# In-Network Aggregation for Vehicular *Ad Hoc* Networks

Stefan Dietzel, Jonathan Petit, Frank Kargl, and Björn Scheuermann

***Abstract*—In-network aggregation mechanisms for vehicular *ad hoc* networks (VANETs) aim at improving communication efficiency by summarizing information that is exchanged between vehicles. Summaries are calculated, while data items are generated in and forwarded through the network. Due to its high bandwidth saving potential, aggregation is a vital building block for many of the applications envisioned in VANETs. At the same time, the specific environment of VANETs calls for novel approaches to aggregation, which address their challenging requirements. In this paper, we survey and structure this active research field. We propose a generic model to describe and classify the proposed approaches, and we identify future research challenges.**

***Index Terms*—VANET, vehicular network, car-to-car, *ad hoc* network, in-network aggregation, data dissemination, survey.**

## I. Introduction

THE CORE idea of vehicular ad hoc networks (VANETs) is to install dedicated short range radio communication (DSRC) units into vehicles, which enable wireless communication between vehicles and with roadside equipment. This new type of communication paves the way for many applications related to active safety, traffic efficiency, and infotainment [1]. As a simple example, consider a vehicle that sends warning messages, because there is a traffic jam on the road. Approaching vehicles that receive the messages can brake safely or use alternative routes, and the information transfer is not hindered by fog, curves, or visual obstacles. Once deployed, VANETs have the potential to significantly reduce accidents, carbon emissions, and waiting times in traffic jams.

For many applications, especially in the active safety area, it is sufficient to exchange information in form of so-called *beacons* between vehicles within immediate vicinity. Hence, communication is single-hop and periodic with high message frequency to support low delays required by safety applications. Essentially, beaconing enhances the knowledge horizon of the local vehicle to cover the direct wireless communication range, which is typically assumed to be about 500 meters in free space. Beaconing protocols are expected to be part of first-day VANET deployments; they are currently undergoing standardization in the EU as so-called cooperative awareness messages (CAMs) [2], as well as in the US [3] and Japan [4]. The underlying physical and medium access control (MAC) layers used for VANET communication are standardized within IEEE 802.11p [5].

For applications that require dissemination of information to a larger number of target vehicles, *geocast* provides geographically limited flooding of messages in a specific destination region. Packets are sent as link-layer broadcast and are selectively re-broadcasted by other vehicles to cover the whole region. Communication is usually multi-hop but event-triggered and less frequent than single-hop beacons. Applications that use geocast are, for instance, an approaching emergency vehicle warning, which informs vehicles in the driving direction of the emergency vehicle to make way. In the EU, decentralized environmental notification messages (DENMs) [2] are standardized to disseminate such event messages, as is a geocast protocol for their dissemination [6].

Together, beaconing and geocasts support applications that either require frequent dissemination of information in a small area or infrequent dissemination of events in a larger area. But even efficient geocast protocols cannot support dissemination of frequent updates from many vehicles in large regions, which is required by applications like traffic information systems or parking spot finders. To enable such applications, information dissemination methods need to consider application semantics to keep bandwidth consumption low while maintaining information quality. Where tolerable, information from multiple sources needs to be combined and aggregated during routing instead of being forwarded unmodified and only being evaluated by receiving vehicles. This is the goal of in-network aggregation protocols for VANETs [7].

### A. Network Capacity Issues

To understand why aggregation is necessary, consider the following example, illustrated in Fig. 1. Vehicles on a stretch of highway disseminate their current speed upstream to improve traffic routing. To allow for consideration of alternative routes, speed information for approximately the upcoming 5 kilometers of road needs to be available to all vehicles. Because we assume no aggregation, that means each vehicle needs to receive the exact speed of each vehicle in its 5-kilometer

Manuscript received June 28, 2013; revised January 19, 2014; accepted April 9, 2014. Date of publication April 24, 2014; date of current version November 18, 2014. This work was supported in part by the European Union's Seventh Framework Programme through the PRESERVE project under Grant 269994.

S. Dietzel is with the Institute of Distributed Systems, University of Ulm, 89081 Ulm, Germany (e-mail: stefan.dietzel@uni-ulm.de).

J. Petit is with the Centre for Telematics and Information Technology, University of Twente, 7522 NB Enschede, The Netherlands (e-mail: j.petit@utwente.nl).

F. Kargl is with the Institute of Distributed Systems, University of Ulm, 89081 Ulm, Germany, and also with the Centre for Telematics and Information Technology, University of Twente, 7522 NB Enschede, The Netherlands (e-mail: frank.kargl@uni-ulm.de).

B. Scheuermann is with the Institut für Informatik, Lehrstuhl für Technische Informatik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany (e-mail: scheuermann@informatik.hu-berlin.de).

Digital Object Identifier 10.1109/COMST.2014.2320091

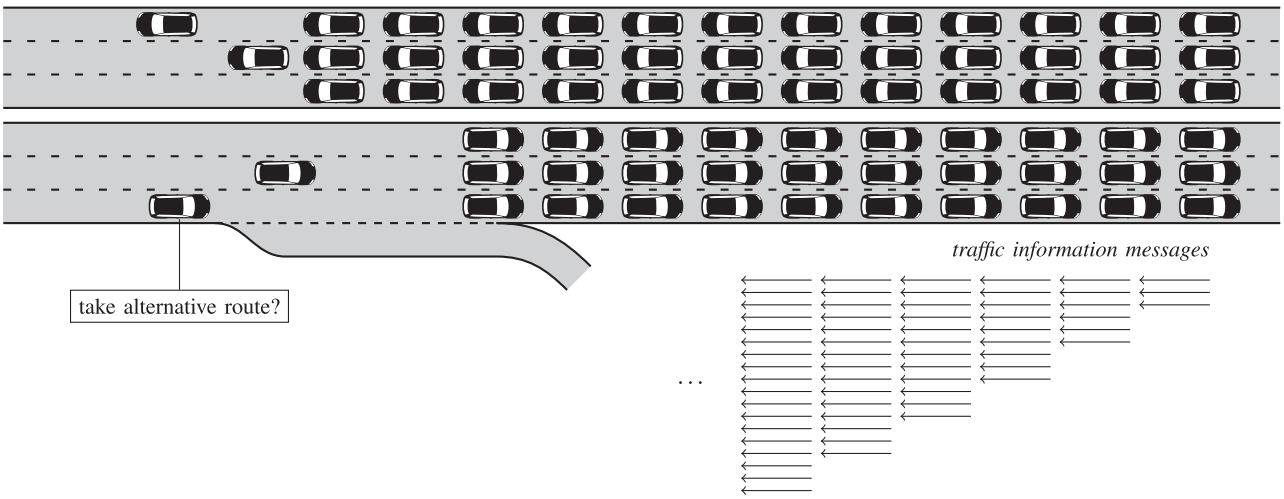



Fig. 1. Traffic information dissemination without aggregation.

vicinity. There is a traffic jam on all 6 lanes of the highway, leading to high vehicle density.

Further assuming an average car length of 5 meters, 1 meter space between vehicles in the traffic jam, and a conservative communication radius of 250 meters, each car will have about 500 vehicles in its direct communication range. As basis for traffic jam detection, we assume that vehicles will exchange beacons such as CAMs.[1] At the minimum rate of 1 Hz, each vehicle will receive up to 500 CAMs per second within its direct communication range. One communication hop farther, vehicles have to forward 1000 messages; 500 received from their direct communication range plus 500 messages received from their 2-hop neighbors. At 5 kilometers distance, already 5,000 messages need to be forwarded per second if basic duplicate detection is in place and each unique message is only forwarded once per vehicle.

Even under idealized conditions, it is unlikely that this amount of messages can be disseminated over large distances. If we assume that only the traffic-situation-relevant subset of CAMs is disseminated farther (that is, only position, vehicle ID, speed, and road ID, all encoded as they are in CAMs), the size of one message is 25 bytes. Further assuming that multiple messages are combined in one packet to save packet header overhead and ignoring wireless transmission collisions, at most 60 such messages can be forwarded per second in high vehicle densities [8] using the IEEE 802.11p MAC layer [5].

That means even within direct communication range, at most 12 percent of the available information can be forwarded. Continuing the calculation, only 1 percent of the available information can be transported to vehicles in 5 kilometers distance. If other applications besides the traffic information system consume bandwidth, the forwarding rate will be even lower. Moreover, it is likely that wireless collisions will occur, further lowering the forwarding rate. Also, our example assumes one-dimensional dissemination due to the highway scenario. In Section II-B we discuss that in cities, where information is disseminated in a two-dimensional plane, the rate decreases quadratically.

[1]The EU foresees DENMs to warn about traffic jams once they are detected, but the detection mechanism for traffic jams is not specified.

### B. Potential of In-Network Aggregation

Because the percentage of information that can be forwarded multi-hop is low, dissemination mechanisms need to implement selection criteria to decide what subset of available information to forward and what to discard. Bad selection can result in skewed information, which does not reflect the real world situation. It is, therefore, crucial for data utility that forwarding criteria are well selected. Even better though, the classical approach of routing information unmodified through the network should be reconsidered. Continuing the above traffic jam example, the core idea is as follows. Instead of forwarding notifications from every single vehicle in the traffic jam, each receiving vehicle assesses whether it is part of the same traffic jam and, if so, only forwards a merged message. Even distant vehicles that receive a notification about two segments of the same traffic jam can merge the messages and only forward the aggregated result. Such merging of different information items can provide bandwidth savings that are superior to schemes that do not modify information in the forwarding phase. For the purpose of this paper, we define the VANET in-network aggregation pattern as follows.

*Definition 1: In-network aggregation* in VANETs is any kind of multi-hop message dissemination where a number of vehicles collaborate to *gain knowledge about real-world phenomena.*

To do so, they exchange messages containing relevant information *derived from atomic sensor readings* or other means of information collection.

During the dissemination of information, atomic *information items are modified and processed* by intermediate vehicles.

This kind of in-network aggregation is especially suitable for all applications that need to build and maintain a knowledge base about large areas and can tolerate approximate information as well as delays in information dissemination. Predominant

use cases are traffic information systems, free parking spot finding, weather information systems, and road condition information systems.

In-network aggregation has also been intensively discussed and successfully applied in the context of (WSNs) and in participatory sensing applications, but their requirements are different: in sensor networks, aggregation mechanisms are means to reduce energy consumption rather than to reduce network capacity usage. Moreover, sensor networks typically transmit data to one or few sinks in a static topology, whereas in VANETs, the topology is highly dynamic and many vehicles are interested in the aggregated information [9]. Section IV-F further elaborates on the relationship of VANET aggregation schemes with proposals for aggregation mechanisms in other domains.

### C. Issues With Infrastructure Support

As an alternative solution to vehicle to vehicle (V2V) network capacity issues, both roadside unit (RSU) communication and use of cellular networks, such as the universal mobile telecommunications system (UMTS) or long term evolution (LTE) have been proposed, as well. For RSU communication, roadside infrastructure has to be deployed in regular intervals along the road network. Then, vehicles can report their traffic reports to RSUs, which are connected to centralized servers using wired network infrastructure. The server aggregates reports and disseminates current traffic information. While the deployment of RSUs is considered in densely-inhabited urban areas, deployment and maintenance costs of 3,000–5,000 US dollars per RSU [10] are widely considered prohibitive for highway coverage. If RSU deployment widens in the future and costs go down, aggregation prior to communicating with RSUs can still help to decrease bandwidth requirements for the connection of RSUs and backend systems.

In addition, the use of cellular networks instead of multi-hop *ad hoc* communication has gained momentum in recent years, and UMTS coverage is becoming a commodity in urban areas. When available, cellular networks can help to disseminate information to vehicles that are hundreds of kilometers away, facilitating early route adaptations. Still, we see two major drawbacks of cellular communication. First, cellular broadband coverage is, again, unlikely to be widely available along highways. Even if highways are covered, UMTS suffers from the breathing cell problem [11], which essentially means that coverage is reduced in scenarios with a high number of users, such as traffic jams. Second, usage of cellular networks, unlike vehicular *ad hoc* communication, incurs costs, which have to be paid by either the vehicle manufacturer, vehicle owner, or driver. Finally, device-to-device (D2D) communication is discussed as part of upcoming LTE cellular standards (cf. [12]). Using D2D, user devices can communicate directly with each other, bypassing the need for LTE base stations. However, such D2D communication is likely to face similar bandwidth issues as IEEE 802.11p-based networks, and therefore benefits from information aggregation.

Therefore, while roadside infrastructure and cellular networks can complement in-network aggregation if they are available, we believe that it is unlikely that they will completely replace the need for aggregation mechanisms.

### D. Our Contributions

We believe that in-network aggregation is an important building block to enable multi-hop information dissemination in vehicular *ad hoc* networks. Strong aspects are cost- and bandwidth-efficient dissemination of up-to-date information in large regions. In this paper, we offer a comprehensive overview of existing protocol proposals, including a discussion of the requirements that they need to meet, and models to categorize and assess them. We highlight promising solutions and point out open issues.

The remainder of the paper is organized as follows. First, we structure different application scenarios for aggregation and analyze their requirements in Section II. Next, we present generic models for VANET in-network aggregation schemes that we will then use to compare the different approaches (Section III). Section IV reviews different aggregation schemes proposed in the recent years. Section V provides a discussion and comparison of the different concepts and presents lessons learned. We differentiate aggregation in other domains in Section IV-F. Section VI summarizes and concludes this paper.

## II. Applications and Requirements

Protocols with aggregation functionality are applicable to a range of use cases for VANETs, and a number of requirements deserve attention. We will first discuss common applications, from which we derive and discuss a broad range of requirements.

### A. Applications

Applications for vehicular networks can be broadly categorized into safety application, traffic efficiency applications, and infotainment applications [13]. Active safety applications are a major use case for VANETs and are likely to be part of first deployments (e.g., [14]). However, safety applications typically require exact data to be transmitted with little to no latency. These requirements directly contradict the aims of in-network aggregation, which we introduced in Section I-B. Namely, in-network aggregation aims to merge data and transmit it with reduced granularity and/or periodicity. However, in-network aggregation is very suitable for a wide range of traffic efficiency applications. In contrast to safety applications, traffic efficiency applications often require periodic multi-hop dissemination of large amounts of information in wide areas, thereby consuming more wireless bandwidth if implemented naïvely. In addition, safety messages can be used as information source by in-network aggregation protocols. Safety messages often contain traffic-efficiency-relevant information, such as vehicle velocity or outside temperature.

Consequently, research on in-network aggregation has focused on traffic efficiency applications. In particular, the following application classes are typically addressed in existing literature.

TABLE I
AGGREGATION USE CASES

| Application | Events | Values | Mentioned in literature |
|---|---|---|---|
| Traffic information systems | traffic jam | average speeds, minimum speeds, travel times | very often |
| Weather information systems | — | average temperature, average visibility | sometimes |
| Road condition warnings | icy road, road construction | — | sometimes |
| Parking spaces | — | # of free spaces | often |

- **Traffic information systems**. Vehicles collaborate to disseminate an approximation of the current traffic situation. Either, only the presence of an event (e.g., traffic jam) is communicated, or actual (average) speed values.
- **Weather information systems**. Just like traffic information, data about average temperature or severe weather conditions can be aggregated. Again, the presence of an event or averages can be used.
- **Road condition warnings**. All forms of road conditions, e.g., stretches of icy roads or bumps, can be summarized using in-network aggregation mechanisms.
- **Parking spaces**. Especially in city scenarios, aggregation can be used to collect the number of available parking spots in different areas.

Of these scenarios, traffic information systems are by far the most-cited use case, followed by aggregate parking space information. Table I summarizes the applications. Regarding the information they collect, we distinguish between *events* and *values*. Events are either present or not; hence, they require only one bit of information to be disseminated. Values represent averages or counts. They require a proper data fusion function (e.g., a duplicate-free average) and need to be represented with higher precision (e.g., as 32-bit double value).

Common to all applications is that they collect information about large scale phenomena and do not impose strict real-time constraints on communication. For instance, traffic information systems often try to identify and characterize traffic jams, which span long stretches of road and consist of a large number of vehicles. As opposed to active safety applications, dissemination delays in the order of minutes are tolerable for traffic information systems.

Further, all applications require dissemination of collected information in large areas, mandating multi-hop dissemination patterns. For instance, traffic information systems require information about traffic jams that are several kilometers away in order to efficiently calculate alternative routes. Likewise, available parking space information needs to be available to vehicles further away to foster proper navigation decisions.

