Tobias Meuser, Daniel Burgstahler, Tobias Rückelt, Doreen Böhnstedt and Ralf Steinmetz. *Hybrid-ProbSense.KOM: Probabilistic Sensing with Hybrid Communication for Gathering Vehicular Sensed Data*. In Proc. Automotive meets Electronics (AmE), IEEE, 2017. ISBN: 978-3-8007-4369-8

Hybrid-ProbSense.KOM: Probabilistic Sensing with Hybrid Communication for Gathering Vehicular Sensed Data

Tobias Meuser, Daniel Burgstahler, Tobias Rückelt, Doreen Böhnstedt and Ralf Steinmetz
Multimedia Communications Lab (KOM), Technische Universität Darmstadt, Darmstadt, Germany
Email: {firstName.lastName}@KOM.tu-darmstadt.de

***Abstract*—Knowledge about vehicles' surroundings gained by the use of local sensors is getting increasing importance. It provides the basis for many advanced driver assistance systems, that are constantly enhanced. However, the functionality of such systems is limited with the sensing range. In order to improve vehicular perception, communication can provide the means. Thereby, a long perception range can be provided by a backend that gathers vehicular sensed information via cellular communication. Such an approach also copes with sparse traffic situations. But since cellular communication is costly, transmission should be minimized. A centrally controlled probabilistic transmission mechanism, as introduced in [1], is able to reduce the amount of data traffic and simultaneously stick to certain quality metrics.**

Based on this we developed an extended approach, named Hybrid-ProbSense.KOM, that is based on cellular and inter-vehicle communication. By the additional use of inter-vehicle communication, we are able to show a further reduction in the overall cellular data traffic of up to 20%.

***Index Terms*—probabilistic sensing, vehicular data collection**

I. Introduction

Many Advanced Driver Assistance Systems (ADAS) require large amounts of sensor data to evaluate the context of the ego-vehicle. As sensor range is physically limited and heavily dependent on external factors, respective assistance systems can only percept the direct environment of the vehicle. To increase the perception range, sensed information can be exchanged via wireless communication. In literature two different approaches have been proposed, namely decentralized approaches, based on wireless ad-hoc communication and often named *DSRC* or *V2V* communication, and centralized approaches based on cellular networks.

Advantages of the decentralized approaches are short latency and independence from external infrastructure. However, requirements on the vehicles themselves are potentially higher than in a centralized approach. V2V communication technology has not been widely deployed yet and may fail in sparse traffic situations. Furthermore, a persistent storage is hard to realize as most network participants are highly mobile. Without infrastructure support, transmission range is highly limited and can only be increased by the use of a multi-hop approach.

In contrast to that, a centralized approach relies on external infrastructure, i. e., cellular networks, and can distribute data easily without consideration of range. Moreover data can be stored persistently on a centralized backend and easily distributed from there. But the use of cellular networks is costly. Thus, data traffic sent via these networks should be minimized. In this paper, we focus on a hybrid approach, combining the advantages of centralized and decentralized approaches.

For the centralized approach, the amount of sent data needs to be limited. Classic traffic load optimization approaches use deterministic behavior and can be classified into three types. *Local preprocessing* reduces the amount of sent data by compression and aggregation. *Data selection schemes* lower the amount of data sent by filtering unnecessary data. In contrast to that, *clustering* based optimization approaches distribute data locally in order to relief the cellular network.

The proposed model is a complete transmission model and based on ProbSense.KOM [1]. As a vehicular network has huge amounts of mobile sensor nodes, i. e., vehicles, most information is potentially sensed multiple times. ProbSense.KOM is purely cellular network based and reduces the amount of sent data by reducing the redundant transmission of the same information using a probabilistic transmission model. As extension, within this work, we consider additional advantages of inter-vehicle communication. We propose Hybrid-ProbSense.KOM, a probabilistic data collection model utilizing both, V2V and cellular communication technology.

