densities, used to obtain delay distributions, while the second setup uses the same configuration but some RSUs are disabled, such that we achieve four different distances from the area of interest until the first encountered RSU. The simulation is based on the S1 motorway, located in the vicinity of Vienna, and 3379 vehicles, each having its own, random route. The total length of the motorway segment considered is 16 km, equalling 32 km when counting both directions. Also each direction consists of 2-3 lanes, depending on the location.

We use Simulation of Urban MObility (SUMO)[15], to provide realistic mobility of vehicles. The road network is downloaded from OpenStreetMaps.org and converted with the netconvert tool provided by SUMO. The route of each vehicle is generated with randomRoutes.py, also part of the SUMO toolkit.

NS-3 is fed with the mobility trace and used to determine the communication between the OBUs and RSUs. Each OBU is associated with a simple application which periodically creates probe samples and buffers them until delivery. The RSUs broadcast announcements periodically, thereby notifying all OBUs within their coverage to report their data. Each RSU records all received probes in a trace file for later processing, containing the probe data and when it was received.

For the delay evaluation we assume that the vehicle performs the measurements by sampling the sensors, and delivers the series of measurements (probe) to the next RSU.

B. Road-side Unit Density and Travel Time

The main difference between a dedicated RSU infrastructure and the existing cellular infrastructure, is the scattered coverage of the former, at least in the early deployment phase. Therefore, the probe latency is mainly due to the travel time between RSUs, or more generally from the measurement point until the next RSU, rather than e.g., communication delay. This travel time may increase in congested traffic situations, and become unbounded for worst case scenarios where the road is completely blocked. However, a reasonably dense RSU based infrastructure can compete with current cellular deployments due to the current limitations of the later, as also shown later.

In order to evaluate the delay, we consider the data sets obtained from a low density scenario, where the RSUs are approximately 7-8 km apart, a high density scenario, where RSUs are located before and after each on/off-ramp, an in-between case resulting in a distance of 800-1500 m, and one where RSUs are located in the middle between the off and on ramp, approximately 3 km apart.

These results are used to discuss general tendencies but also used as input for the aBox algorithm. We also compare the results to the delivery delay for a taxi/cellular FCD collection method published in [2]. Newer studies on FCD...
exist, however only this one publishes detailed probe delay data.

Fig. 2 shows the results for the three RSU distributions, compared to the cellular delay distribution. For the later, the sampling rate is in general lower than in the RSU case due to network costs, reduction of data amount and multiple probes are aggregated into one message before sending. The long tail of the low density RSU setup is due to vehicles which are congested, and therefore take longer to arrive at the RSU. In summary, the results show that RSU based FCD reporting can compete with cellular based at medium densities. Also, given a high penetration rate of vehicles, the delay could be further decreased if Vehicle2Vehicle (V2V) communication is used, however this approach requires additional modelling to determine the communication overhead.

C. Aggregation Delay

The total latency, that is the travel time and aggregation delay, of the vehicle velocity probes is a critical parameter in particular for traffic control applications.

In order to evaluate the total latency we have built a scenario in which the traffic on the probed road segment gets congested due to an event on the road. This situation causes an increase of the probe delivery. The results from the previous section provides data for setting the aBox parameters. The $\Delta \cdot d$ box dimension is $\Delta = 60$ seconds and $d = 40$ meters. The CP results to be sent to the FCD application are represented by box plot in Fig. 3. It can be seen a relative good fit to the average of the raw data (ground truth). More insight about the operation of the algorithm is obtained from the delay distribution in Fig. 4; neglecting the first transient periods of the simulation in which the traffic ramps up, the delay values stabilize to a low value corresponding to the free flow in Fig. 3. The number of received probe messages is sufficient (larger than $N=10$ messages), so that some messages are discarded. During traffic congestion, the number of messages, that arrive in time, decreases below $N$ and the timeout $T_1 = 2$ minutes is used to move to the next aggregation box. Large delays are caused by empty boxes (and trigger the timeout $T_2 = 10$ minutes). Thus the delay of the output stream is variable, but can be tuned with the parameters $N$, $d$ and $\Delta$.

D. Accuracy of the Measurement Aggregation

In this section we broaden the previous results and investigate how the penetration rate and the sampling rate affect the accuracy of the estimated speed. We compare the accuracy of two algorithms, aBox and $k$-newest, as a function of the probe input stream variation and then generalize the observations. The core difference between $k$-newest and aBox is that $k$-newest provides continuously updated information, e.g., once every minute, without taking the differences in the probe creation time into account, whereas aBox aggregates the data with respect to creation time. In aBox, the estimates are more accurate, but the rate of sending the aggregated data may vary as the delay in Fig. 4 shows.