All foreseen applications can tolerate lossy approximation of data. In fact, even if raw data is available, application algorithms will most likely summarize it before deriving decisions. For instance, if traffic information systems work on raw data, that is, speed information about each single vehicle on a road, they will first analyze it to detect possible traffic jams and only base their navigation decisions on the derived traffic jams. Hence, if the exact requirements of applications are known, aggregation mechanisms can be tailored to a point where they can save a great amount of communication bandwidth by employing semantic data compression, without sacrificing data utility for applications.

In the following, we provide an overview of what we consider the key aspects of and requirements for VANET data aggregation, point out the challenges resulting from them, and discuss results and solution approaches.

### B. Data Reduction

In VANET protocols, data reduction is often done in a distance-based manner. With increasing distance from the source of a measurement, the provided information becomes increasingly coarse. Hence, it can be described and transmitted with a lower number of bits per second on the medium. The exact meaning of "coarse" may vary significantly in that context: for instance, updated measurements can be provided less often, thereby essentially reducing the temporal resolution of what is provided to network participants (e.g., [15], [16]). It is also conceivable to reduce the spatial resolution by summarizing measurements from larger and larger geographical areas into single aggregates with increasing distance (e.g., [17], [18]). Or, data representations with a lower accuracy, and thus a smaller size, can be used for measurement data from larger distances (e.g., [19]–[21]). All these approaches—alone or in combination—in essence reduce the network bandwidth that is spent to convey information about a certain part of the real world.

This observation leads to a very generic perspective on VANET data aggregation: an aggregation scheme can be characterized by how much local network bandwidth it spends on spreading information that stems from a source at a given distance. This characterization is independent from how aggregation is performed in detail and independent from the specific application and protocol used to generate, transport, and make use of the information. A "source" in this context is an atomic item about which information can be obtained and distributed by the vehicles—for instance, a single segment of a road. Scheuermann *et al.* [22] introduce a *bandwidth profile* to capture the network bandwidth usage. A non-negative function $b : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is a bandwidth profile of a given aggregation mechanism $\mathcal{A}$ if $\mathcal{A}$ ensures that information about a source at distance $d$ is provided at least with bandwidth $b(d)$. Under very generic assumptions, which hold regardless of the specific application and protocol, it is shown in [22] that aggregation schemes can, in general, only deal with the finite network resources if their bandwidth profiles are in $o(1/d^2)$. That is, any aggregation mechanism in VANETs must reduce the amount of network bandwidth spent on data from a source at distance $d$ asymptotically faster than $1/d^2$. It is also shown that this bound is tight, i.e., it is possible to build aggregation mechanisms with bandwidth profiles that come arbitrarily close to that bound and can still under all circumstances deal with finite locally available bandwidth.

We can distinguish two approaches to achieve the necessary data reduction: *syntactic* compression and *semantic* compression. For the purpose of our survey, we define syntactic compression as a transformation of input values $a_1, \ldots, a_n \rightsquigarrow b_1, \ldots, b_m$ such that given $b_1, \ldots, b_m$, $a_1, \ldots, a_n$ can be fully reconstructed. All mechanisms that allow only partial or no reconstruction of $a_1, \ldots, a_n$ employ semantic compression. We use the term "semantic" to denote that such mechanisms usually employ knowledge about the information contained in $a_1, \ldots, a_n$ to calculate a meaningful representation.

Consider difference encoding as an example for syntactic compression: $b_1 := a_1$ serves as baseline, whereas $b_2 := a_1 - a_2$, $b_3 := a_1 - a_3, \ldots, b_n := a_1 - a_n$ encode only the differences of subsequent values. This scheme allows for compression, because $b_2, \ldots, b_n$ can be represented efficiently if the variation of $a_1, \ldots, a_n$ is low. In addition, all input values can be fully reconstructed. The average function, i.e., $b = (a_1 + \ldots + a_n)/n$, is a semantic compression function. For large $n$, much bandwidth is saved if only the average is communicated, but the input values cannot be reconstructed using only their average. Moreover, average calculation is only meaningful if the input values' semantics are known. If input values represent velocities, their average is a useful abstraction; if input values are geo-coordinates, it might not be. *Cluster-based accurate syntactic compression of aggregated data in VANETs* (CASCADE) [23]–[26] is an example for a syntactic compression scheme. Instead of absolute position and speed values for individual vehicles, the scheme transmits averages of position and speed over a group of vehicles. In addition, the individual cars' values are encoded as deltas of the average. Clearly, these differences will typically come from a much smaller range than the absolute values. Thus, they can be encoded with a smaller number of bits. The combined record will therefore be smaller than the sum of the sizes of individual records. Moreover, no information is lost in the compression process. The size of the records, however, will still increase linearly with the number of included vehicles. Thus, the gain is only a constant factor, and the bandwidth profile is not low enough to scale to arbitrarily large areas.

Other schemes employ semantic compression. Examples are TrafficView [27] and the fuzzy-logic-based scheme by Dietzel *et al.* [16]. The idea is comparable to other media compression algorithms, like JPEG [28] or MPEG Audio Layer III (MP3) [29]: by exploiting knowledge about which information is the most relevant for an application, details can be ignored without loss of information utility. For instance, a traffic information system likely does not need information about the exact velocities of all vehicles on a single congested stretch of road. Instead, the average velocity, or even just the (binary) information that the speed is close to zero, will suffice. The cost of semantic compression is that it is not fully reversible. Therefore, compression functions have to be chosen carefully with application requirements in mind.

Closely related to the means of data compression is the way in which (compressed) information is represented for dissemination. Once aggregated, information needs to be encoded in packets, which are then sent over the wireless channel to facilitate dissemination to other vehicles. Basically, aggregation schemes can aim to reduce the size of these packets, reduce the number of packets, or a combination of both. All approaches have their distinct advantages. If packet size is reduced, the channel load is reduced and collisions are less likely to occur. If fewer packets are transmitted, channel contention is reduced.

Reduction of packet size can be achieved using syntactic compression of information, as discussed above. For instance, disseminating a normal beacon containing a vehicle's speed once and only broadcasting speed differences afterwards reduced the size of subsequent packets. Similarly, reduction of the amount of packets can be realized by simply filtering redundant information. For instance, a vehicle could only disseminate its current speed after a substantial change. Most aggregation mechanisms, however, employ a combination of packet size and packet count reduction; examples are [15]–[17], [19]–[21], [30], [31].

A basic pattern that we already discussed in Section I-B is that vehicles take information received from several other vehicles, merge it, and only further disseminate the result. As a result, the number of packets is reduced, because only one instead of many packets is forwarded. Moreover, the merged information may be coarser than the original information, depending on application requirements. Therefore, the size of packets is reduced, as well. To achieve a trade-off between packet count and packet size reduction, many schemes combine several merged information items into one packet before transmitting it.

### C. Overhead Reduction

As soon as an aggregation mechanism summarizes information from several vehicles, it needs a way to describe the area and time that the summarized information is about. These identifiers lead to another problem: a reduced amount of data used to describe information will not be helpful if a much higher amount of (meta-)data is required to describe the area to which the aggregated value refers. Practical aggregation mechanisms therefore depend on efficient means to encode the scope of an aggregate in both time and space.

For one-dimensional roads, such as highways, the encoding problem is manageable: two points suffice to describe an interval on the road. To distinguish different roads, a road ID can be added. Similarly, two points suffice to describe an (axis-aligned) rectangle in a city environment. To further reduce the overhead, some schemes impose a fixed subdivision on the road network. Instead of specifying intervals and rectangles explicitly, unique identifiers suffice to describe aggregate regions, but the underlying subdivision structure needs to be known to all network participants.

To further reduce the required overhead, many schemes impose a fixed hierarchy of aggregate areas, such as shown in Fig. 2. Caliskan *et al.* [17] were the first to come up with such a scheme: in order to disseminate information on the current parking situation in a city, they subdivide the city area by a hierarchical quadtree structure.

While rectangles or hierarchies of rectangles are well-suited for parking information, they are not ideal to capture the traffic situation in a city. Traffic jams do not expand in rectangles on

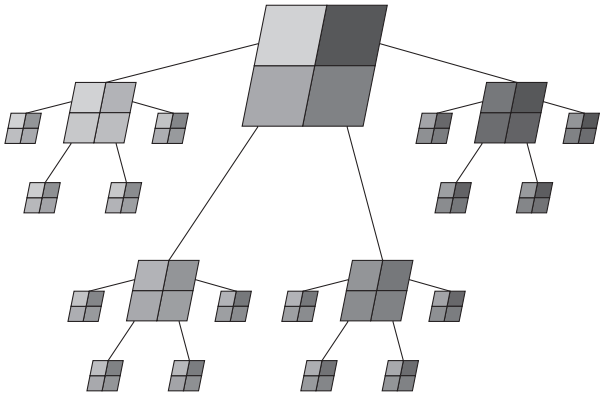


Fig. 2. A quadtree is used to implement aggregation hierarchy.

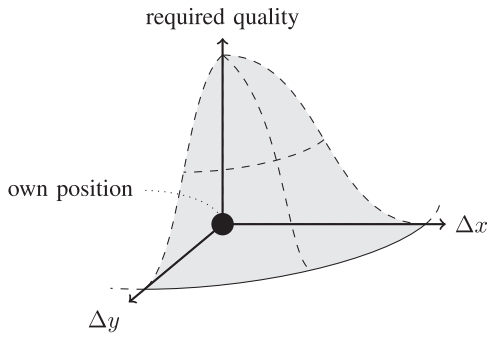



Fig. 3. Exemplary data quality requirements of a parking spot finding application.

a map, but along streets. As a result, encoding the exact areas of traffic jams is more complex. Suppose there is a traffic jam around a highly frequented inner city street crossing. Such a traffic jam can expand along all streets leading up to the intersection, without being present on (geographically very close) roads or lanes leading away from the intersection. Encoding an aggregated information item about the whole traffic jam is not trivial in these situations, because road and lane identifiers, as well as information about the traffic jam length need to be included for all lanes that are part of the traffic jam. Such data encoding for the aggregation area likely diminishes bandwidth savings.

Therefore, aggregation schemes need to find a balance between data reduction and overhead reduction. For instance, to achieve ideal data reduction, a large traffic jam spanning multiple streets in the city center should be encoded in one summarized message. But to achieve minimal overhead, it might still be advisable to use multiple aggregated summaries. Fixed grids, road network maps, and hierarchies can help to reduce overhead when disseminating aggregates between vehicles. But their maintenance and synchronization should be kept in mind when calculating a scheme's overhead.

### D. Preservation of Data Utility

While reduction of data and overhead are necessary to cope with bandwidth requirements, it is equally important to ensure that the data utility after aggregation, sometimes referred to as quality of information (QoI) [32], [33], still meets application requirements. Note that metrics to judge data utility cannot be generalized but instead depend on the requirements of a particular application. Generally speaking, an application will disseminate information from $n$ different dimensions, like time, location, number of parking spots, average speed, and so forth. And for each dimension, a different data granularity is required. Moreover, the granularity requirements differ depending on context, such as distance to the local vehicle, as depicted in Fig. 3. Given such a function, it is the goal of an aggregation mechanism to fulfill the granularity requirements.

In addition to using the right level of aggregation, the functions used for merging information should also be chosen carefully to support data utility. The main requirements can be summarized as order and duplicate insensitivity [34]. For instance, consider an aggregation mechanism that averages speeds in a region. Here, the speed of each vehicle should be counted once per time period, but not more. Otherwise, the average value will be biased towards the reading of some vehicles. Likewise, the order in which items are aggregated should not change the result. Order insensitivity is especially important for distributed, hierarchical aggregation schemes.

Tightly related to duplicate insensitivity, schemes should be able to cope with data that changes over time. For instance, if a scheme only allows one observation per vehicle, it should be able to always select the newest observation from each vehicle to be counted. Ideally, schemes should provide update mechanisms for aggregates to allow fresh information to be integrated without recreating the aggregates from scratch—because otherwise it is impossible to update the aggregate with new sensor readings in a node where only the summarized aggregate is available.

Finally, the semantics of data absence should be well defined by a scheme. Consider an aggregation scheme that only aggregates binary information about traffic jams. Naïvely implemented, the absence of data can either mean "there is no information available" or "there is no traffic jam at that position." Due to the delay-tolerant and lossy nature of communication in a sparsely connected, dynamic, wireless network of moving vehicles, aggregation scheme designers should therefore consider to also disseminate negative information.

The issues discussed so far deal with quality loss introduced by the aggregation mechanism itself. But, like other schemes that deal with collaborative sensor data collection, aggregation schemes also have to cope with sensor faults or biases. The extent to which aggregation results are affected by such faults largely depends on the aggregation function used. An arithmetic average can be influenced more easily by outliers than a median or other functions that specifically detect and filter outliers. Independent of specific fusion functions, multiple sources that contribute to aggregated values can be used for outlier detection and actually benefit the utility of the results. Such use of aggregation specifically to filter faulty values and correct biases, however, is not actively pursued by existing work.

### E. Flexibility

From the requirements formulated above, it is obvious that an aggregation mechanism needs to be able to adapt to different

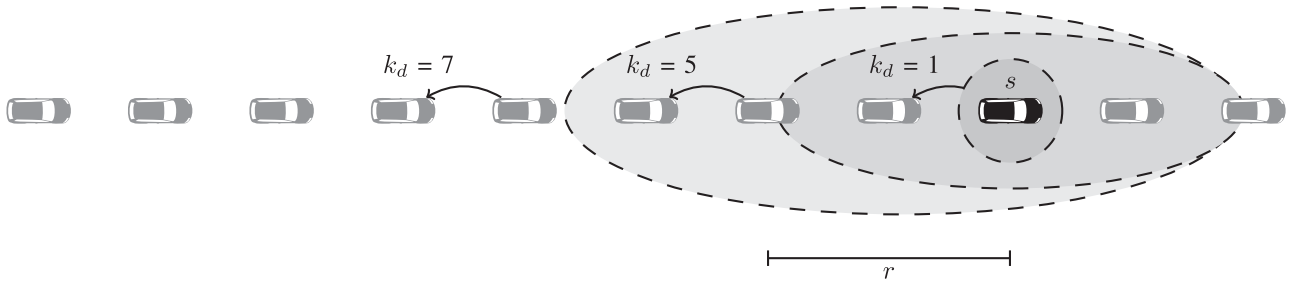



Fig. 4. Exemplary $k$-anonymity over distance calculation: in the 1-hop neighborhood of vehicle $s$, marked by the communication range $r$, $k_d$ equals 1, because each vehicle receives $s'$' atomic observations; vehicles further away receive only aggregated information, which increases $k_d$.

situations. Section II-B mandates that the amount of communicated data per time unit has to be reduced at least quadratically with increasing distance to an observation. Similarly, we argued in Section II-D that different applications require a level of aggregation that matches their specific requirements. In essence, fulfilling both requirements often means that very coarse-grained aggregation is mandatory for information far away, while information close to the own vehicle needs to be represented in a much more fine-grained way. Thus, schemes that use simple road segmentation approaches are in general not flexible enough.

An obvious extension is to define a certain segmentation of the road network on the lowest level and then impose a hierarchical data structure on top of that, as discussed in Section II-C. While allowing for flexible adaptation of granularity relative to the distance of events, such fixed hierarchies still have shortcomings. First and foremost, the granularity does not adapt to the information contents. For instance, suppose there is a nearby stretch of road with very homogenous traffic. Here, a compact representation would be possible without significant loss in accuracy. With a fixed granularity, though, this cannot be realized. The opposite situation would, for example, be two lanes where one is congested while the other one flows freely. If this information is over-aggregated, an important basis for routing decisions is lost.

Therefore, aggregation mechanisms benefit from being able to adapt their granularity to location, time, and information contents to make best possible use of data compressibility while considering the information loss tolerance of the specific application. This trade-off can be modeled as a problem of decision theory (cf. [35, Ch. 36]): given an imperfect knowledge about the surroundings, which (aggregation) decision maximizes utility of information? TrafficView [19]–[21] is a scheme where the utility—that is, the minimization of information loss—is explicitly modeled and used for aggregation decisions. Also, Kumar and Dave [36], [37] use multi-criteria based decision making for aggregation decisions. Other works employ machine learning concepts such as clustering [38] or fuzzy-logic-based clustering and decision making (cf. [39, Ch. 13, 15]); examples for approaches using fuzzy logic are [15], [16], [40].