The concept is to share sensed information locally with vehicles in the direct surrounding. Once this information is received by a neighbor vehicle, this vehicle stores that information. In case a vehicle senses an event and decides to transmit, all respective received data records are appended as list of witnesses. At the backend this can be interpreted as a set of individual event detections. As a result, in case of optimizing data transmissions against a minimal detection rate, the respective transmission probability can be reduced. We evaluated our hybrid approach Hybrid-ProbSense.KOM, by the use of Simulation of Urban MObility (SUMO) [2] with the open-source TAPAS-Cologne scenario [3], and can show that it outperforms a purely cellular based approach.

The remainder of this paper is structured as follows. After presenting relevant related work in Section II, we give a system overview about our hybrid approach in Section III. Following this we present our evaluation in Section IV, beginning with a description of our simulation setup. Finally we conclude our paper with a summary in Section V.

II. RELATED WORK

In this Section we provide an overview of existing approaches in literature for the related fields of mobile sensing and information dissemination in vehicular networks. Our considered scenario can be classified into the field of *Mobile Sensing* with elements of *Urban Sensing* and *Mobile Crowd Sensing (MCS)*. In *MCS*, huge amounts of distributed mobile devices are used to build large scale sensing applications [4]. Such sensing applications can be classified according to three criteria: the scale of the sensing application, the degree of user interaction and the processing of sensor information. With Hybrid-ProbSense.KOM, our goal is to use a probabilistic sensing approach. A probabilistic approach has the potential to reduce data traffic by sacrificing deterministic results. However, current probabilistic approaches in literature only try to increase data quality by prioritizing correct values over bad ones. Hossain et al. used a probabilistic approach to increase the quality of information by validating the information correctness using different environmental parameters like distance between sensor and sensed event [5]. In our work we use probabilities vice versa to discard a certain percentage of redundancy.

To exchange information with other vehicles, both cellular networks and V2V communication techniques are available. The main purpose of V2V communication is the exchange of information in traffic safety scenarios, i. e., collision warning systems [6]. Current issues are the lack of deployment and the limited range. Particularly the limited range is an issue in sparse traffic situations.

In contrast to this, cellular networks offer a high communication range and dense network coverage. However, due to their transmission costs they should be used advisedly. In order to lower the amount of information transmitted via cellular networks, different techniques have been proposed in literature, namely *clustering*, *data selection schemes* and *local preprocessing*.

Clustering algorithms are mainly based on three basic clustering algorithms: *Lowest ID clustering algorithm (LID)*, *Highest Degree clustering algorithm (HD)* and *Weighted clustering algorithm (WCA)*. In *LID* every node is assigned an unique id, which is broadcasted via Ad-Hoc. As cluster head the node with the lowest ID is selected [7]. *HD* is using the node with the highest node degree as cluster head, which minimizes the amount of clusters [8]. *WCA* is a generalization of HD, which uses a weighting function to select the optimal cluster head. This weighting function has several properties like node degree, transmission power, mobility and battery power [9].

Lowest Relative Mobility Clustering Algorithm (MOBIC) is an algorithm to efficiently cluster in mobile networks [10]. The concept of *MOBIC* is based on *LIT*. Instead of an unique ID, a performance indicator is used to form the cluster. This indicator is based on the relative speed to the neighbors and the node mobility. Different approaches used MOBIC to evaluate the performance of their clustering algorithm [11], [12]. Clustering algorithm face issues in VANETs currently, as V2V technology has not been widely deployed. In order to be able to use clustering algorithm properly, a close to 100% V2V deployment rate is required.

Data selection schemes filter and aggregate information in order to lower the network load. One of the most basic data selection schemes is the *send-on-delta* approach. The concept is to only transmit a sensor value if it differs from the previously transmitted value above a threshold [13]. This approach has been enhanced for sensor networks. Suh introduced the prediction based send-on-delta approach [14]. A linear prediction method has been used to further optimize the send-on-delta approach. The value is only transmitted if the delta between the actual and its predicted value excesses a certain threshold. Unfortunately all of those approaches have been optimized for stationary sensor networks. Due to the rapid changes of sensor location in VANETs, this strategy cannot be deployed easily.