The accuracy is quantified by calculating the mean square error of the ground truth mean speed values and the estimated mean speed values for the two aggregation approaches:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$  \hspace{1cm} (1)

where $Y_i$ is the ground truth and $\hat{Y}_i$ is the aggregated estimate. The ground truth calculation is based on all samples available, including vehicles without OBUs, samples that were not delivered and so on. The estimated values (determined by the aggregation algorithm) depend on the penetration rate, sampling rate, whether the OBU passed a RSU and whether the OBU actually generated a probe in the area of interest.

Based on the original simulation, we define an AOI and activate only RSUs in a distance of 0, 1, 3 and 6.5 km, respectively, away, and filter the collected probe data such that it varies the penetration rate of vehicles equipped with OBUs, resulting in less vehicles which are capable of producing and reporting probe data. Similarly, the sampling rate of each vehicle is reduced, which has an impact on the probability that the OBU samples at a high enough resolution, which in our case mean whether it samples within the defined area or outside of it. For the $k$-newest approach we additionally vary the value of $k$ to determine the optimal number of values. The permutations which are investigated are show in Table II.

<table>
<thead>
<tr>
<th>TABLE II</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMULATION CONFIGURATION:</td>
</tr>
<tr>
<td>OBU penetration rates: (%)</td>
</tr>
<tr>
<td>Sampling rates: (Hz)</td>
</tr>
<tr>
<td>k values</td>
</tr>
</tbody>
</table>

Fig. 2. Delay distributions for the three RSU densities compared with the cellular delay pattern.
Fig. 3. Estimated speed from aggregation using the aBox Algorithm. Filled profile curve is the ground truth using all samples.

Fig. 4. Number of measurement samples used for aggregation, and delivery delay for aBox algorithm.

Fig. 5. Mean square error for different values of $k$ of the $k$-newest aggregation algorithm for the full coverage scenario.

Fig. 6. Mean Square Error for different sampling rates as a function of vehicle penetration rate using aBox algorithm.

Each combination is repeated 30 times and at each a random subset of the vehicles is selected corresponding to
the desired FCD penetration ratio and the mean of the MSE is calculated. For the sampling rate we reduce the sampling frequency, varying it from 1 Hz to $\frac{1}{10}$ Hz. Additionally, we remove the initial 10 minutes from the simulation and the last 20 minutes, which consists of populating the road network and stabilizing the flow of vehicles and the dispersion of the vehicles, respectively.

First we evaluate the impact of the value of $k$ in $k$-newest. Fig. 5 shows the comparison between the evaluated $k$ values for the case where the area of interest is within coverage of the RSU and a sampling rate of 1 Hz, making it the best case. The remaining combinations have been omitted for brevity. Instead we focus on the aBox aggregation method. We set this time $k = 20$ as lower values create bias towards high speed moving cars, whereas higher values require a high penetration rate to perform efficiently.

Fig. 6 shows the results for aBox, for the case where the area of interest is within coverage of a RSU, meaning that messages are delivered almost as soon as they are collected. The results for the remaining three RSU distributions have not been depicted as they are practically identical to aBox. Fig. 6 shows that at low penetration rate of 5%, the sampling rate has a large impact; it more than halves the mean square error, due to the reduced number of probes which fall within our area of interest. When looking at the effect of penetration rate, it is surprising that already at a penetration rate of 20% and at a sampling rate of 1 Hz we have nearly minimized the mean square error. This indicates that high resolution data would be possible at even low penetration rates. The amount of probe data generated is not discussed, since with Controlled Probing, the data traffic load is distributed among the various areas of interest and would not pose any scalability problems.

V. CONCLUDING REMARKS

In this work we have introduced CP, a novel method to collect FCD, targeted in time and on a location area, using the dialogue between the vehicle and the roadside infrastructure.

It is important to emphasize the distributed character of command distribution, collection and processing (aggregation) that contribute to a better scalability of the system. The latter allows to significantly increase the sampling rate of sensors in the vehicle, a feature that opens new application fields.

For traffic control application, the collection of vehicle speed values and their latency have been discussed in detail. The potential increase of the aggregated probe latency in CP when compared to the state of the art taxi-fleet measurement seems to be matched up, even at medium densities, but at a lower operational cost and a higher level over detail.

Additionally, it has been shown that when the requirements of the probe results are high, CP can provide accurate results at penetrations rates of as little 20% by allowing to increase the sampling rate without for a specific area, without over-loading the entire system with unnecessary communication and processing of probe data.

Further work will include the alignment with the ISO work and the hybrid probing of events, as well as the realization of traffic control mechanisms based on the CP information.

ACKNOWLEDGEMENT

This work was performed in the FTW project ITS Evolution, a scientific cooperation co-funded by ASFINAG, Kap- sch TrafficCom AG and Vienna University of Technology (Inst. of Telecommunications, Inst. of Computer Languages). The Austrian Competence Center "FTW Forschungszentrum Telekommunikation Wien GmbH" is funded within the program COMET - Competence Centers for Excellent Technologies by BMVIT, BMWFJ, and the City of Vienna. The COMET program is managed by the FFG.

REFERENCES