While scalability criteria provide a hard upper limit for the bandwidth, application requirements on data quality should be kept whenever possible. The key challenge for such an application-tailored, lossy compression scheme is, of course, that the data is constantly changing and, in general, not all the finer-grained information is available in one single vehicle.

### *F. Privacy*

Privacy has been highlighted as an important property for VANET deployment [41]. Aggregation has intrinsic privacy benefits, because information is summarized more and more with increasing distance to the participating vehicles. Thus, the further away an observer is from a target vehicle, the less information she gets about the exact position, speed, and other information items from the observed vehicle. Dietzel [42] proposes an adapted $k$-anonymity metric to quantify the intuitive privacy notion of aggregation. The $k$-anonymity metric, introduced by Sweeney [43], is mainly used in the database domain to express that table entries are indistinguishable from $k-1$ other entries. Dietzel [42] extends the $k$-anonymity concept to $k$-anonymity over distance, meaning the indistinguishability of information about a target vehicle from other information items in the same aggregated information item, depending on the distance of an observer (see Fig. 4). The paper focuses on highway scenarios. Hence, the distance to the observer is expressed as a one-dimensional value. In the best case, $k$-anonymity over distance increases linearly:

$$k_d^{\max} = \max\left(1, \lfloor d/r \rfloor \cdot \mathcal{N}\right) \tag{1}$$

where $d$ is the distance between the observer and the observed vehicle in meters, $r$ is the 1-hop communication radius, and $\mathcal{N}$ is the average number of 1-hop neighbors of a vehicle.

Note that this best case calculation depends on a number of assumptions. Most importantly, a mechanism achieving the best case must use a high level of aggregation, including hierarchical aggregation. In fact, it needs to aggregate all information available, independent of the actual contents of the information. Considering that a mechanism actually needs to adapt flexibly to different situations (see Section II-E), the best case will only be achieved for homogenous traffic situations in practice.

Moreover, the best case calculation assumes a perfectly non-invertible aggregation function. That is, even if an attacker knows some exact values that were used when calculating aggregated values, those values should not make it easier for her to calculate the exact values of a target vehicle. Unfortunately, such an assumption will not hold for commonly used aggregation functions. For instance, suppose that the average is calculated as the aggregation function. Then, the more single values are known, the more constrained the distribution of the remaining unknown values becomes, i.e., the more is known also about them. In the extreme case, if all single values but one are known, the unknown value can even be calculated

exactly. Similarly, the possible range for minima and maxima is diminished if some single values used in the calculation are known.

Finally, all direct neighbors of a vehicle will always have access to exact information, because it is used to bootstrap the aggregation process. Therefore, it is important to design aggregation schemes in such a way that these direct neighbors cannot easily abuse this exact information.

### G. Information Integrity Protection

In contrast to the intrinsically higher level of privacy, the resilience of aggregation mechanisms against malicious data manipulation is generally lower than the resilience of comparable schemes using exact information. Applying Raya and Hubaux' [44] categorization, our attacker model is a *rational, active insider attacker*. That is, the attacker's goal is to create messages suggesting a specific traffic or other situation, which diverts from the real world in a way beneficial to the attacker. The attacker is assumed to possess valid key material issued by a public key infrastructure (PKI) (e.g., [45] or [3]) to create signatures on her messages. To achieve maximum impact, it is conceivable that an attacker will falsify messages that claim to contain aggregated information from other vehicles about large regions.

As a result, driving efficiency systems may not work correctly. For example, routing systems would suggest suboptimal routes due to fake traffic jams. False route proposals can lead to lower user acceptance of such systems. Even though aggregation mechanisms are usually not used to support active safety mechanisms, it is also conceivable that manipulated data might lead to dangerous situations on the road. For instance, if vehicles miss information about an upcoming traffic jam, they might drive too fast and not be able to brake in time.

Due to the nature of aggregation, typical entity centric security mechanisms for VANETs [46] are not directly applicable. Normally, vehicles that receive a message use the attached signature and certificate to ascertain that the sender is a vehicle. In case of insider attackers, this check at least limits the possibility of so-called Sybil attacks. In a Sybil attack, a single attacker pretends to represent multiple vehicles, as explained in [47]. Moreover, vehicles often collect multiple reports about the same event and use majority votes [48]–[50] to decide whether to trust the reports. For aggregation, this approach does not work: for the same reasons that the granularity of the information itself needs to be reduced, which we discussed in Section II-B, it is not possible to transfer the signatures of all vehicles that contributed to an aggregate. Moreover, the information that was originally signed may have changed during the aggregation process, because forwarding vehicles merged it with other information.

## III. Generic Model

To structure our subsequent discussion of different aggregation schemes, we will now introduce terminology and models based on Dietzel *et al.* [51], [52]. First, we define what kind of information is represented by aggregation schemes and how it is represented. Then, we introduce a reference architecture, which identifies the major components that most aggregation protocols share.

### A. Information Representation

We will briefly discuss the main information items used by aggregation mechanisms. Essentially, aggregation mechanisms deal with time-dependent information about geo-spatial regions. The original sources for information are in-vehicle sensors that sample their surroundings. As stated in Definition 1, this information is disseminated over multiple hops and possibly modified and merged on the way. All information in an aggregation scheme is represented by *aggregates*.

*Definition 2:* An *aggregate* is a tuple that contains information about a specific geographic region at a specific time period. More specifically

$$A := (\mathcal{L}, \mathcal{T}, v, Q) \in \mathcal{A}. \tag{2}$$

We distinguish three types of information in the tuple.

1) $\mathcal{L}$ and $\mathcal{T}$ are the *locator* of the aggregate. They identify a geographical region and a time period, respectively. Typical examples for $\mathcal{L}$'s representation are 2-dimensional areas or a road identifier and a location interval; $\mathcal{T}$ is usually a timestamp or a time interval.
2) $v$ is the *primary* value of the aggregate. It conveys values of in-vehicle sensors or observations done by vehicles. Examples are speed, temperature, road conditions, and parking spots.
3) $Q = q_1, \ldots, q_n$ are the *auxiliary* values of the aggregate. Such values are used by many aggregation schemes to denote the certainty or quality of information after it has been aggregated. Auxiliary values can either relate to one of the primary values (e.g., standard deviation of an average), or they relate to the aggregate as a whole (e.g., count of observations summarized in the aggregate).

The set of all possible aggregates is denoted by $\mathcal{A}$.

Usually, aggregation schemes use information items from single vehicles to bootstrap the aggregation process. These items form a subset of all possible aggregates.

*Definition 3:* An *atomic observation* is a tuple that is composed of a vehicle's local sensor values at one point in time:

$$o := (L, T, v, Q) \in \mathcal{O} \subset \mathcal{A}. \tag{3}$$

Like for aggregates, $L$ and $T$ are a geographical and a temporal locator, respectively. But for observations, they identify a specific point in space and time where $o$ was observed. For instance, $L$ can be a global positioning system (GPS) coordinate and $T$ can be a timestamp. In addition, the observation contains a primary value $v$ which represents an exact sensor reading or observed value.

The auxiliary values of observations are trivial, that is, the standard deviation is 0, count of contained observations is 1, and so forth. We included $Q$ in the definition to show that observations can be regarded as a special case of aggregates. In practical applications, the auxiliary values of observations are often omitted.

*Example 1:* To get a better understanding of aggregates and atomic observations, consider a traffic information system. The purpose of the system is to inform vehicles about average speed on different parts of the road network. To achieve that, vehicles broadcast atomic observations with speed reports, which are later aggregated and further disseminated in summarized form. For simplicity, we assume location is a one-dimensional value measured in meters and the timestamps are given in seconds relative to a fixed starting value. Three vehicles each create an atomic observation and broadcast it

$$o_1 = (500, 10, 50), \tag{4}$$

$$o_2 = (510, 20, 55), \tag{5}$$

$$o_3 = (510, 30, 60). \tag{6}$$

For $o_1$, $L = 500$, $T = 10$, and $v = 50$. Here, $v$ represents the vehicles' velocity. As stated above, we omit the auxiliary values for now. Suppose now, the traffic information calculates the merged information using all three observations. That will result in an aggregate

$$A = ([500, 510], [10, 30], 55, (5, 3)). \tag{7}$$

After merging, both the location and the time are stated as an interval: $\mathcal{L} = [500, 510]$ and $\mathcal{T} = [10, 30]$. The velocity is given as the average of the three atomic observations. Furthermore, two auxiliary values inform about the aggregate's quality: $q_1 = 5$ is the standard deviation of the average velocity and $q_2 = 3$ indicates that 3 atomic observations have been merged to create the aggregate. Note that this specific combination of parameters—intervals, average, standard deviation, and count—are only used for this example. While many schemes employ similar mechanisms to aggregate atomic observations, more complex operations are possible and will be discussed in Section IV-B.

Both atomic observations and aggregates can be extended to contain more than one value. For instance, consider an application that aggregates both average velocity and average outside temperature. Since most existing aggregation schemes focus on one type of value exclusively, we omitted the possibility for multiple primary values in our definitions.

Finally, we need to model the knowledge base of vehicles. Each vehicle has access to some subset of all observations and aggregates. The subset is composed of the vehicle's own observations and aggregates, as well as all observations and aggregates received from other vehicles.

*Definition 4:* The *world model* of a vehicle is the entirety of all information available to a vehicle $x$ at time $T$, represented by a set of aggregates

$$\mathcal{W}(x, T) := \{A_1, \ldots, A_n\} \subseteq \mathcal{A}. \tag{8}$$

The world model can contain several information items with overlapping geographical and temporal regions. Moreover, items can contain inconsistent information due to faulty sensors or malicious attacks. The world model is therefore commonly filtered before using it for application decisions or further dissemination.

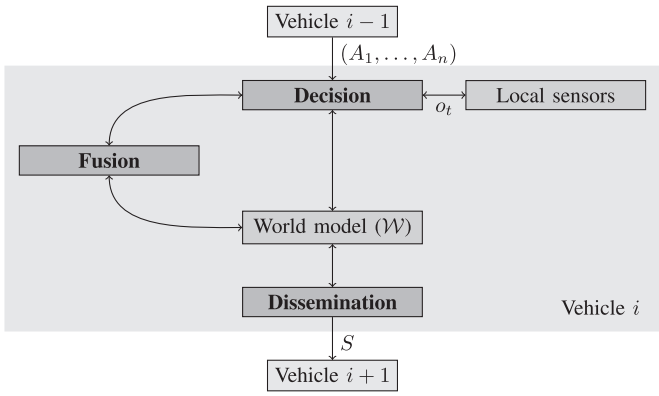



Fig. 5. Generic aggregation protocol architecture.

### B. Reference Architecture

Even though many different aggregation schemes have been proposed, we can identify a common set of components that each scheme comprises. Namely, each aggregation scheme needs to *decide* whether any number of items is "close enough" to be aggregated. Then, a *fusion* algorithm is necessary, which combines items into one new information item. Finally, information needs to be *disseminated* to other vehicles. Due to the mobility of vehicles, combined with further dissemination by receiving vehicles, which perform the same steps, the described components suffice to achieve efficient multi-hop dissemination of information. In addition, a world model data structure is needed that manages all information available to a vehicle and provides efficient querying and updating mechanisms.

Existing schemes use different arrangements of these components. Fig. 5 shows the most common arrangement. New information, either observations or aggregates, is forwarded to the decision component and compared with already known information. The decision component decides whether the items can be aggregated and, if so, forwards them to the fusion component. If not, they are directly added to the world model. The goal is to make sure that only items that are "similar enough" by a suitable metric are fused and other items are kept separately in order to preserve data quality. The fusion component then performs the actual aggregation and adds the result to the world model. Finally, the dissemination component selects a subset of the world model for further dissemination, creates messages, and disseminates them to other vehicles.

Note that in this arrangement, new information is immediately aggregated before it is added to the world model. Because many fusion methods impose some quality loss and local storage is comparatively cheap, it can be argued (cf. [17]) that all new information should be added to the world model first. Then, the dissemination component triggers the decision and fusion process. No matter which variant is chosen, the set of required components remains the same. We will now discuss the required functionality for each component.

*Decision:* The decision component compares a number of information items, i.e., members of the powerset of all possible aggregates $\mathcal{P}(\mathcal{A})$, and groups similar items for aggregation

$$\mathsf{Decision} : (\mathcal{P}(\mathcal{A}), \mathfrak{C}) \to \{yes, no\}. \tag{9}$$

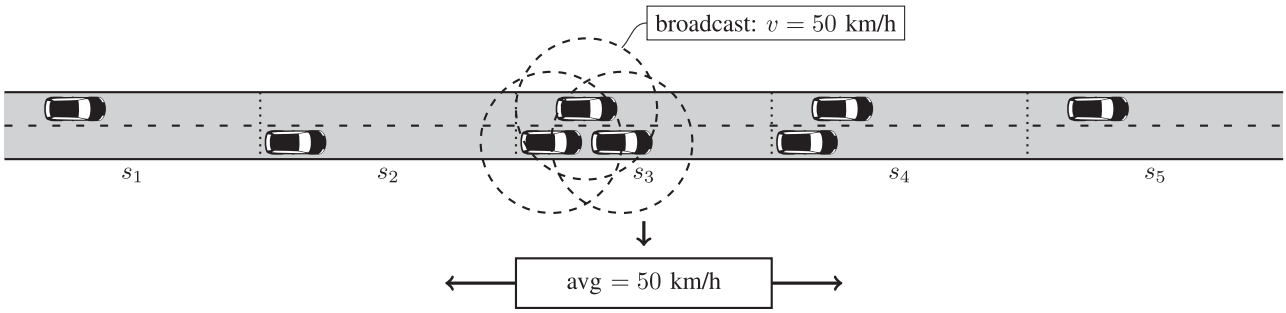



Fig. 6. Overview of SOTIS. Each vehicle broadcasts the exact speed to direct neighbors; average speeds per segment are disseminated further.

Inputs to the decision function can be new atomic observations from local sensors, new aggregates and atomic observations received from other vehicles, and aggregates or atomic observations already present in the world model. Besides the actual information items, the decision function is possibly influenced by context $\mathfrak{C}$, such as the current time, location, and driving direction of the own vehicle. Common decision criteria are geographical relation of information items and similarity of contained values, such as speed. Moreover, the decision needs to take into account temporal correlation, as well as movement direction to address the dynamic nature of VANETs. Essentially, the decision components try to group similar information in a way that is comparable to clustering mechanisms (cf. [35]). Information that is redundant or irrelevant will be subsumed in existing information or kept as separate information, respectively. Irrelevant information can then be pruned during world model maintenance. A simple decision method is to group all items from a particular road segment, thereby reducing the number of packets that need to be disseminated. More elaborate decision mechanisms reduce information granularity more and more with increasing distance between the deciding vehicle and the information region. Also, more complex rule sets can be used, e.g., fuzzy-logic-based decision rules.

*Fusion:* The fusion component performs the actual merging of information items once they have been grouped by the decision function

$$\mathsf{Fusion} : \mathcal{P}(\mathcal{A}) \rightarrow \mathcal{A} \tag{10}$$

where $\mathcal{P}(\mathcal{A})$ is the powerset of all possible aggregates. Fusion can be a lossless or a lossy process. An example for lossless fusion is the simple concatenation of values. Most existing schemes use lossy fusion in order to save more bandwidth. Several fusion functions can be used for different values contained in an information item. For instance, geographical coordinates might be fused by calculating their bounding box, that is, the smallest axis-aligned rectangle that contains all given coordinates. Speed values can be fused by calculating their average. Both examples are lossy in the sense that the original atomic values cannot be reconstructed given the result of the fusion. However, lossy fusion can help to reduce the size of packets, because the merged information may be represented more compactly.

*Dissemination:* The decision component selects a—possibly modified—subset of the world model for dissemination

$$\mathsf{Dissemination} : (\mathcal{W}, \mathfrak{C}) \rightarrow S \subset \mathcal{W}. \tag{11}$$

In order to preserve bandwidth, only the most relevant items in the world model are selected for dissemination instead of disseminating the whole world model. Like the decision function, the dissemination selection strategy takes into account the current context $\mathfrak{C}$ of the disseminating vehicle. For instance, geographically closer information can be given preference for dissemination. Moreover, the history of previously selected information items is used by some schemes to prioritize information that has not been disseminated for a while. Dissemination is most often periodic, but additional strategies, like carry-and-forward, are possible. Moreover, dissemination can further reduce the number of packets on the wireless channel because several aggregates are usually concatenated within one packet for transmission.