In contrast to that, CarTel is an approach developed for vehicular networks. Nodes are not allowed to transmit data themselves, but data is actively requested by the server [15]. The communication with the server is performed via wireless access technology. Mobile devices are able to request data from the server. If the data is stored in the server cache, the request is immediately responded. In this case, the server sends a sensing request to vehicles eligible for sensing. For data delivery, CafNet's data delivery mechanism is used [16]. This approach is useful if data is requested by external devices. If the vehicles themselves request data from this server, unnecessary overhead would be produced.

Our approach achieves a traffic reduction by lowering the data redundancy on server side. To achieve that, we discard valid information with a certain probability unregarded the quality of this information. This approach can only be used in large scale sensing scenarios. If the sensing scenario is too small, fluctuation and already small redundancies are an issue. Compared to ProbSense.KOM, Hybrid-ProbSense.KOM can deal with smaller scenarios by exploiting V2V communication technology.

III. SYSTEM OVERVIEW: HYBRID APPROACH

Our goal in this work is to collect road related information on a centralized server as efficient as possible. Efficient in this context means with as few data traffic as possible while sticking to predefined quality gates. Each with cellular technology deployed vehicle is a sensor node, building together a huge sensor network.

Our concept is based on the assumption that each event's quality can be measured. This quality is used in our approach to reduce the data traffic by defining a desired quality.

In large scale sensing systems, information is often sensed multiple times. Due to a lack of coordination between vehicles, information is therefore also transmitted multiple times. The quality gate defines how many event transmissions are required to guarantee sufficient quality. In highly trafficked regions, events might be transmitted more often than required. If the quality is too high, the produced traffic is higher than the optimum. In this case the number of transmissions of this event type can be reduced to reduce network traffic. Same goes for the case, if the quality is too low. In this case, the number of event transmission needs to be increased to ensure data quality.

To realize such a behavior without addressing each vehicle independently, we use a probabilistic transmission approach. All transmission probabilities are controlled on server side and distributed from there. Once the server detects a gap between the desired quality and the actual quality, an adjustment is performed.

As this approach requires homogeneously distributed traffic, the map is divided into virtual geo cells. The geo cells are independent from each other and each geo cell is assigned its own transmission probability matrix.

For our approach we do the following assumptions:

1) Every vehicle has a fully working event processing system deployed. Therefore not only sensor data might be detected, but also complex events, i. e., traffic signs.
2) Each vehicle has a cellular network connection deployed and the central server is always available.
3) There are no connection issues through the cellular network. Each packet transmitted by a vehicle will arrive at the server.
4) Some vehicles have also V2V communication technology available. If a vehicle broadcasts a message locally, each vehicle inside the communication range deployed with V2V communication technology receives this message.

A. Difference to ProbSense.KOM

Our hybrid data collection approach uses both cellular networks and V2V communication technology. The assumption is that the deployment rate of V2V technology is increasing over the next years. Therefore, our presented approach extends ProbSense.KOM [1] by the additional usage of V2V communication. The changes on ProbSense.KOM are performed mainly on client side. Thus, the server side does not need any knowledge about the current technology equipment of each vehicle. Moreover, Hybrid-ProbSense.KOM also copes with partial penetration of V2V equipped vehicles.

B. Events types

We distinguish between two event categories in reference to [17]. *Discrete events* are location based and can be detected multiple times in a sampling period, which are considered as different event instances, i. e., the detection of a traffic sign. In contrast to this, *continuous events* are detected for a certain period and are not considered as multiple instances, i. e., temperature detection. For *discrete events* we use the *detection latency* as quality metric. The detection latency is defined as the time between the occurrence of an event and the detection on server side. Latency is no valid quality metric for *continuous events* though, as the occurrence of a continuous event cannot be detected. Therefore we use a defined *event density* as quality metric, which describes the number of events in a certain time and area.

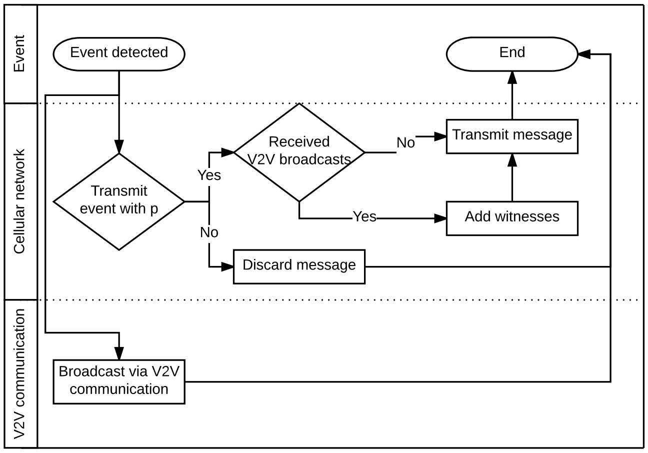


Figure 1: Client transmission process

C. Client side

The main improvement of the proposed model is located at the client side. In ProbSense.KOM, each vehicle measures and transmits its values independently from each other. Hybrid-ProbSense.KOM addresses this issue: Once a vehicle detects an event, the event is broadcasted to all surrounding vehicles. This is done regardless if this event shall be transmitted to the server side or not. The transmission process of the client is pictured in figure 1.

For discrete events this mechanism is especially useful due to their specific location. For continuous events the effects should be rather small, as two transmitting vehicles are more seldom in V2V communication range. Once a vehicle receives a broadcast message from another vehicle, it can append this information to its own event message. As such a mechanism was not provided by HERE [18], we extend the data model by a field 'additional witnesses'. It is then evaluated by the server in order to increase the amount of received information.

D. Server side

The server side behavior is quite similar to the one proposed in ProbSense.KOM [1] and can be divided into two steps: the measurement period and the probability adjustment.

1) Measurement Period: In the measurement period the server remains passive and listens for events transmitted by vehicles. The length of the measurement period is dependent on the event type. For continuous events, the measurement period length can be rather short as data quality can be measured even after short time intervals. The period length for singular events needs to be longer, as all events need to arrive at the server at least two times. Otherwise the latency could not be calculated. If the measurement period length is set too short for singular events, it leads to incorrect results, as the short latency events have already arrived, but the ones with long latencies are not known to the server. The length of the measurement period for singular events is based on the maximum latency, the required redundancy on server side and an factor for the maximum accepted traffic inhomogeneity.

In ProbSense.KOM the server used the number of events and the time of arrival to calculate the probabilities. Compared to ProbSense.KOM, the server in our approach does not only count the number of incoming events, but also evaluates the field 'additional witnesses' (compare III-C). If any additional witnesses are present, than this event is not only counted once but multiple times. This has an impact on the calculated latency as well as the density:

The latency is decreased, as an event arrives at the server not only a single time but multiple times. The server is now able to decrease the transmission probability more than normally, as the required number of events is received faster.

For continuous events the impact of this optimization is rather small for short ranged V2V communication. Anyhow, in some cases vehicles sensing the same continuous event are in V2V communication range. Then the V2V communication has also a positive impact on the density.

2) Probability Adjustment: After the measurement period a probability adjustment is performed. During this step the server calculates the new transmission probabilities based on the measured values of the preceding measurement period.

Every geocell is divided into 4 subcells in order to achieve better results. Instead of calculating the probability for the geocell as a whole, the probability for each subcell is calculated. To ensure data quality, only the highest probability among all subcells is used as transmission probability.