While implementing all components, existing aggregation schemes often focus on one of the reference architecture components. Therefore, we will group our following discussion of existing schemes according to the component that resembles their main contribution.

## IV. State of the Art

Many different aggregation schemes for VANETs have been proposed in the past decade. One of the earliest mechanisms is *self-organizing traffic information system* (SOTIS) [30], [53], which was originally introduced in 2003. We will use the SOTIS scheme as an introductory example before continuing to discuss schemes that focus specifically on decision, fusion, or dissemination. We discuss SOTIS in more detail, because it is well suited to understand the basic concepts of in-network aggregation, as well as to discuss the impact of several requirements we discussed in Section II. We conclude our state of the art analysis with an outlook on aggregation in other domains than VANETs.

SOTIS' core idea is to impose a fixed segmentation on the road network, which correlates with the wireless communication range, and only disseminate information with segment granularity in larger areas. Each segment is uniquely identified by a segment ID together with a road ID; both are assumed to be globally known. Fig. 6 shows how the SOTIS system works. In their 1-hop neighborhood, vehicles send periodic beacons containing their current position $p$, road ID $r$, a timestamp $t$, and the current velocity $v$; the observation tuple is $o := ((p, r), t, v)$. Receiving vehicles calculate the average $v$ of all speed reports of one road segment $s$ to decide on a road segment's traffic status; the aggregate tuple is $A := ((s, r), t, v)$.

All such traffic status summaries are again disseminated periodically. In contrast to individual vehicles' data, summary reports are disseminated over multiple hops. To save storage space, receiving vehicles only keep the newest summary report per road segment, assuming that this is the most accurate one. The main aggregation scheme components are implemented as follows:

- *Decision*. Atomic observations are selected for aggregation if and only if their geographic identifier is in the same road segment. Aggregates are selected for further aggregation if and only if their geographic region (i.e., road segment) is the same.
- *Fusion*. The fusion algorithm is shown in Algorithm 1. Atomic observations are merged by creating a new summary record about a road segment. The function `GetSegment` is used to determine the fixed segment ID corresponding to a given position. The time stamp is set to the current time. All atomic speed values are averaged. Aggregates are not merged further; given two aggregates, the fusion function will drop the older aggregate.
- *Dissemination*. Atomic observations are disseminated only if they were created by local sensors. In addition, a fixed number of aggregates that represent surrounding road segments are disseminated.

**Algorithm 1:** SOTIS:Fusion($A_1, \ldots, A_n$)

**Input**: A set of aggregates $\{A_1, \ldots, A_n\} \subset \mathcal{A}$.
**Result**: An aggregate $A$ that represents the merged data of all aggregates.

**if** $A_1, \ldots, A_n \in \mathcal{O}$ **then**
  $A \leftarrow$ ((`GetSegment(`$p_1$`)`, $r_1$), `GetCurrentTime()`, $\frac{1}{n}\sum_{i=1}^{n} v_i$)
**else**
  $A \leftarrow A_{\text{argmax}_i(t_i)}$
**end**
**return** A

By disseminating only summarized information about road segments, self-organizing traffic information system (SOTIS) achieves a much higher awareness of the current traffic situation than dissemination of atomic information would allow for. Moreover, self-organizing traffic information system (SOTIS) reduces the number of packets that are sent over the wireless channel. However, the implementation has a number of drawbacks. First, self-organizing traffic information system (SOTIS) uses only one level of aggregation hierarchy. Thus, communication overhead is reduced by a constant factor, but still grows linearly with the area that information is communicated about. Therefore, self-organizing traffic information system (SOTIS) does not scale to large areas, as proven in [22]. Furthermore, data quality can suffer depending on the exact segment size chosen. Suppose a segment size of 1,000 meters is chosen to allow a wide range dissemination of the traffic status. Now a segment where all cars drive at approximately 80 km/h will look the same after aggregation as a segment where 30 cars stand still and 120 cars drive 100 km/h. Yet, those two traffic situations are likely to result in different routing decisions if the exact values were known. If smaller segments are chosen, the average is more likely to be accurate, but it can still be skewed. Thus, self-organizing traffic information system (SOTIS) is not able to keep data quality requirements. In addition to possible over-aggregation, there is also no duplicate filtering in place. When vehicles calculate the segment average in the fusion step, they might consider more than one report with different timestamps per vehicle, hence biasing the aggregated average. Finally, self-organizing traffic information system (SOTIS) does not look at the possible influence of malicious attackers on the aggregation results.

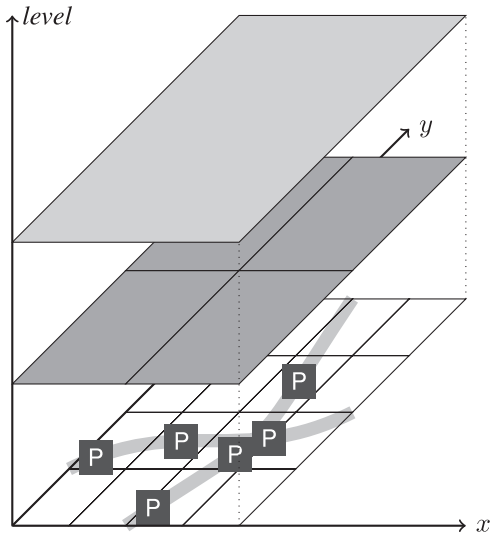



Fig. 7. Hierarchical aggregation based on distance of observations.

Following SOTIS' publication, several improved protocols were proposed. These papers usually focus on one of the reference architecture components and implement an improved version. In addition, proposals often focus on a specific aspect of an aggregation component, which relates to one or more of the requirements discussed in Section II. We will now discuss proposals for each component in turn, concentrating on the main novel aspect of each paper discussed.

### A. Decision

*1) Hierarchy:* One of the major drawbacks of SOTIS is the missing hierarchical aggregation structure. A flat aggregation structure confines the aggregation benefit to be constant. Caliskan *et al.* [17] present an aggregation scheme that uses a quadtree structure to disseminate information about free parking spots in urban scenarios. That is, at the lowest level, the city is divided into non-overlapping cells. On each higher level, four lower-level cells are combined to form one higher-level cell. Fig. 7 shows the corresponding aggregation process. The idea is to disseminate more exact information about free parking spots in the local vicinity and only disseminate coarse information in larger areas. For instance, simulations show that at 7000 m distance to a parking spot, more than 60% of the information is represented with the coarsest granularity. As atomic observations, current occupancy is reported by payment units at each parking area. Vehicles further aggregate the available parking spaces whenever information of all four grid cells is available. Because more accurate lower level aggregates are available for

the direct vicinity, navigation decisions can be refined as a vehicle approaches its target area.

The proposed mechanism is able to better adapt the used bandwidth, and it reduces the number of packets that are communicated further, due to the introduced hierarchy. However, the proposed grid structure is static and does not adapt to the local situation. For instance, parts of a city with a high street density or a high demand for parking spots might benefit from a finer resolution than suburban areas. This cannot be reflected by the static quadtree structure.

Thinking beyond the parking space application, an approach along these lines is not well suited for traffic information, because the aggregation hierarchy describes larger and larger two-dimensional squares, whereas traffic jams and other traffic information is best described using subsets of a reference road network. Lochert *et al.* [18] define an aggregation hierarchy that is closer to the needs of a traffic information system. Instead of dividing a city according to a quadtree structure, a hierarchical set of landmarks and interconnections between them is defined. At the lowest level, all road intersections are landmarks, and the interconnections between these landmarks are the road's segments in the real world. The exchanged information describes the travel times currently required along these road segments. Travel time serves as a combined metric of speed and distance, supporting routing algorithms well. On higher and higher aggregation levels, the set of landmarks is restricted to smaller and smaller sets of central points in the road network (i.e., major intersections). The landmarks on higher levels are connected by virtual long-range interconnections. The exchanged information is the travel time along the currently best route. Simulations show that reductions of average travel time of up to 10% can be achieved using the proposed protocol.

Like for the quadtree approach, the hierarchy of landmarks needs to be globally known. In the context of road networks, this problem is worse, because changes to the landmarks can occur more often, especially on lower levels. Moreover, the landmark scheme is based on a fixed hierarchy and a fixed set of landmarks on all levels. It therefore cannot flexibly adapt to the current situation. The following papers introduce more flexible aggregation decision functions. In addition to possible hierarchic aggregation, these adapt to quality requirements flexibly, resulting in a dynamic segmentation of the road network.

*2) Data Utility:* The TrafficView system introduced by Nadeem *et al.* [19] (also in [20] and [21]) introduces data quality considerations for aggregation decisions. The goal of TrafficView is to disseminate traffic status information over larger areas. In a later paper [27], the authors extend their system by comparing different data dissemination strategies, taking into account traffic driving in the opposite direction. Atomic information items consist of vehicle IDs, their current position, current time, and speed. Aggregated information items describe clusters of close-together vehicles using a list of vehicle IDs, their averaged position, speed, and the time of the aggregate's oldest input information. Note that while position, speed, and time are aggregated in a lossy fashion, the list of IDs will still grow linearly in the number of vehicles. To reach aggregation decisions, the road ahead is divided into a number of regions, and for each region the aggregation rate and the portion of space to use are configured. The aggregation rate determines the number of atomic values that should be combined into aggregates. The portion can be used to express the importance of different regions and determines the region size. Inside the regions, aggregation decisions are made using a cost-based metric. The "cost" of aggregation is designed such that it is high whenever aggregation would introduce a large error. Then, information items are combined such that minimal cost is induced. Results show that the proposed protocol increases visibility for vehicles: 50% of all vehicles know about the surrounding 575 m of road when using TrafficView; their visibility is approximately 525 m without.

The idea of cost-based aggregation is used by many following papers (e.g., [16], [18], [54]), because it allows the aggregation system to adapt to different traffic situations. This approach is comparable to advanced audio or image compression techniques, such as MP3 or JPEG, in the sense that it uses intrinsic properties of the data to be compressed to achieve high (lossy) compression rates with minimal quality loss. However, the specific implementation used by TrafficView still leaves the list of vehicle IDs in an aggregate as a linearly growing component. Moreover, only an average position value is stored for all vehicles contained in an aggregate. For large-scale events, such as long traffic jams, the TrafficView aggregates would both contain large lists of vehicle IDs, and the aggregate's average position would not characterize a traffic jam well. Therefore, TrafficView is only suitable for aggregating a small number of vehicle records, increasing the visibility in the local scope.

*3) Flexibility:* Taking TrafficView's flexible road segmentation one step further, Dietzel *et al.* [16] introduce a fuzzy-logic-based aggregation scheme. The main idea is to completely abandon any fixed road segmentation. Instead, similar to TrafficView, information items are aggregated dependent on whether they are "similar". Because similarity metrics can depend on application requirements and can potentially depend on several influences, they are expressed using a set of fuzzy logic rules [55]. For instance, the authors take the distance between two information items and the standard deviation of the aggregated speed average into account for their aggregation decisions. These real-valued influencing factors are then evaluated using a set of fuzzy logic rules to reach aggregation decisions. An example set of rules is:

if **STANDARD_DEVIATION** is LOW and
**LOCATION_DIFFERENCE** is SMALL then
**AGGREGATION_DECISION** is YES

In this example, the standard deviation and the location difference are weighed off against each other. The fuzzy logic rules abstract from specific value comparison, making it easier to express rule sets for aggregation decisions. Because there is no underlying fixed road segmentation, aggregated areas can be of any size. For instance, a few slow-moving vehicles can be represented by a small aggregation area while, at the same time, long traffic jams can be represented by one large aggregate. Simulation results in the paper show that the mean deviation from the real situation is 5 km/h after 60 s simulation time with

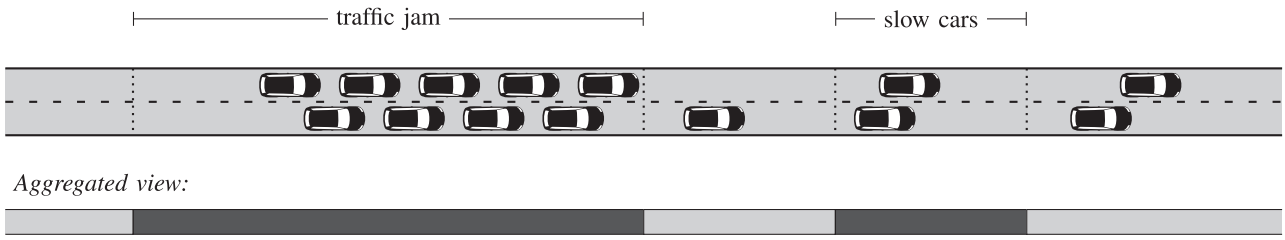



Fig. 8. An example road with a long stretch of homogenous traffic due to a traffic jam and a short stretch of slow speed due to slow vehicles.

flexible aggregation, whereas it is approximately 18 km/h with fixed segments. Fig. 8 shows an example aggregated view of a road section.

Kumar and Dave [36], [37] propose to apply ideas from multi-criteria decision making systems (MCDMs), originally proposed in the operations research domain [56], and use them to achieve flexible aggregation decisions. In a two-step process, first all information available is evaluated for similarity using criteria like position distance or speed difference. A $k$-d-tree [57] data structure is used to index information items using their similarity scores, which allows for efficient range queries. Then, application preferences and requirements are used to calculate the best aggregation strategy among several available strategies. The proposed approach is very flexible; however, the authors only vaguely discuss how to select suitable criteria for specific applications, such as traffic information systems. Zhang *et al.* [38] discuss fuzzy clustering as an alternative way to achieve flexible aggregation decisions. Each vehicle calculates similarities between known values and then clusters similar messages. Only summaries of the aggregated clusters are further disseminated. In their evaluation, the authors argue that their scheme performs significantly better than schemes using fixed segmentation, because vehicles that are close-by but part of different traffic situations are not aggregated using the clustering technique.

An approach that also aims at a dynamic, situation-dependent depiction of a road is proposed by van Eenennaam and Heijenk [54]. Borrowing ideas from run length encoding and pulse code modulation (cf. Salomon [58] and Waggener [59]), the authors try to represent the traffic situation on a long stretch of the road using only a small subset of representative atomic values. All information is forwarded multi-hop; however, not every vehicle adds its own information to the list of forwarded values. Instead, a threshold function is used so that only vehicles with a speed that deviates significantly from the last entry in the list add their current speed and position. The resulting sampling approximation is shown in Fig. 9. When interpreting the sampled subset, each entry is valid for the whole road interval between itself and the following entry in the list. In order not to over-emphasize outliers, each new sample added to the list does not represent a single vehicle alone, but represents the averaged speed of a set of close-by vehicles. Thus, the number of samples needed is further reduced, because local noise in form of differing speeds is filtered out. Schwartz *et al.* [60] extends the presented scheme to work with different traffic situations on multiple lanes, as well as to support road networks instead of only a single road.

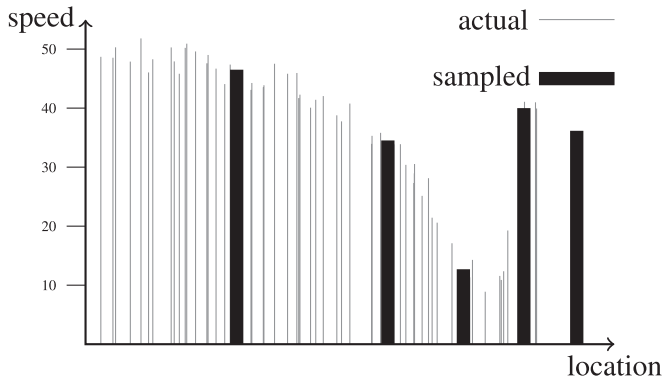



Fig. 9. Traffic approximation using a subset of representative values.

### B. Fusion

*1) Lossless:* Given a positive decision to merge two items of information, we need a method to actually perform the fusion. Fusion methods can be categorized into *syntactic* and *semantic* approaches, as explained in Section II-B. *Syntactic* aggregation uses techniques to compress the data from multiple vehicles in order to fit the data into a single communication packet. This compression results in lower overhead than sending each message individually. In *semantic* aggregation, the information from individual vehicles is summarized. For instance, instead of reporting the exact location of five vehicles, only the existence of five vehicles is reported [61]. While syntactic compression allows for lossless reproduction of the original data, bandwidth savings are limited. Semantic compression offers higher compression rates at the cost of information loss.