Probability calculation is different for discrete and continuous events:

For *continuous events* the probability calculation is based on the measured density δ_{t-1} of the last measurement period and the expected density δ_{exp}. The formula for the adjusted probability for a continuous event is:

$$p_t = \min\left(p_{t-1} * \frac{\delta_{exp}}{\delta_{t-1}}, 1\right) \tag{1}$$

In contrast to that, the probability calculation for *discrete events* is more complex. The calculation is based on the latency, but the quality gate shall be matched with a probability ρ.

The set of latencies L cannot be measured directly, but it can be estimated by interpolating between the last and the first received event. As the fluctuation of the quantile is too high to ensure reliable results, the average latency $\bar{l}$ is used.

To be able to calculate the probability from the average latency $\bar{l}$, the average number of vehicles $\overline{n_l}$ passing an event of the specific type before latency exceedance is required.

To calculate the average amount, the relation between l and n needs to be determined. This relation can be ascertained using the average measured latency $\bar{l}$ and the average number of vehicles n_a passing that event in time. $\overline{n_a}$ cannot be measured however, but it can be estimated using the average number of received events $\overline{n_r}$ and p as shown in equation 2.

$$\overline{n_a} \approx \overline{n_r} \times p^{-1} \tag{2}$$

As only the average number of events $\overline{n_a}$ would be used for the calculation, events being trafficked less than average would not be considered. That has been addressed by introducing the maximum traffic inhomogeneity α. α is a scaling factor, which can be set accordingly to consider events with less-than-average traffic. The final relation between latency and density is shown in equation 3.

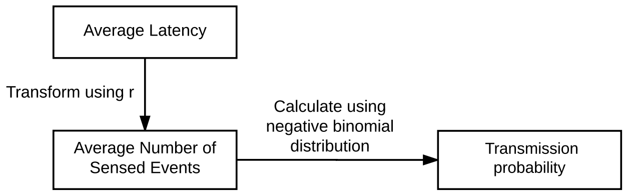


Figure 2: Probability calculation

$$r(l, n) \approx l \times \alpha \times {n_a}^{-1} \tag{3}$$

Now the amount of vehicles $\overline{n_l}$ passing within the l_{exp} can be calculated. In order to determine the new transmission probability using $\overline{n_l}$, we use the negative binomial distribution in equation 4. It describes the nature of discrete events properly, as a certain amount τ has been transmitted to the server after n tries. The corresponding probability density function is defined in equation 5.

$$NB_p(n) = \binom{n-1}{\tau-1} \times p^{\tau} \times (1-p)^{n-\tau}, p \in [0,1] \tag{4}$$

$$\Omega_p(n) = \sum_{i=\tau}^{n} \left[\binom{n-1}{\tau-1} \times p^{\tau} \times (1-p)^{n-\tau}\right], p \in [0,1] \tag{5}$$

The probability can now be calculated by determining the value for the probability p_t, for which the latency is expected to be lower than l_{exp} in ρ cases. Equation 6 displays the calculation of p_t. If p is not defined for some reason, 1 is assumed to ensure functionality.

$$p_t = \min\left[p \;\middle|\; \Omega_p^{-1}(1-\rho) = \overline{n_l}\right] \tag{6}$$

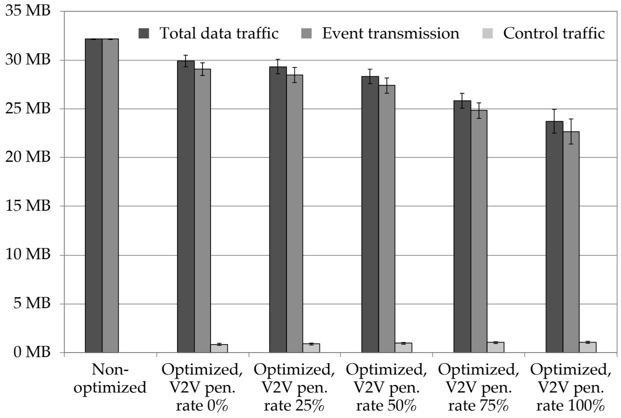
In figure 2 the whole probability calculation process is displayed.

3) Probability Dissemination: The probability dissemination in Hybrid-ProbSense.KOM is the same as in ProbSense.KOM.

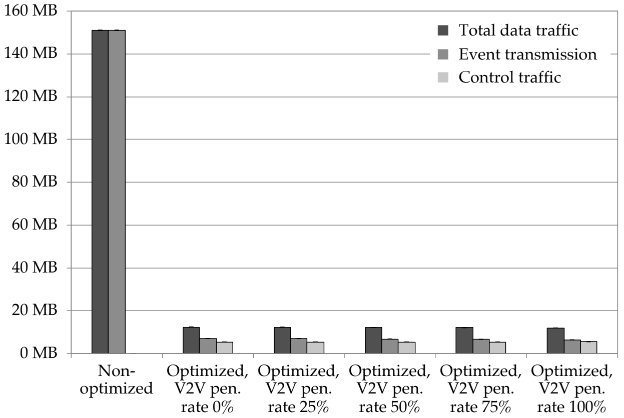
There are two ways for the server to provide the current probabilities to the vehicles:

The first way to transmit the probability is unicast based. Whenever a vehicle transmits an event to the server, the server checks the last location of the vehicle. If the transmitting vehicle has moved to a different geocell, the probability is transmitted in response. Moreover each vehicle without a set of probabilities, i. e., a vehicle that has just been started, automatically request the current probability set once it detects an event.

The second way uses the Geocast technology to disseminate the transmission probabilities. This way is used if the probabilities have been adjusted by the server (compare III-D2).



(a) Traffic for Discrete Events with different V2V deployment rates.



(b) Traffic for Continuous Events with different V2V deployment rates.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

We evaluated Hybrid-ProbSense.KOM by implementing the described model in Java. The server side is fairly similar to the one of ProbSense.KOM, with the exception that the field *additional witnesses* is evaluated. On the client side several changes have been performed. As soon as a vehicle detects an event, it broadcast the event via V2V communication and eventually transmits it to the server. In our simulation, cellular network is assumed to be always available. The broadcasting in V2V has been implemented as fixed-range broadcasting, i. e., all vehicles in a certain radius around the sending vehicle receive the broadcast. The communication range for V2V communication has been set to 150 meters, which is realistic for rural areas. In cities the communication range is typically less, however the used scenario has huge rural areas.

For the movement of vehicles, we used the Simulation of Urban Mobility (SUMO)[1]. The used scenario is the TAPAS Cologne Scenario in Version 24.0, which is one of the largest freely available scenarios. The Cologne Scenario simulates the traffic of Cologne and its rural areas. Its total size is about $30 \times 30km^2$ and the amount of vehicles caps at about 13000 vehicles. As Hybrid-ProbSense.KOM reduces the redundancy on server side, the amount of vehicles on the street has direct impact on the simulation results. With higher numbers of vehicles on the streets the results would further improve.

Two events have been simulated: a speed-sign event (discrete) and a temperature event (continuous). As discrete events are bound to a specific location, we placed 100 of them randomly on the map. This random placement is weighted on the traffic density of the streets to skip the streets with the 5 % lowest traffic density. This is to ensure to prevent event placement on completely untrafficked streets.

The speed-sign event should arrive at the server in time in 95% of all cases, i. e., we tolerate an error of 5%. In all simulations, the produced total traffic with its components *data traffic* and *control traffic* has been investigated. Data traffic is the traffic that is produced by the vehicles to transmit the sensed events to the server. In contrast, control traffic is traffic produced by the server to adjust the probabilities. This includes probability adjustments via unicast as well as adjustments via geocast. It is assumed, that geocast produces the same traffic as multiple unicast connections to all vehicles in that specific region. The results of the simulation have been compared to an opportunistic approach, in which all detected events are always transmitted to the server.