Well-known syntactic data compression techniques, such as ZIP [62] or LZW [63], are too generic and achieve too little bandwidth gain to be useful for data aggregation in VANETs. Still, approaches exist that try to achieve sufficient compression using syntactic, invertible compression. For instance, CASCADE [23], [24] uses a variation on differential coding, which is a technique that is frequently used in the multimedia domain to encode audio signals (cf. [64]). Algorithm 2 shows CASCADE's implementation of difference encoding. First, a location parameter (such as mean or median) is calculated for all items that should be aggregated. Namely, the positions of all aggregates $A_i$ are regarded as corners of a polygon and its center $(X, Y)$ is calculated, as indicated by the `GetCenter` function. Likewise, $V$ is calculated as the median of all speed values that should be merged. Then, differential encoding is used to encode the differences between atomic observations and overall cluster data. That is, for each coordinate its $x$ and $y$

axis difference to the center location $(X, Y)$ is stored and for each speed the difference to the median speed $V$ is stored. The resulting aggregate consists of the location parameters $(X, Y)$ and $V$, as well as all the calculated differences.

**Algorithm 2:** CASCADE:Fusion($A_1, \ldots A_n$)

**Input**: Aggregates $A_1, A_2, \ldots, A_n \in \mathcal{A}$.
**Result**: An aggregate $A$ that represents the merged data of all aggregates.

$(X, Y) \leftarrow \texttt{GetCenter}(A_1, \ldots, A_n)$ ;
$V \leftarrow \texttt{Median}(v_1, \ldots, v_n)$ ;
**foreach** $i \in \{1, \ldots, n\}$ **do**
  $\Delta x_i \leftarrow X - x_i$ ;
  $\Delta y_i \leftarrow Y - y_i$ ;
  $\Delta v_i \leftarrow V - v_i$ ;
**end**
$A \leftarrow ((X, Y), V, \Delta x_1, \ldots, \Delta x_n, \Delta y_1, \ldots, y_n, \Delta v_1, \ldots, \Delta v_n)$
**return** $A$

The authors claim that CASCADE achieves a compression ratio of at least 86%. Instead of aggregating all differences as in CASCADE, van Eenennaam and Heijenk [54] propose to consider only data deviating from the baseline by more than a defined threshold. Their approach reduces packet size and storage space, and thus the communication overhead.

*2) Lossy:* Also, trying to filter out less relevant information, Zooming [65] is a technique based on discrete cosine transform (DCT) [66] that aims at reducing the computation overhead. A DCT transforms a sequence of data points to the frequency domain. There, they are expressed in terms of a sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications, notably for lossy compression of audio (e.g., MP3) and images (e.g., JPEG). Fig. 10 shows the process for VANETs. After applying DCT, coarse-granular information is represented by low frequency information and details of an aggregate are represented by high frequency information. After filtering out high frequency information, a zoomed-out aggregate can be obtained that represents a coarser view of the information and can be efficiently encoded. This zoom-based approach permits to provide fine-grained data for nearby drivers and coarse-grained data for drivers further away. The authors claim that their DCT-based approach performs significantly better than a fixed hierarchical grid (e.g., [17]): with 50% of the communication overhead, they increase the visibility of surrounding parking spots by 67% over [17] (in terms of available information with the finest granularity within 600 m distance).

Besides efficient compression, a notable problem of data fusion methods is that of duplicate message due to redundant sensor readings or cycles while forwarding information. Due to the decentralized nature of aggregation mechanisms, vehicles often receive aggregates and cannot tell whether they have already contributed to the contained information. As a result, vehicles might add their own observations multiple times. The result is a biased aggregate that reduces information quality, as discussed in Section II-D.

*3) Duplicate Filtering:* Lochert *et al.* [31] propose to use a modified version of Flajolet-Martin sketches (FM sketches) [67] to solve the duplicate counting problem. Their approach is later refined in [68]. Originally, FM sketches were used to estimate the number distinct elements in large databases with constant per-element effort and without the need to pass through the data multiple times. The price for these features is the loss of exact counting: FM sketches provide a probabilistic approximation of the number of distinct elements.

Fig. 11 shows the data structure used for FM sketches. A bit field $S = s_1, \ldots, s_w$ of length $w \geq 1$ is used as an approximation of a positive integer. The bit field is initialized to zero at all positions. To add an element $x$ to the sketch, the element is hashed by a hash function $h$ with geometrically distributed positive integer output, where $P(h(x) = i) = 2^{-i}$. The entry $s_{h(x)}$ is then set to one. The key to duplicate-insensitivity of the FM sketch is that, regardless of the number of times an identical object is inserted, the same bit in the sketch is always set. To obtain the estimated value of the sketch, the length of the first uninterrupted sequence of 1 bits $l$ is counted. Fig. 11 shows an example sketch of size $w = 8$ with $l = 3$ bits set to 1 before the first 0 occurs. The estimate is $E = 2^l/\varphi$, where $\varphi \approx 0.77351$ is a constant. Multiple sketches can be used to further reduce estimation error.

In order to automatically remove old information from aggregates, Lochert *et al.* [68] modified the original FM sketches. They use counters of $n$ bits length instead of single bits at each index position. These counters represent a time to live (TTL) in the range $[0, \ldots, 2^{n-1}]$ for that bit. The use of counters serves to combine duplicate insensitivity with the possibility to continuously update aggregates, as discussed in Section II-D. In their paper, Lochert *et al.* [68] use FM sketches for counting. However, the sketches can likewise be used for sums and, by combining sum and counting sketches, for calculating averages.

Zekri *et al.* [69] also rely on FM sketches for duplicate-insensitive data fusion. However, their scheme does not use the TTL-adapted version. Instead, it keeps a number FM sketches, each representing a time interval for a road segment. This approach allows to keep a history of events for each road segment. However, keeping separate FM sketches for each time interval results in high storage requirements; the authors note that the size of their data structure is 22 megabytes for a reasonable area of interest and one week of historic data.

### C. Dissemination

*1) Clusters:* After being aggregated the data is disseminated. Central challenges for dissemination are how to select the set of forwarding vehicles and the set of target vehicles such that all interested vehicles get the information they require while only a minimal set of vehicles forwards messages to avoid too much redundancy.

A typical approach used in WSNs [70] is to introduce a hierarchy, which often resembles a tree, among all nodes and use it for dissemination. Such an approach works well if few central nodes need to collect information from a large set of sensors. However, it is not suitable for VANETs where each vehicle both observes information and wants to receive information.

Cluster-based aggregation, for instance employed by Raya *et al.* [71], requires the election of a cluster head that will be responsible for aggregating, controlling, and sending

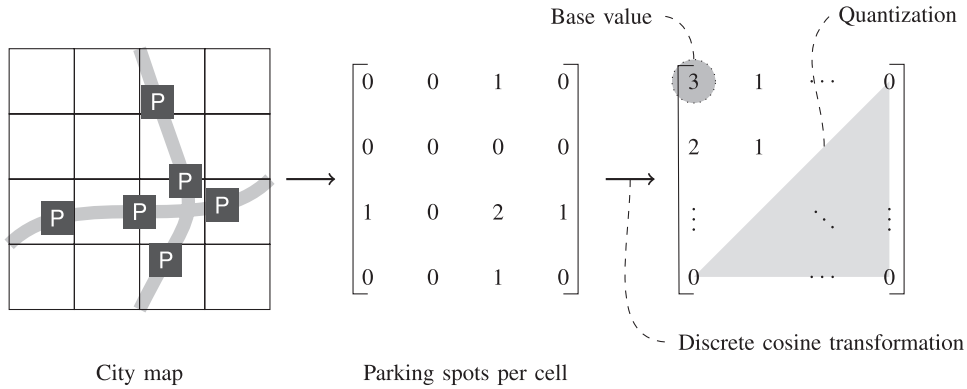



Fig. 10. Application of a discrete cosine transform (DCT) to traffic data. From left to right: map with streets, corresponding parking spots per cell, and parking spot information after DCT transformation. After the DCT, a quantization is applied that filters high-frequency information.

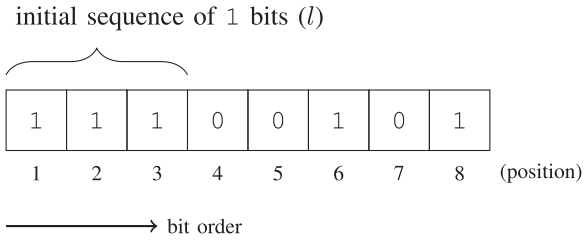



Fig. 11. Example FM sketch.

information to members of the cluster. In their paper, Raya *et al.* use the one-hop communication range as the basis for cluster formation. Vehicles that are not cluster heads only disseminate their local observations and receive aggregated information. The cluster head appears as a single point of failure and the efficiency of the dissemination relies on the stability of the clusters. If cluster membership changes frequently, the bandwidth required for maintenance may limit the bandwidth savings of aggregation. Aiming to eliminate frequent cluster reorganization, Dietzel *et al.* [72] introduce data-similarity based clustering, which—inspired by aggregation decisions—merges vehicles with similar speeds.

Saleet *et al.* [73] propose the *region-based location service management protocol* (RLSMP), which considers the use of message aggregation and geographical clustering to minimize the number of location updates and querying messages for location management of vehicles. The map is segmented according to a fixed grid, which is known by all vehicles. Each cell in the grid contains a particular vehicle, called a cell header, that is responsible for aggregating the location information about all vehicles in the same cell. The authors indicate that, although message aggregation will improve scalability, it can also lead to more packet collisions and retransmissions, because larger packets are transmitted. Moreover, delays are possible as the information must be buffered for aggregation before being sent.

*2) Relevance:* Due to the maintenance overhead of hierarchy-based solutions, other proposals use more decentralized approaches for information dissemination. A common approach is to use relevance criteria to decide whether to disseminate information in a certain region or not. The distance between target vehicles and the area that the disseminated information is about [74] is a prime example for such relevance criteria.

Cuckov and Song [75] propose a structureless information dissemination scheme. It works in three phases: observation, one-hop dissemination, and aggregate dissemination. Local views of vehicles are disseminated in a fixed region using geocast. At points where new information is available to be added, information is aggregated. With increasing distance to the event location, the information resolution is reduced. Similarly, Dietzel *et al.* [16] employ relevance criteria as proposed in Eichler *et al.* [76] during message dissemination. The idea is that all vehicles keep a local world model with all known information. However, the size of the world model will be too large to be disseminated periodically with high frequency. Therefore, a set of weighing functions is introduced to prioritize information. Criteria used are distance to the event, timeliness of information, and others. Whenever information is to be disseminated, only the subset of the world model with the highest weights is selected.

*3) Abnormality:* But even in a relevance-based scheme, bandwidth is possibly wasted to communicate information of low utility. For instance, it is arguable whether information about regions of free-flowing traffic needs to be disseminated at all or whether it is enough to disseminate information about abnormal traffic situations, such as traffic jams. Chen [77] propose a dissemination scheme where only vehicles with abnormal information communicate. Other vehicles remain silent. All vehicles driving on the same road segment are regarded as a cluster to generate traffic message about this segment. To reduce communication overhead, only abnormal traffic data is spread to nearby road segments. For this inter-segment dissemination, epidemic routing is used [78]. By employing event-driven and periodic mechanisms, the abnormal traffic messages are disseminated on time to vehicles that might need it. Of course, any such scheme has an inherent drawback: a vehicle which does not receive information on a specific region cannot know whether the situation is normal, or whether communication has not (yet) succeeded.

Shafiee and Leung [79] also propose an anomaly-based protocol, which considers fixed size and fixed locations for segments. As a requirement, each vehicle is assumed to be equipped with a digital map that includes average speed values for each road segment. In addition, vehicles continuously broadcast their average speed to 1-hop neighbors. When the difference between a beacon-based average value and its

corresponding value in the map exceeds a threshold, this violation will be notified to other one-hop neighbors through the next beacon messages.

*4) Carry-and-Forward:* The approaches discussed so far mostly address the issue of how to select the most suitable information for dissemination in order to provide the most useful information to receiving vehicles. An orthogonal problem during data dissemination is how to maximize the chance that aggregatable information meets in the dissemination process. Most schemes use periodic beaconing using fixed beacon intervals for dissemination, which may not be optimal. In contrast, Catch-Up as described in [80], [81] proposes to adaptively change the forwarding delay according to the current context. The idea is that vehicles wait before forwarding information in order to aggregate it with redundant or similar information about the same events. Results show that Catch-Up reduces the number of reports by approximately 50% at 3 km distance to the observed event when compared to randomized waiting times. Data aggregation is based on fixed size road segments. Because of the introduction of a delay, Catch-up is not suitable for delay-sensitive applications.

### D. Information Integrity Protection

Proper protection of data integrity is traditionally a weak point of naïve aggregation scheme implementations, as discussed in Section II-G. Mohanty and Jena [82] have published a survey on secure aggregation. However, the existing survey lacks clear discussion of requirements for secure aggregation, as well as a categorization for different security approaches. We therefore provide an overview of current integrity protection approaches in the following.

*1) Trusted Hardware:* Picconi *et al.* [61] proposed one of the first mechanisms to secure aggregation in VANETs. Their mechanism is based on probabilistic, interactive verifications. Whenever a car receives an aggregated record, it reads the claimed number of participants $n$. A random number $r \in \{1, \ldots, n\}$ is selected, and the sender of the aggregate is challenged to provide the $r$th atomic observation, including a signature of the original observer as proof that the aggregation was performed correctly. Given a correctly signed atomic observation, the receiver checks whether it is plausible that the atomic observation was used in the aggregate. Typically, this means that the location must be within the aggregate area, among other checks. If an attacker takes into account some atomic observations but alters others, she is detected with probability $f/n$ where $n$ is the number of claimed atomic observations in the aggregate and $f$ the number of fake observations.

As it is explained above, the protocol's integrity relies on a form of commitment scheme. First, the message with the claimed aggregate is sent. Then, the receiver replies with a second message asking for atomic observations. Thirdly, the sender reveals the requested information. Requiring 3 rounds of communication, however, can be difficult in highly mobile vehicular networks. Therefore, the authors propose to use a tamper-proof device. The tamper-proof device acts as a proxy for the receiver inside the sender's car. To send an aggregate, the sending car forwards it to its own tamper-proof device. The device then performs the random challenge and broadcasts the results. Even if the software outside the tamper-proof device cannot provide a response to the challenge, the aggregate is sent as a proof of malicious behavior. It is therefore crucial that an attacker cannot prevent the tamper-proof device in her own controlled car from sending packets. Otherwise, the malicious behavior proofs can be kept from being sent. Given the numerous possibilities of jamming attacks, this guarantee is hard to achieve.

Ibrahim *et al.* [25], [26] propose a security mechanism tailored to their CASCADE protocol (see Section IV-B). They assume that trusted computing is employed to tamper-proof the aggregation algorithms, i.e., fusion, decision, and dissemination. Therefore, an attacker cannot modify information about larger areas. Putting the whole aggregation algorithm inside tamper-proof hardware is a possibility that Picconi *et al.* [61] discussed as well. However, they argue that maintaining the whole aggregation logic inside tamper-proof devices is too costly.

Assuming the aggregation process to be tamper-proof, Ibrahim *et al.* only assume the on-board GPS devices to be unprotected. That enables an attacker to disseminate false location information. To detect these attacks, a combination of signal strength measurements and additional laser distance measurements is employed. This is an example for purely plausibility-based attack detection, and it only works in the local vicinity, which explains why the authors employ trusted computing to secure dissemination of aggregated reports in wider areas. To exclude misbehaving vehicles from the network, quarantine messages are employed. These messages are received by a trusted component in attacker vehicles which then disables the attacker's radio devices.

*2) Cryptography:* The previous approaches use a combination of interactive detection and plausibility-based detection. However, both approaches rely on trusted hardware. In contrast, Raya *et al.* [71] propose a mechanism that employs cryptographic protection mechanisms without the need for trusted hardware. Their scheme assumes that the underlying aggregation mechanism works similar to SOTIS (see beginning of Section IV). Having agreed on an average value per road segment, the goal is to protect that value against further modification. To achieve this protection, Raya *et al.* discuss three different signing mechanisms, which trade off between computational overhead and communication overhead. Fundamentally, the scheme is limited by the requirement to have fixed segments and non-hierarchical aggregation. In all variants of the signatures, unique identities are required to thwart Sybil attacks [47], which reduces driver privacy.