A. Evaluation Results

The figures 3a and 3b show the amount of traffic produced for a certain event type. The non-optimized approach is an complete transmission approach, i. e., every sensed event is transmitted. As described above, the impact on the V2V deployment rate shall be investigated. To show the impact of the penetration rate of V2V technology, we used deployment rates of 0%, 25%, 50%, 75% and 100%. A deployment rate of 0% equals the standard ProbSense.KOM approach.

Figure 3a displays the traffic produced by the different approaches. It can be seen that ProbSense.KOM has only small optimization potential of around 7% in this scenario. This is accounted to the relatively low number of vehicles on the streets. However, the traffic decreases with increasing V2V deployment rate. For 25% V2V deployment rate, the effects are very small, as vehicles are rarely in V2V communication range. From 50% V2V deployment rate on, the improvement increases of a maximum of 26% compared to the non-optimized approach and around 20% compared to ProbSense.KOM. The data traffic is the main factor for the high total traffic, the control traffic in only around 5% of the total traffic produced. In all simulations, the latency stayed within the desired maximum latency of $600s$ in 95% of all cases.

For continuous events, the traffic reduction is very high, as shown in figure 3b. The V2V deployment rate had almost no impact on the produced results however. With a V2V deployment rate of 100%, Hybrid-ProbSense.KOM produces only around 3% less traffic than ProbSense.KOM. This is expected, as vehicles are rarely in V2V communication range when transmitting a continuous event. The total traffic reduction

---

[1] www.sumo.dlr.de

is huge however. Compared to the non-optimized approach, ProbSense.KOM and Hybrid-ProbSense.KOM achieve a traffic reduction of around 92%. The data traffic is again the biggest part of the total traffic, but the control traffic is around 45% of the total traffic. This is due to the more frequent updates of continuous event probabilities, which can be performed as no certain time interval is required in order to measure the data quality correctly like for discrete events.

V. CONCLUSION

In this work, we introduced Hybrid-ProbSense.KOM, a hybrid probabilistic data collection approach. The vehicles on the streets are mobile sensors and collect road related information. This information can be categorized into two data types: discrete events and continuous events. While discrete events occur spontaneously at specific locations (i. e., speed signs), continuous events have no specific location but are a region based measurement (i. e., temperature). Our goal was to further optimize the performance of ProbSense.KOM [1] using single-hop V2V communication. Different to most current approaches in literature, the usage of V2V is optional in our model. As a 100% deployment rate of V2V communication cannot be assumed, Hybrid-ProbSense.KOM is designed to benefit even from less than 100% V2V deployment rate.

Once a vehicle detects an event, the event broadcasted via V2V communication. This enables cooperation between vehicles. Instead of only transmitting the detected event, each vehicle also adds the vehicles that have detected the same event. Therefore the redundancy on server side is higher than in ProbSense.KOM, which resulted in higher optimization potential. The improvement of our model is mainly important for discrete events, as vehicles are often in communication range when detecting the same event. For continuous events however, vehicles are very seldom in communication range and therefore no further optimization could be achieved.

In the evaluation was shown, that Hybrid-ProbSense.KOM is able to lower data traffic of an opportunistic transmission model in all scenarios. For continuous events, Hybrid-ProbSense.KOM achieved an data traffic reduction of around 88% over the opportunistic transmission model. The performance for continuous events was as expected not dependent on the V2V deployment rate. In case of singular events, HybridProbSense.KOM achieves a traffic reduction of 25% over the opportunistic transmission model. Comparing ProbSense.KOM with Hybrid-ProbSense.KOM, the latter was able to further decrease the transmitted data amount by around 19% with a 100% V2V deployment rate. But first optimization could already be achieved with 50% V2V deployment rate.

Further optimization potential can be accomplished by lowering the quality gates or increasing the amount of vehicles. In our future work, we want to focus on further optimization using V2V communication technology by introducing client side probability adjustments and multi-hop V2V communication. Even without 100% deployment rate, we are confident about further optimization potential.