Similar to Raya *et al.*, Dietzel *et al.* [40] rely on cryptographic signatures as a basic trust anchor. However, the scheme is based on a more flexible underlying aggregation scheme [16], which we discussed in Section IV-A. Additionally, probabilistic verification and plausibility checks are employed to reduce the number of signatures added to each aggregate. The core idea is to keep a subset of all underlying atomic observations including their signatures for each aggregate. If an aggregate is not modified, the atomic observation values should relate closely to the aggregated values; for instance, their speed should be

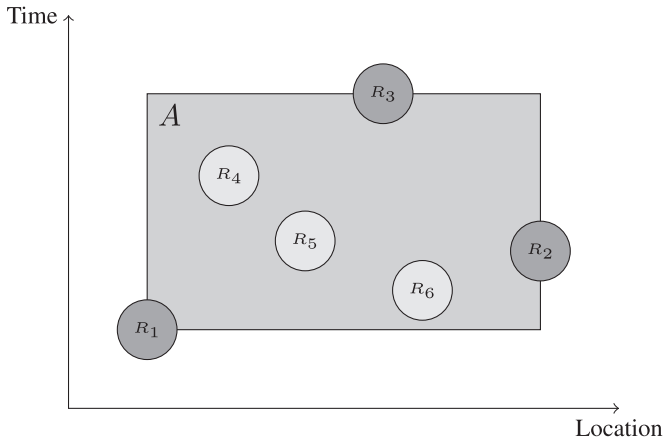



Fig. 12. Distribution of signed atomic observations for probabilistic verification.

similar. Because the atomic observations are signed, an attacker cannot easily forge them. Fig. 12 illustrates the concept using an example aggregate $A$, which is represented as rectangle to indicate the time period and location interval it covers. $R_1$, $R_2$, and $R_3$ are included to serve as proof for the aggregate's extend, because they originate from the aggregate's geographic borders. In addition, $R_4$, $R_5$, and $R_6$ serve as proof that the claimed conditions also exist in the whole area covered by the aggregate.

Each vehicle that receives an aggregate checks that both the distribution of signed atomic observations is sufficiently uniform and that the values in the observations match the claimed values in the aggregate. As a result, the confidence in the aggregate's correctness is not absolute, but represented as a percentage. Depending on the desired security guarantees, the number of atomic observations per area can be adjusted to be larger or smaller.

Molina-Gil *et al.* [83] enhance the scheme by introducing probabilistic verification to cope with the processing load due to signature verification. For each message received, only a random subset of the presented signatures is checked. Further, the authors propose to adjust the granularity of included atomic observations according to the type of road. For instance, less observations should be added on highways.

*3) FM Sketches:* All mechanisms discussed so far can be applied to arbitrary fusion functions. However, many newer aggregation mechanisms specifically employ various kinds of sketches for data fusion. These data structures offer desirable properties, like duplicate insensitivity (cf. Section IV-B). Consequently, several newer secure aggregation schemes explore the possibilities to secure sketches against manipulation.

Garofalakis *et al.* [84] offer a straightforward security mechanism for FM sketches. The authors assume an aggregation mechanism where only binary events are aggregated (e.g., "an accident occurred at position $x$") and all vehicles can either claim to witness the event or not. Moreover, it is assumed that an upper bound for the total number of vehicles in the network is known. For each bit set to 1 in a sketch, a proof is kept that contains the node ID that set the bit to 1, the bit position in the sketch, the vehicle's atomic observation (i.e., sensor values), and the vehicle's signature on these values. This approach protects against inflation of the FM sketch value to the extent

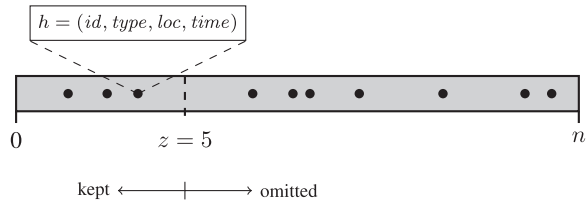



Fig. 13. Structure of a secured $z$-smallest approximation.

that an attacker needs to provide a valid signature from different vehicles for each bit set to 1. Deflation protection is achieved by adding a second FM sketch that counts the complement value, that is, $N - v$ where $N$ is the expected upper bound of the value and $v$ is the value to be counted.

Hsiao *et al.* [85] propose to use $z$-smallest probabilistic counting instead of FM sketches to proof the occurrence of events. The idea of $z$-smallest is that, given $n$ elements uniformly distributed between 0 and 1, the $z$-smallest element gives an approximation of $n$ by calculating $z/c$ where $c$ is the value of the $z$-smallest element, as shown in Fig. 13. To protect against inflation, each vehicle signs a hash of its vehicle id, the event type, location segment, and time. Only the $z$-smallest signatures are kept with the aggregate. The idea is that an attacker cannot produce enough signatures on hashes that fall into the $z$-smallest values. Therefore, an attacker cannot artificially increase the result. There is no deflation protection in this scheme, because the authors argue that an attacker will only try to produce fake events, such as a fake accident, and not try to hide events.

Han *et al.* [86] further optimize the bandwidth usage of secure FM sketches. The assumption about the underlying aggregation mechanism is that an average value, e.g., a speed average, is calculated for fixed road segments. Moreover, the authors assume that vehicles communicate their reports exclusively to roadside units, which forward them to a centralized traffic management center (TMC). Thus, the vehicles are only assumed to share a key with the TMC. Only the TMC—and not vehicles inside the network—can verify the proofs on the presented aggregates. Using these assumptions, all explanations in the paper assume symmetric cryptography, i.e., message authentication codes (MACs), as signatures.

For inflation protection, the same approach as Garofalakis' is used: whenever a vehicle sets one of the FM sketch bits to 1, it attaches a signature on the event id, event segment, time, and vehicle ID. However, some of the signatures are merged to save additional space. Namely, signatures on the initial uninterrupted sequence of 1-bits in the FM sketch are combined. Only signatures on 1 bits that are not part of the initial sequence are kept separate, because the position of these bits cannot be predicted. To achieve deflation protection, the vehicle that calculates the aggregate initializes a hash chain using its signature on the aggregate segment ID and time. Then, a one way function is applied $i$ times where $i$ is the length of the initial 1 bit sequence, forming a hash chain. If the aggregate is merged with other aggregates, and the 1 bit sequence gets longer, any vehicle can apply the one way function again. However, an attacker cannot invert the one way function to deflate the sketch estimate. The TMC can, because it shares a

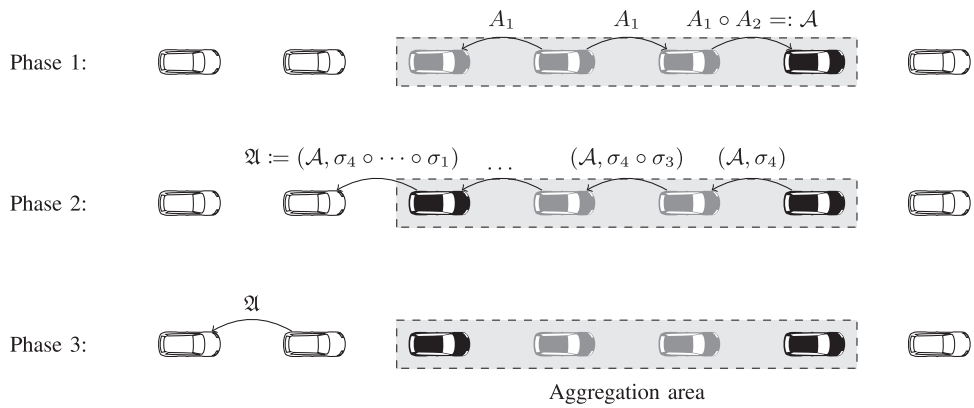



Fig. 14. Overview of SeDyA's three phases.

key with each car, recalculate the initial signature on the event and time and verify the hash chain. The size of the hash chain is constant, as opposed to the combined size of all signatures on the complement FM sketch in Garofalakis' proposal.

The SeDyA mechanism [87] further improves the previous approaches. The authors employ a 3-phase approach for combining flexible aggregation with cryptographic integrity protection, as shown in Fig. 14. First, vehicles disseminate atomic observations about their current position, time and speed. Observations are forwarded and aggregated with other observations until the edge of a homogenous speed stretch is detected. In this phase, integrity-protected FM sketches, like in Garofalakis' proposal are used. In the second phase, the preliminary aggregate is disseminated again to all participating vehicles. In this phase, each vehicle signs the preliminary aggregate to attest its correctness. Finally, a third phase is used to disseminate the finalized, signed aggregate to other vehicles that are not driving within the aggregation area. Even though the approach uses multi-signatures [88] and identity-based cryptography [89] to keep security overhead low, the protocol's three phases impose a considerable bandwidth overhead, although they achieve good protection against aggregate manipulations while still allowing for dynamic road segmentation.

Because cryptographic proposals fail to provide good security against insider attackers with low bandwidth requirements, researchers have started to analyze different integrity protection approaches, which rely on data consistency checking. One important factor to consider is data redundancy: if aggregates are received through many forwarding paths in parallel, it is less likely that an attacker can control all of them. Dietzel *et al.* [90] have analyzed the dissemination redundancy of a simple aggregation protocol, showing the future potential of such data-consistency-based security mechanisms.

### *E. Privacy*

Unlike security, few papers deal with privacy aspects of in-network aggregation explicitly. As discussed in Section II-F, insecure aggregation is commonly considered beneficial for privacy, but beyond initial research, no thorough analysis has been done. Secure aggregation, however, eliminates many of the privacy-preserving aspects of aggregation, namely, almost all previously presented schemes add several cryptographic signatures to aggregates to detect attacks on integrity. To eliminate Sybil attacks, unique identities are often assumed. For all vehicles whose signatures are added to the aggregate using unique identities, privacy is lost. For other VANET communication protocols, which do not aggregate information, so-called pseudonym schemes are often used to protect vehicle identities. In pseudonym schemes, each vehicle essentially uses a number of signing keys in turn to avoid being tracked by correlating multiple messages that were signed with the same key (cf. [13, Ch. 9]). However, a secure aggregation scheme that fully supports such pseudonym schemes does, to the best of our knowledge, not exist yet.

CARAVAN [91] introduces group keys to aid location privacy. Data aggregation is considered to further help communication overhead. The assumed underlying aggregation is a non-hierarchical aggregation within segments that span one communication hop. Information is forwarded to roadside units and then further to a traffic management center. Nodes do not directly disseminate aggregated information in the network. Within this setting, vehicles within 1-hop communication range form a group and elect a group leader, which is changed periodically. Having established a group encryption key, all group members only communicate their exact observations to the group leader and encrypt the communication using the group key. Then, the group leader aggregates the information and broadcasts the result to the nearest roadside unit. Because the group members do not need to broadcast information themselves, they cannot be tracked as easily.

### *F. Aggregation in Other Domains*

Besides the wide body of work in the context of VANETs, in-network aggregation has been intensively discussed in other domains, including WSNs and participatory sensing applications. However, the aims of aggregation, as well as the employed algorithms, are very different in these areas.

Aggregation mechanisms in WSNs focus on reducing the energy consumption of the nodes to increase network lifetime. From an algorithmic point of view, several formal definitions and requirements are similar to vehicular networks. For instance, order and duplicate insensitivity—originally formalized in [34]—has led to several schemes using FM sketches [67] (cf. Section IV-B). Communication patterns and requirements

TABLE II
OVERVIEW OF AGGREGATION SCHEME PROPERTIES

| | Scheme | SOTIS [30] | Caliskan et al. [17] | Lochert et al. [18] | TrafficView [19] | Dietzel et al. [16] | MCDM-based [36, 37] | Zhang et al. [38] | van Eenennaam and Heijenk [54] | CASCADE [23] | Zooming Chang et al. [65] | Sketches [31] | Zekri et al. [69] | RLSMP [73] | Cuckov and Song [75] | Chen [77] | Shafiee and Leung [79] | Catch-up [80] | Picconi et al. [61] | Secure CASCADE [25] | Raya et al. [71] | Dietzel et al. [40] | Molina-Gil et al. [83] | Garofalakis et al. [84] | Hsiao et al. [85] | Han et al. [86] | CARAVAN [91] | SeDyA [87] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Focus | – | Decision | | | | | | | Fusion | | | | Dissemination | | | | | Security | | | | | | | | | |
| | Section | IV | IV-A | | | | | | | IV-B | | | | IV-C | | | | | IV-D | | | | | | | | | |
| | Main use case[1] | T | P | D | T | T | T | T | T | T | P | P | P | T | T | T | T | T | T | T | T | T | T | B | B | T | T | T |
| **Decision** | Flexible segments | ○ | ○ | ○ | ● | ● | ● | ● | ● | ● | ○ | ○ | ○ | ● | ● | ● | ○ | ○ | ● | ● | ○ | ● | ● | ○ | ○ | ○ | ○ | ● |
| | Hierarchical | ○ | ● | ● | ○ | ● | ● | ● | ● | ○ | ● | ● | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ● | ○ | ● | ○ | ○ |
| **Fusion** | Lossless | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | Duplicate filter | ○ | ○ | ○ | ● | ○ | ○ | ○ | ● | ○ | ○ | ● | ● | ○ | ○ | ○ | ○ | ● | ○ | ● | ○ | ○ | ○ | ● | ● | ● | ○ | ○ |
| **Dissemination** | Infrastructure needed | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ○ |
| | Context-adaptive | ● | ● | ● | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ● | ● | ● | ● | ● | ○ | ○ | ● |
| | Node hierarchy[2] | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ● | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ● | ○ | ○ | ○ | ○ | ○ | ● | ● |
| | Abnormality-based | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | Adaptive delay | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| **Security** | Integrity protected | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● |
| | Unique identities | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ● | ● | ● | ● | ● | ● | ○ | ○ |
| | Trusted hardware | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ● | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

[1] T: traffic information, P: parking spots, D: driving times, B: binary events.
[2] A subset of vehicles is selected to perform special actions; an example are cluster heads in a clustering mechanism.

are, however, different in sensor networks. For instance, the energy demands of wireless communication are not an issue in vehicular networking, where, as pointed out above, available network capacity is the key limiting factor. This, in turn, is not a bottleneck in most sensor network applications. Moreover, the underlying communication pattern is also very different: while sensor networks typically transmit data to one (or few) sinks, aggregation mechanisms in VANETs aim to provide aggregated information to all (or most) vehicles. These differences imply the need for very different protocol designs. Finally, sensor networks are mostly static, and therefore able to build and maintain static aggregation structures. Vehicular networks, in contrast, are highly dynamic, their topology changes very rapidly. For all these reasons, the requirements, the application setting, and the objectives for aggregation mechanisms in VANETs are very different from those in sensor networks. Consequently, the solutions that have been developed are also very different, and it is an open challenge to design protocols that work equally well in both areas. For an overview on aggregation mechanisms in sensor networks see, for instance, [92].

The need for aggregation mechanisms has also been pointed out in the context of participatory sensing applications, most notably by Shi *et al.* [93]. There, the term aggregation is used in the sense of aggregation functions in databases: an application (in the most simple setting located on a central server) formulates an aggregate query whenever it is interested, for instance, in the average temperature in some area. Mobile devices of participating users are then used to collect the raw data necessary to answer the aggregate query. The focus of the work in [93] is then on computing the aggregate function in such a way that the querying instance cannot reveal private data of individual users. In accordance with the differing requirements of this application domain, the work does not primarily aim for a reduction of the used communication bandwidth, as it is the central goal of aggregation in VANETs.

## V. DISCUSSION AND OPEN ISSUES

Table II shows an overview of all schemes discussed in Section IV. For each scheme, we state the most important attributes in the categories decision, fusion, dissemination, and security. Looking at different aggregation protocol components, we see common trends but also open challenges, such as integration of aggregation mechanisms with other dissemination means. Moreover, we will discuss how results from aggregation mechanisms for the VANET domain apply to other research domains.

When aggregation was first introduced in the VANET research field, the immediate idea was to combine measurements of fixed road segments using 1-hop broadcast and to only disseminate the aggregated information over multiple hops. Segment size was largely inspired by wireless communication range for practical reasons. Because the communication overhead reduction achieved by such a flat scheme is bounded by a constant factor, it is easy to see that the first schemes do not scale to larger areas. Therefore, the addition of multiple levels of hierarchy was a natural consequence. Especially the reduction of aggregation accuracy with growing distance is a

common approach. The rationale is that vehicles do not need exact information about areas far away to make decisions. As the vehicles approach a certain region, they will receive more and more accurate information about the region.

A different line of research has addressed a second problem of predefined road segmentation. Depending on the current traffic or environment situation, fixed segments can be either too small or too large. In case of a large-scale phenomenon like a large traffic jam, small segments waste bandwidth. In case of small phenomena like a spot of icy road, fixed segments will conceal important details. Flexible decision systems adapt to these changing conditions and aggregate more or less accurately accordingly. However, this flexibility comes at a cost. A scheme with pre-defined road segmentation and fixed hierarchical aggregation can provide dependable guarantees about the expected information horizon available to a vehicle. Flexible schemes on the other hand provide information with higher utility. But due to the changing size of aggregates, it is hard to estimate how much information will be available to a vehicle.

It is an open challenge to combine the flexibility of dynamic road segmentation schemes with the predictability of fixed hierarchies. Following the argument of reduced granularity for information further away, a promising approach could be to use flexible aggregation for information in the direct vicinity and to switch to a fixed hierarchical approach for information further away.

In terms of data fusion, the main observation is the necessity for duplicate-insensitive fusion mechanisms, namely, counting, summing, and averaging. The most-proposed mechanism to achieve duplicate insensitivity are FM sketches or variations thereof. For example, modifications of FM sketches have been proposed to cope with measurements that are only valid for a limited period of time. The cost of using such sketches is twofold. First, sketches approximate the real values within certain error bounds, but they do not provide exact results. Second, the size of a sketch is larger than a simple integer value, assuming common sketch parameters. Moreover, FM sketches are known to have high error bounds for small counts, which can be a problem in some network density settings. Finally, no dependable error bounds have been calculated for using sketches to calculate average values. The simple approach is to keep two sketches, one sum and one count, and divide them by each other to obtain the average value. However, this approach doubles the space requirement, and the resulting error bound for the average value is not well investigated.

An alternative way to achieve duplicate insensitivity without FM sketches is to keep a list of vehicle IDs for all vehicles that participate in a sketch. Vehicles can then determine whether to add their observation by checking if their ID is contained in the attached list of IDs. This approach may seem naïve at first sight, but recall that all existing approaches for integrity protection of aggregates depend on a list of IDs of some sort (e.g., certificates of all participating vehicles) for aggregate verification. If such a secure aggregation mechanism is used, it can be an alternative to use the ID list for an exact count and only rely on FM sketches for duplicate-insensitive sums.

Of course, integrity mechanisms that rely on such participant ID lists pose a problem in terms of scalability. Therefore, it remains an open challenge to optimize duplicate insensitive fusion, especially for more complex operations.

While duplicate-insensitivity has been tackled by existing schemes in various ways, we believe that the potential to actively correct faulty or biased sensor values has not yet been exploited. Given that aggregation protocols combine information from a large number of different sources, it is likely that only a small fraction is faulty or biased. However, it is an open challenge to design information fusion functions that can detect outlying values and biases and actively filter them.

In contrast to decision and fusion mechanisms, the dissemination mechanism of almost all existing schemes is fairly simple. Vehicles typically broadcast their world model, or a subset thereof, in regular, periodic intervals. Subset selection is usually based on relevance criteria; geographic distance to the sending vehicle is the predominant selection method implemented. For scalability, it is crucial that the set of relevance criteria is carefully selected. It remains an open challenge to select proper dissemination regions and corresponding information routing schemes in a city scenario, because routes of vehicles are more complex and dissemination in circular areas might require too much bandwidth. Moreover, dissemination could be improved by implementing more adaptive timings instead of periodic dissemination. For instance, information dissemination could be delayed to increase chances of aggregation and delivery success, as proposed in [80], [81], [94]. In addition to routing criteria, MAC layer protocols could be tailored to aggregation protocols, for instance, by prioritizing more important aggregated information. In addition, dissemination could react dynamically to available bandwidth in the network by adapting aggregation decisions, i.e., aggregating more coarse or more fine-grained depending on available bandwidth.

Especially in recent years, several proposals for secure in-network aggregation have been made. However, existing cryptography-based approaches are limited to aggregation of binary events on fixed road segments. On the other hand, consistency-based approaches mainly rely on trusted computing primitives to achieve security. Both more flexible cryptographic approaches and more reliable consistency-based approaches remain an open challenge. Considering the constraints of current cryptographic approaches, it is well conceivable that consistency checks will play an important role in future secure aggregation schemes.

Beyond challenges of specific aggregation components, we see the support for multiple application data types at the same time as one of the main challenges. It is well accepted in the research community that a number of foreseen VANET applications will require multi-hop message dissemination [7], [51], [52], [95], [96] and that their bandwidth consumption will be an important problem to solve. Because aggregation mechanisms only scale well if they know the semantics of the data they combine, however, current proposals are only able to cope with one type of application data at a time. For instance, they only average speed or count free parking spots. The straightforward options to support multiple applications are infeasible: running multiple aggregation protocols in parallel is detrimental to the bandwidth gained by using aggregation in the first place, and putting several values with different underlying

application quality requirements in one aggregate will likely result in reduced information utility. At the same time, multiple applications could be beneficial for information quality. For instance, bad weather conditions could indicate low average speeds, because drivers more carefully. Yang *et al.* [97] present work that discusses such interrelations, but their work does not consider in-network aggregation.

Orthogonal to this problem, aggregation efficiency could be further improved by shifting towards more application-specific values in the aggregates. The proposal of Lochert *et al.* [18] shows a possible step in this direction: instead of average speeds, the authors aggregate estimated travel times. The underlying observation is that average speed information, which is provided by many proposed aggregation schemes, is often further aggregated in local applications. Ultimately, a local navigation system will use it to make routing decisions. A custom-tailored aggregation scheme could therefore only disseminate routing decision clues instead of average speeds to save bandwidth. However, these very application specific approaches are inhibitive to the goal of supporting multiple applications at the same time. Future mechanisms will have to find a compromise—or an efficient encoding mechanism to support both goals at the same time.

Being a multi-hop information dissemination mechanism, aggregation schemes are rivaled by Geocast dissemination mechanisms, cellular data networks including recent proposals for peer-to-peer network overlays [98], [99], and roadside infrastructure units. While these approaches all come with their unique drawbacks, their combination can benefit the overall application. For instance, cellular network coverage can be problematic in high vehicle densities on highways. On the other hand, in-network aggregation can be problematic for dissemination of information in very large areas (i.e., dozens of kilometers). Current research fails to embrace multi-modal information dissemination, which could combine the advantages of different approaches without suffering from their drawbacks. We see the main strength of aggregation protocols in providing almost realtime information about the extended vicinity of a vehicle. Inside an aggregation protocol, vehicles can determine when an aggregate has reached a stable state. At this point, aggregated information can be further disseminated using, for instance, cellular networks. This way, the backend network benefits from reduced network load due to pre-filtering of information. Likewise, such stable aggregates could be disseminated using efficient Geocast in a highway scenario to benefit from the higher bandwidth efficiency of Geocast.

Integrating different types of networks and dissemination means, such as V2V, cellular, and RSUs, mainly increases complexity in the dissemination component of an aggregation scheme. The dissemination component has to adapt to different dissemination strategies depending on context. The decision and fusion components however, can provide benefits independent of dissemination strategies. For instance, if cellular network coverage would be available and all vehicles utilize it for traffic information collection, the available network capacity will likely still not suffice. Pre-aggregation of information from the vehicles' direct neighbors can help to reduce the amount of information that needs to be communicated using the cellular network. Likewise, pre-aggregated information helps to reduce the communication load of RSUs to backend services. Therefore, we see cellular networks and RSU coverage as a chance to improve information dissemination, but at the same time, we see the necessity to still perform aggregation *decisions* and information *fusion*, and we see it as a challenge to design such hybrid aggregation protocols.

In terms of standardization, aggregation protocols have usually been regarded as stand-alone protocols in existing work. If implemented that way, they would duplicate a large number of information that is already contained in standardized messages such as ETSI's CAMs in Europe. An alternative could be to base an aggregation protocol on existing CAMs and only use additional messages for dissemination of aggregated information. It is an open challenge to better integrate aggregation protocols with existing standards to maximize reuse of disseminated information. Moreover, standardized metrics for judging quality of aggregated information and performance of aggregation protocols could help to choose and combine suitable protocols.

We summarize the main future challenges for in-network aggregation schemes as follows.

- Dependable bandwidth usage profiles and upper bounds for flexible aggregation schemes.
- Duplicate insensitivity with low error bounds and little bandwidth usage, especially for average calculations.
- Dissemination that explicitly supports aggregation (e.g., delay and forward strategies tailored to aggregation).
- Scalable information integrity protection while preserving vehicle anonymity.
- Schemes that support multiple information types (e.g., average speed and road conditions) with possibly differing quality requirements.
- Impact assessment of application dependent data types such as travel times instead of average speed.
- Combination of different dissemination modalities (e.g., vehicle to vehicle, roadside infrastructure, cellular networks) for optimal bandwidth use.
- Integration of protocols with standardization activities and common metrics for aggregation quality.

Many interesting proposals for aggregation protocols exist in literature. However, most existing schemes focus on particular aspects of aggregation, and fail to provide an optimal solution for all components. A combination of (an improved version of) FM sketches with a flexible aggregation scheme in the direct vicinity and an hierarchical overlay for larger areas, possibly considering cellular networks to further improve dissemination range, still poses an interesting research challenge. From an application point of view, the support of different application data with different quality requirements is the main challenge to solve.

## VI. Conclusion

Vehicular networks are currently approaching their initial deployment. As a first set of applications, standardization bodies mainly foresee safety functions using 1-hop messages or simple event notifications, which are forwarded over multiple hops. While often raised as an important issue, in-network

aggregation is not currently being implemented. The research literature offers a number of proposals for suitable aggregation mechanisms with varying degrees of flexibility, scalability, and integrity protection. However, most proposed mechanisms are tailored to a specific use case or aim to solve a specific subproblem of in-network aggregation, as discussed in Section V.

In this survey, we have presented a generic architecture and used it to categorize different aggregation mechanisms and assess their suitability for solving particular challenges. We see it as one of the major future challenges to further investigate generic aggregation protocols, which are able to integrate information from different domains, such as traffic information, weather, and parking spots. Besides integrating different domains, a future mechanism could integrate cellular networks in specific situations to ease long-range dissemination of information that was previously aggregated using car-to-car communication. More generic mechanisms especially need to be able to dynamically select aggregation areas based on current traffic or road status and need to adapt their aggregation approach to changing situations. Further, integrity protection of aggregated information is still an open challenge. Existing proposals usually trade-off flexibility for better protection, which, in turn, also incurs higher bandwidth usage.

Once VANET deployments reach larger and larger scale, it is important that research in aggregation mechanisms has progressed enough so that current, simpler information dissemination protocols can be complemented with more advanced aggregation mechanisms. Besides their immediate application to VANETs, the mechanisms we discussed in this survey can serve to inspire research in other research domains, such as data mining or distributed sensing applications. Because of the challenging nature of VANETs, solutions developed for this domain have the potential to be applicable to a wide range of use cases dealing with large sets of highly dynamic data.

## REFERENCES


[1] M. Emmelmann, B. Bochow, and C. Kellum, *Vehicular Networking: Automotive Applications and Beyond*. Hoboken, NJ, USA: Wiley, 2010, ser. Intelligent Transport Systems.

[2] *Intelligent Transport Systems (ITS); Vehicular Communications; Basic set of Applications; Part 2: Specification of Cooperative Awareness Basic Service*, ETSI EN 302 637-2, 2013.

[3] *Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages*, IEEE Std. 1609.2, 2013.

[4] *DSRC Implementation Guide—A Guide to Users of SAE J2735 Message Sets Over DSRC*, SAE Int., Washington, DC, USA, Feb. 2010, v20.

[5] *Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) Specifications (802.11p),* 2007.

[6] *Intelligent Transport Systems (ITS); Vehicular Communications; Geonetworking; Part 4: Geographical Addressing and Forwarding for Point-to-Point and Point-to-Multipoint Communications*, ETSI TS 102 637-4, 2011.

[7] E. Schoch, F. Kargl, M. Weber, and T. Leinmuller, "Communication patterns in VANETs," *IEEE Commun. Mag.*, vol. 46, no. 11, pp. 119–125, Nov. 2008.

[8] E. Schoch and F. Kargl, "On the efficiency of secure beaconing in VANETs," in *Proc. 3rd ACM Conf. Wireless Netw. Security*, 2010, pp. 111–116.

[9] S. Yousefi, M. Mousavi, and M. Fathy, "Vehicular ad hoc networks (VANETs): Challenges and perspectives," in *Proc. 6th Int. Conf. ITS Telecommun.*, 2006, pp. 761–766.

[10] N. Banerjee, M. D. Corner, D. Towsley, and B. N. Levine, "Relays, base stations, and meshes: Enhancing mobile networks with infrastructure," in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw.*, 2008, pp. 81–91.

[11] K. L. Thng, B.-S. Yeo, and Y. Chew, "Performance study on the effects of cell-breathing in WCDMA," in *Proc. 2nd Int. Symp. Wireless Commun. Syst.*, 2005, pp. 44–49.

[12] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[13] H. Hartenstein and K. Laberteaux, *VANET: Vehicular Applications and Inter-Networking Technologies*. Hoboken, NJ, USA: Wiley, 2010, ser. Intelligent Transport Systems.

[14] *Intelligent Transport Systems (ITS); Communications Architecture*, ETSI EN 302 665, 2010.

[15] S. Dietzel, E. Schoch, B. Bako, and F. Kargl, "A structure-free aggregation framework for vehicular ad hoc networks," in *Proc. 6th Int. WIT*, Hamburg, Germany, 2009, pp. 61–66.

[16] S. Dietzel, B. Bako, E. Schoch, and F. Kargl, "A fuzzy logic based approach for structure-free aggregation in vehicular ad-hoc networks," in *Proc. 6th ACM Int. Workshop VANET*, 2009, pp. 79–88.

[17] M. Caliskan, D. Graupner, and M. Mauve, "Decentralized discovery of free parking places," in *Proc. 3rd ACM Int. Workshop VANET*, 2006, pp. 30–39.

[18] C. Lochert, B. Scheuermann, C. Wewetzer, A. Luebke, and M. Mauve, "Data aggregation and roadside unit placement for a VANET traffic information system," in *Proc. 5th ACM Int. Workshop VANET*, 2008, pp. 58–65.

[19] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "TrafficView: Traffic data dissemination using car-to-car communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 8, no. 3, pp. 6–19, Jul. 2004.

[20] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "TrafficView: A scalable traffic monitoring system," in *Proc. IEEE Int. Conf. Mobile Data Manage.*, 2004, pp. 13–26.

[21] S. Dashtinezhad *et al.*, "TrafficView: A driver assistant device for traffic monitoring based on car-to-car communication," in *Proc. VTC-Spring*, May 2004, vol. 5, pp. 2946–2950.

[22] B. Scheuermann, C. Lochert, J. Rybicki, and M. Mauve, "A fundamental scalability criterion for data aggregation in VANETs," in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw.—Mobicom*, 2009, pp. 285–296.

[23] K. Ibrahim and M. Weigle, "CASCADE: Cluster-based accurate syntactic compression of aggregated data in VANETs," in *Proc. IEEE Globecom Workshops*, Nov. 2008, pp. 1–10.

[24] K. Ibrahim and M. C. Weigle, "Optimizing cascade data aggregation for VANETs," in *Proc. 5th IEEE Int. Conf. Mobile ad hoc Sens. Syst.*, Sep. 2008, pp. 724–729.

[25] K. Ibrahim and M. Weigle, "Towards an optimized and secure cascade for data aggregation in VANETs," in *Proc. 5th ACM Int. Workshop VANET*, 2008, pp. 84–85.

[26] K. Ibrahim, M. Weigle, and G. Yan, "Light-weight laser-aided position verification for cascade," in *Proc. Int. Conf. WAVE*, Dearborn, MI, USA, Dec. 2008, pp. 1–9.

[27] T. Nadeem, P. Shankar, and L. Iftode, "A comparative study of data dissemination models for VANETs," in *Proc. 3rd Annu. Int. Conf. Mobile Ubiquitous Syst.*, 2006, pp. 1–10.