ACKNOWLEDGMENTS

The work presented in this paper was partly funded by the German Federal Ministry of Education and Research (BMBF) under grant no. 01IS12054.

REFERENCES

[1] D. Burgstahler, T. Meuser, U. Lampe, D. Böhnstedt, and R. Steinmetz, "ProbSense.KOM: A Probabilistic Sensing Approach for Gathering Vehicular Sensed Data," in *Proceedings of the 3rd International Conference on Mobile Services (MS 2016)*. IEEE, Jun 2016, pp. 9–16.

[2] D. Krajzewicz, G. Hertkorn, C. Rössel, and P. Wagner, "Sumo (simulation of urban mobility)-an open-source traffic simulation," in *Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM20002)*, 2002, pp. 183–187.

[3] S. Uppoor and M. Fiore, "Large-scale urban vehicular mobility for networking research," in *Vehicular Networking Conference (VNC), 2011 IEEE*. IEEE, 2011, pp. 62–69.

[4] H. Ma, D. Zhao, and P. Yuan, "Opportunities in Mobile Crowd Sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.

[5] A. Hossain, P. Biswas, and S. Chakrabarti, "Sensing Models and its Impact on Network Coverage in Wireless Sensor Network," in *Proc. Third International Conference on Industrial and Information Systems*, 2008.

[6] X. Yang, J. Liu, N. H. Vaidya, and F. Zhao, "A Vehicle-to-Vehicle Communication Protocol for Cooperative Collision Warning," in *Proc. First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, 2004.

[7] C. R. Lin and M. Gerla, "Adaptive Clustering for Mobile Wireless Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 15, no. 7, pp. 1265–1275, 1997.

[8] M. Gerla and J. T.-C. Tsai, "Multicluster, Mobile, Multimedia Radio Network," *Wireless Networks*, vol. 1, no. 3, pp. 255–265, 1995.

[9] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: A Weighted Clustering Algorithm for Mobile Ad Hoc Networks," *Cluster Computing*, vol. 5, no. 2, pp. 193–204, 2002.

[10] P. Basu, N. Khan, and T. D. Little, "A Mobility Based Metric for Clustering in Mobile Ad Hoc Networks," in *Proc. 2001 International Conference on Distributed Computing Systems Workshop*, 2001.

[11] Y. Zhang, J. M. Ng, and C. P. Low, "A Distributed Group Mobility Adaptive Clustering Algorithm for Mobile Ad Hoc Networks," *Computer Communications*, vol. 32, no. 1, pp. 189–202, 2009.

[12] M. M. C. Morales, C. S. Hong, and Y.-C. Bang, "An Adaptable Mobility-Aware Clustering Algorithm in Vehicular Networks," in *Proc. 2011 13th Asia-Pacific Network Operations and Management Symposium*, 2011.

[13] M. Miskowicz, "Send-on-Delta Concept: An Event-Based Data Reporting Strategy," *Sensors*, vol. 6, no. 1, pp. 49–63, 2006.

[14] Y. S. Suh, "Send-on-Delta Sensor Data Transmission With a Linear Predictor," *Sensors*, vol. 7, no. 4, pp. 537–547, 2007.

[15] B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden, "CarTel: A Distributed Mobile Sensor Computing System," in *Proc. 4th International Conference on Embedded Networked Sensor Systems*, 2006.

[16] K. W. Chen, "CafNet: A Carry-and-Forward Delay-Tolerant Network," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

[17] S. Ahn and D. Kim, "Proactive context-aware sensor networks," in *European Workshop on Wireless Sensor Networks*. Springer, 2006, pp. 38–53.

[18] HERE, "Vehicle Sensor Data Cloud Ingestion Interface Specification v2.0.2," 2015. [Online]. Available: https://lts.cms.here.com/static-cloud-content/Company_Site/2015_06/Vehicle_Sensor_Data_Cloud_Ingestion_Interface_Specification.pdf

All online references in this paper were last accessed and validated in January 2017.