[28] *Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines*, ISO/IEC 10918-1, 1994.

[29] *Generic Coding of Moving Pictures and Associated Audio Information—Part 3: Audio*, ISO/IEC 13818-3, 1998.

[30] L. Wischhof, A. Ebner, H. Rohling, M. Lott, and R. Halfmann, "SOTIS—A self-organizing traffic information system," in *Proc. 57th IEEE Semiannu. VTC-Spring*, 2003, pp. 2442–2446.

[31] C. Lochert, B. Scheuermann, and M. Mauve, "Probabilistic aggregation for data dissemination in VANETs," in *Proc. 4th ACM Int. Workshop VANET*, 2007, pp. 1–8.

[32] C. Feng, R. Zhang, S. Jiang, and Z. Li, "QoI-based data gathering and routing guidance in VANETs," in *Web-Age Information Management*, vol. 7419, Z. Bao, Y. Gao, Y. Gu, L. Guo, Y. Li, J. Lu, Z. Ren, C. Wang, and X. Zhang, Eds. Berlin, Germany: Springer-Verlag, 2012, ser. Lecture Notes in Computer Science, pp. 87–98.

[33] E. Gelenbe and L. Hey, "Quality of information: An empirical approach," in *Proc. 5th IEEE Int. Conf. MASS*, 2008, pp. 730–735.

[34] S. Nath, P. B. Gibbons, S. Seshan, and Z. R. Anderson, "Synopsis diffusion for robust aggregation in sensor networks," in *Proc. 2nd Int. Conf. Embedded Netw. Sensys*, 2004, pp. 250–262.

[35] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[36] R. Kumar and M. Dave, "Knowledge based framework for data aggregation in vehicular ad hoc networks," in *Computational Intelligence and Information Technology*, vol. 250, V. Das and N. Thankachan, Eds. Berlin, Germany: Springer-Verlag, 2011, ser. Communications in Computer and Information Science, pp. 722–727.

[37] R. Kumar and M. Dave, "A framework for handling local broadcast storm using probabilistic data aggregation in VANET," *Wireless Pers. Commun.*, vol. 72, no. 1, pp. 315–341, Sep. 2013.

[38] L. Zhang, D. Gao, W. Zhao, and H.-C. Chao, "A multilevel information fusion approach for road congestion detection in VANETs," *Math. Comput. Modell.*, vol. 58, no. 5/6, pp. 1206–1221, Sep. 2013.

[39] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.

[40] S. Dietzel, E. Schoch, B. Könings, M. Weber, and F. Kargl, "Resilient secure aggregation for vehicular networks," *IEEE Netw.*, vol. 24, no. 1, pp. 26–31, Jan./Feb. 2010.

[41] F. Schaub, Z. Ma, and F. Kargl, "Privacy requirements in vehicular communication systems," in *Proc. IEEE Int. Conf. Privacy, Security, Risk, and Trust (Passat)/Symp. Secure Comput.*, Vancouver, BC, Canada, 2009, pp. 139–145.

[42] S. Dietzel, "Privacy implications of in-network aggregation mechanisms for VANETs (invited paper)," in *Proc. 8th Int. Conf. WONS*, Bardonecchia, Italy, 2011, pp. 91–95.

[43] L. Sweeney, "k-Anonymity: A model for protecting privacy," *Int. J. Uncertainty, Fuzziness Know.-Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.

[44] M. Raya and J.-P. Hubaux, "The security of vehicular ad hoc networks," in *Proc. 3rd ACM Workshop Secur. Sens. Netw.*, 2005, pp. 11–21.

[45] *Intelligent Transport Systems (ITS); Security; Security Header and Certificate Formats*, ETSI TS 103 097, 2013.

[46] F. Kargl *et al.*, "Secure vehicular communication systems: Implementation, performance, and research challenges," *IEEE Commun.*, vol. 46, no. 11, pp. 110–118, Nov. 2008.

[47] J. Douceur, "The Sybil attack," in *Proc. 1st IPTPS*, 2002, pp. 251–260.

[48] M. Raya, P. Papadimitratos, I. Aad, D. Jungels, and J.-P. Hubaux, "Eviction of misbehaving and faulty nodes in vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1557–1568, Oct. 2007.

[49] X. Zhuo, J. Hao, D. Liu, and Y. Dai, "Removal of misbehaving insiders in anonymous VANETs," in *Proc. 12th ACM Int. Conf. MSWiM*, 2009, pp. 106–115.

[50] P. Golle, D. Greene, and J. Staddon, "Detecting and correcting malicious data in VANETs," in *Proc. 1st ACM Workshop VANET*, 2004, pp. 29–37.

[51] S. Dietzel, F. Kargl, G. Heijenk, and F. Schaub, "On the potential of generic modeling for VANET data aggregation protocols," in *Proc. 2nd IEEE VNC*, Jersey City, NJ, USA, 2010, pp. 78–85.

[52] S. Dietzel, F. Kargl, G. Heijenk, and F. Schaub, "Modeling in-network aggregation in VANETs," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 142–148, Nov. 2011.

[53] L. Wischhof, A. Ebner, and H. Rohling, "Information dissemination in self-organizing intervehicle networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 90–101, Mar. 2005.

[54] M. van Eenennaam and G. Heijenk, "Providing over-the-horizon awareness to driver support systems," in *Proc. 4th IEEE Workshop V2VCom*, 2008, pp. 1–7.

[55] L. A. Zadeh, G. J. Klir, and B. Yuang, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*. Singapore: World Scientific, 1996.

[56] S. Zionts and J. Wallenius, "An interactive programming method for solving the multiple criteria problem," *Manage. Sci.*, vol. 22, no. 6, pp. 652–663, Feb. 1976.

[57] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, Sep. 1975.

[58] D. Salomon, *Data Compression: The Complete Reference*. New York, NY, USA: Springer-Verlag, 2007, no. 10.

[59] W. M. Waggener, *Pulse Code Modulation Techniques: With Applications in Communications and Data Recording*. New York, NY, USA: van Nostrand Reinhold, 1995, ser. Electrical Engineering.

[60] R. S. Schwartz *et al.*, "Using V2V communication to create over-the-horizon awareness in multiple-lane highway scenarios," in *Proc. IEEE IV Symp.*, Jun. 2010, pp. 998–1005.

[61] F. Picconi, N. Ravi, M. Gruteser, and L. Iftode, "Probabilistic validation of aggregated data in vehicular ad-hoc networks," in *Proc. 3rd Int. Workshop VANET*, 2006, pp. 76–85.

[62] *ZIP File Format Specification*, PKWARE Inc., Dayton, OH, USA, 2012, APPNOTE.TXT version 6.3.3.

[63] T. Welch, "A technique for high-performance data compression," *Computer*, vol. 17, no. 6, pp. 8–19, Jun. 1984.

[64] C. Cutler, "Differential quantization of communication signals," U.S. Patent 2 605 361, Jul. 29, 1952.

[65] G.-Y. Chang, J.-P. Sheu, and C.-Y. Chung, "Zooming: A zoom-based approach for parking space availability in VANET," in *Proc. IEEE 71st VTC-Spring*, 2010, pp. 1–5.

[66] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, no. 1, pp. 90–93, Jan. 1974.

[67] P. Flajolet and G. N. Martin, "Probabilistic counting algorithms for data base applications," *J. Comput. Syst. Sci.*, vol. 31, no. 2, pp. 182–209, Oct. 1985.

[68] C. Lochert, B. Scheuermann, and M. Mauve, "A probabilistic method for cooperative hierarchical aggregation of data in VANETs," *Ad Hoc Netw.*, vol. 8, no. 5, pp. 518–530, Jul. 2010.

[69] D. Zekri, B. Defude, and T. Delot, "A cooperative scheme to aggregate spatio-temporal events in VANETs," in *Proc. 16th Int. Database Eng. Appl. Symp.*, 2012, pp. 100–109.

[70] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[71] M. Raya, A. Aziz, and J.-P. Hubaux, "Efficient secure aggregation in VANETs," in *Proc. 3rd Int. Workshop VANET*, 2006, pp. 67–75.

[72] S. Dietzel, M. Balanici, and F. Kargl, "Towards data-similarity-based clustering for inter-vehicle communication (short paper)," in *Proc. IEEE VNC*, Dec. 2013, pp. 238–241.

[73] H. Saleet, O. Basir, R. Langar, and R. Boutaba, "Region-based location-service-management protocol for VANETs," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 917–931, Feb. 2010.

[74] B. Xu, A. M. Ouksel, and O. Wolfson, "Opportunistic resource exchange in inter-vehicle ad-hoc networks," in *Proc. Mobile Data Manage.*, 2004, pp. 4–12.

[75] F. Cuckov and M. Song, "Geocast-driven structureless information dissemination scheme for vehicular ad hoc networks," in *Proc. IEEE 5th Int. Conf. NAS*, 2010, pp. 325–332.

[76] S. Eichler, C. Schroth, T. Kosch, and M. Strassberger, "Strategies for context-adaptive message dissemination in vehicular ad hoc networks," in *Proc. 3rd Annu. Int. Conf. Mobile Ubiquitous Syst. Netw. Serv.*, Jul. 2006, pp. 1–9.

[77] W. Chen, "VANETs-based real-time traffic data dissemination," in *Proc. IEEE Int. Conf. WCNIS*, 2010, pp. 468–472.

[78] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Dept. Comput. Sci., Duke Univ., Durham, NC, USA, 2000, Tech. Rep.

[79] K. Shafiee and V. C. M. Leung, "A novel localized data aggregation algorithm for advanced vehicular traffic information systems," in *Proc. IEEE ICC Workshops*, 2009, pp. 1–5.

[80] B. Yu, J. Gong, and C.-Z. Xu, "Catch-Up: A data aggregation scheme for VANETs," in *Proc. 5th ACM Int. Workshop Veh. Inter-Netw.*, 2008, pp. 49–57.

[81] B. Yu, C.-Z. Xu, and M. Guo, "Adaptive forwarding delay control for VANET data aggregation," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 1, pp. 11–18, Jan. 2012.

[82] S. Mohanty and D. Jena, "Secure data aggregation in vehicular-adhoc networks: A survey," *Procedia Technol.*, vol. 6, pp. 922–929, Jan. 2012.

[83] J. M. Molina-Gil, P. Caballero-Gil, C. Hernández-Goya, and C. Caballero-Gil, "Data aggregation for information authentication in VANETs," in *Proc. 6th Int. Conf. IAS*, 2010, pp. 282–287.

[84] M. Garofalakis, J. M. Hellerstein, and P. Maniatis, "Proof sketches: Verifiable in-network aggregation," in *Proc. IEEE 23rd Int. Conf. Data Eng.*, 2007, pp. 996–1005.

[85] H.-C. Hsiao, A. Studer, R. Dubey, E. Shi, and A. Perrig, "Efficient and secure threshold-based event validation for VANETs," in *Proc. ACM Conf. WiSec*, 2011, pp. 163–174.

[86] Q. Han, S. Du, D. Ren, and H. Zhu, "SAS: A secure data aggregation scheme in vehicular sensing networks," in *Proc. IEEE ICC*, Cape Town, Africa, 2010, pp. 1–5.

[87] R. W. van der Heijden, S. Dietzel, and F. Kargl, "SeDyA: Secure dynamic aggregation in VANETs," in *Proc. 6th ACM Conf. WiSec*, 2013, pp. 131–142.

[88] A. Boldyreva, "Threshold signatures, multisignatures and blind signatures based on the Gap-Diffie-Hellman-group signature scheme," in *Proc. Public Key Cryptogr.*, 2003, vol. 2567, pp. 31–46.

[89] M. L. Das, "A key escrow-free identity-based signature scheme without using secure channel," *Cryptologia*, vol. 35, no. 1, pp. 58–72, Jan. 2011.

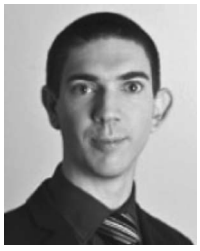
[90] S. Dietzel, J. Petit, G. Heijenk, and F. Kargl, "Graph-based metrics for insider attack detection in VANET multi-hop data dissemination protocols," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1505–1518, May 2013.

[91] K. Sampigethaya *et al.*, "CARAVAN: Providing location privacy for VANET," in *Proc. 3rd ESCAR Conf.*, 2005, pp. 1–15.

[92] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 14, no. 2, pp. 70–87, Apr. 2007.

[93] J. Shi, R. Zhang, Y. Liu, and Y. Zhang, "PriSense: Privacy-preserving data aggregation in people-centric urban sensing systems," in *Proc. 29th Conf. INFOCOM*, 2010, pp. 1–9.
[94] N. Kumar, N. Chilamkurti, and J. J. P. C. Rodrigues, "Learning automata-based opportunistic data aggregation and forwarding scheme for alert generation in vehicular ad hoc networks," *Comput. Commun.*, vol. 39, pp. 22–32, 2013.
[95] F. Bai and B. Krishnamachari, "Exploiting the wisdom of the crowd: Localized, distributed information-centric VANETs," *IEEE Commun. Mag.*, vol. 48, no. 5, pp. 138–146, May 2010.
[96] ETSI, Intelligent Transport Systems (ITS); Vehicular Communications; Basic set of Applications 2010.
[97] J.-Y. Yang, L.-D. Chou, C.-F. Tung, S.-M. Huang, and T.-W. Wang, "Average-speed forecast and adjustment via VANETs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4318–4327, Nov. 2013.
[98] J. Rybicki *et al.*, "Challenge: Peers on wheels—A road to new traffic information systems," in *Proc. 13th Annu. Int. Conf. MobiCom*, Sep. 2007, pp. 215–221.
[99] J. Rybicki, B. Scheuermann, and M. Mauve, "Peer-to-peer data structures for cooperative traffic information systems," *Pervasive Mobile Comput.*, vol. 8, no. 2, pp. 194–209, Apr. 2012.

**Stefan Dietzel** received the Diploma degree in computer science in 2008 from the University of Ulm, Ulm, Germany, where he is currently working toward the Ph.D. degree. Between 2010 and 2012, he was with the Distributed and Embedded Systems Security Group, University of Twente, Enschede, The Netherlands. Since 2012, he has been with the Institute of Distributed Systems, University of Ulm. His research interests include message dissemination mechanisms, in general, and in-network data aggregation, in particular, as well as security and privacy aspects of vehicular communication.



**Jonathan Petit** is a Postdoctoral Fellow with the Services, Cybersecurity and Safety Group, University of Twente, Enschede, The Netherlands. He received the Ph.D. degree in networks, systems and architecture from the University of Toulouse, Toulouse, France, in 2011. He is a Technical Coordinator of the European FP7 PRESERVE project. His research interests include security and privacy, intelligent transportation systems, and wireless and vehicular communication.

**Frank Kargl** serves as the Chair of Distributed Systems at the University of Ulm, Ulm, Germany. He is also a Part-Time Professor with the University of Twente, Enschede, The Netherlands. His research interests include dynamic and cooperative distributed systems and their security and privacy, with a special focus on cooperative intelligent transportation systems. In this area, he participated in a number of projects, such as SeVeCom and PRECIOSA. He is currently a Coordinator of the ongoing PRESERVE project that aims to make security and privacy in V2X a reality for the upcoming ITS deployment. He coauthored more than 100 peer-reviewed publications, and he is actively contributing to the cooperative ITS community through participation in bodies, such as the C2C-CC, and as a Cochair of events such as ACM WiSec, ACM VANET, IEEE WiVeC, or IEEE VNC. Other areas of his research include hardware and system security, privacy, and distributed computing.

**Björn Scheuermann** is a Professor and Chair of Computer Engineering with the Humboldt-Universität zu Berlin, Berlin, Germany. He received the B.S. degree in mathematics and computer science and the Diploma degree (German M.S. equivalent) in computer science from the University of Mannheim, Mannheim, Germany, both in 2004. In 2007, he received the Ph.D. degree in computer science from the Heinrich Heine University, Dusseldorf, Germany, where he became a Junior Professor in 2008. He was an Associate Professor with the University of Würzburg, Würzburg, Germany, where he was the Head of the Telematics Group. He was also an Associate Professor in practical computer science/IT security with the University of Bonn, Bonn, Germany. In October 2012, he joined the Humboldt-Universität zu Berlin. His scientific focus is on performance, design, and security aspects of computer networks. Within this field, he works, for instance, on car-to-car communication, the performance of privacy-preserving communication, and network hardware design.