Comparison of Relevance Estimation Mechanisms for Cooperative Awareness Messages in VANETs

Jakob Breu
Research and Development
Mercedes-Benz
Sindelfingen, Germany
Email: jakob.breu@daimler.com

Michael Menth
Department of Computer Science
University of Tübingen
Tübingen, Germany
Email: menth@uni-tuebingen.de

***Abstract*—Cooperative Awareness Messages (CAMs) regularly distribute state information about vehicles in Vehicular Ad-hoc Networks (VANETs). Hardware modules in series vehicles have limited processing capabilities so that only a subset of received CAMs can be processed which may be chosen by a relevance estimation mechanism (REM). In this work, we review several REMs that differ in complexity. We compare them using various approaches in different road scenarios. The results show that more complex REMs provide similar results in simple road scenarios but clearly improved results for demanding road scenarios.**

I. Introduction

In the oncoming years *Vehicular Ad-Hoc Networks* (VANETs) are expected to be one of the innovations with the biggest impact towards improved driving safety, more economic driving and smoother traffic [1]. The basis of this system is a derivation of the 802.11 family of standards called 802.11p [2]. It has been optimized for a connectionless data transmission to enable minimum latencies in a dynamic environment. The Cooperative Awareness Message (CAM) is one of several standardized message formats on the application layer [3], [4]. CAMs are transmitted frequently up to 10 times per second by all participating vehicles and contain basic information about the sender's current state. Each vehicle can maintain an internal map of its environment and support various applications by receiving CAMs from surrounding vehicles.

After decades of research and standardization, VANETs have evolved to a state in which the deployment into series vehicles is imminent [5], [6]. The car manufacturers face new challenges in the development for the mass market. One of these challenges is the processing of all received messages. In a recent study we have shown that depending on the local vehicle density, movement patterns, penetration rate of VANET technology and road topology, high rates of messages may be received [7], [8]. Increased message rates can induce a high processing load on the application layer. For example, each message may be stored in an internal database, fused with other sensor data, and map-matched for precise positioning to realize sophisticated applications. To fulfill industrial requirements regarding cost, robustness and stability, the overall system has to be designed to deal with these message rates. To that end, we proposed to filter the most relevant messages and introduced REMs of different complexity in [9], [10].

Additionally, REMs have been proposed to enable different types of intelligent broadcasting algorithms [11], [12] and in-network data aggregation concepts [13]. The basic idea is to handle highly relevant information in a preferred way while information of low relevance is discarded or merged into aggregates.

As REMs differ in complexity but series implementations should be efficient, the benefit of more complex REMs must be quantified to decide whether a simple or more complex REM should be chosen in practice. To that end, we develop an *observed relevance* of a vehicle for a receiver, which is based on its actual trace over time, and use it as ground truth to compare the quality of REMs. We further suggest several approaches to visualize and quantify the outcome of this comparison. The message orders induced by different REMs can be diagrammed to identify the effects of each mechanism. The rank correlation measures differences of the message orders in a scalar value. In addition we quantify the fraction of erroneously filtered messages. These methods are applied to simulated traffic data and the results are discussed.

The remainder of this paper is organized as follows. Section II gives an overview on proposed REMs. In Section III the simulation environment is introduced. In Section IV, the comparative methods are defined and the results for the simulated traffic are discussed. Section V concludes and gives prospects on future work.

II. Relevance Estimation Mechanisms

The determination of the relevance of a message depends on available information, processing capabilities, and the definition of relevance. A message's relevance regarding crash avoidance applications differs from the relevance regarding parking space allocation. A number of REMs have been proposed for message routing and forwarding in future VANETs. Most of them use spatio-temporal approaches, combining location and time information with weight factors to construct an overall relevance function [12], [14], [11].

In the following sections we review a set of REMs. The used parameter values are given in Table I.

©IEEE, 82nd IEEE Vehicular Technology Conference (VTC Fall), Boston, September 2015

Parameter	Value	Parameter	Value
d_{min}	10 m	α	0.015
γ	0.3821	Δd_{max}	1000 m
t_{max}	10 s	β	0.15
		Δt_{max}	10 s

TABLE I
PARAMETERS USED FOR THE REMS.

A. Distance-based Relevance Estimation

A commonly proposed relevance mechanism is the Euclidean distance of sending and receiving vehicle at the time of the message transmission as a relevance value [15]. This approach does not take the movement of both vehicles into account. We call it *distanceRE*:

$$R_{\text{distance}} = \frac{1}{\max(d_{\text{min}}, d(t_0))}, \tag{1}$$

where d_{min} is the direct vicinity of the receiving vehicle and $d(t_0)$ is the Euclidean distance of sending and receiving vehicle at the message transmission time t_0.

B. Movement-based Relevance Estimation

For crash avoidance applications the relevance of a CAM corresponds to the current and future distance of the sending and receiving vehicles. Our proposal for a general relevance estimation formula is

$$R_{\text{general}} = \max_{t_0 \leq t \leq t_{\text{max}}} \left(\frac{1}{\max(d_{\text{min}}, d(t))} \cdot \frac{1}{(\frac{1}{\text{s}} + t)^{\gamma}} \right), \tag{2}$$

where $d(t)$ is the distance between sending and receiving vehicle at time t (unit s) and γ is a time penalty coefficient [9]. To limit the prediction time, t_{max} should be set to a reasonable time range. The distance function $d(t)$ predicts the future distance of sending and receiving vehicle based on their speed and heading information at time t_0.

1) Static Movement Extrapolation: One assumption for the distance prediction is that both vehicles move with constant speed and heading given at t_0 [9]. We call this approach *staticRE* in the remainder of this work as it assumes static movement.

2) Dynamic Movement Extrapolation: The assumption of static movement ignores the fact that vehicles often change their speed and heading. Therefore, this assumption may lead to relevance estimation values which differ significantly from the corresponding observed relevance values. In such a situation it might happen that messages will be discarded even if their real relevance is high enough.

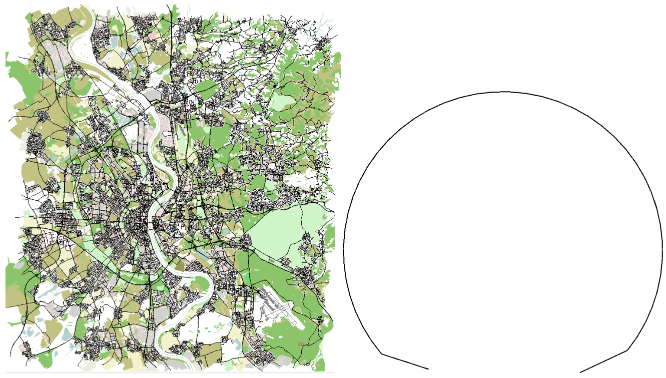
To eliminate this shortcoming, we proposed an extended distance function in [10]. In this concept, the movement is assumed to be dynamic in a way that the relevance is maximized. To this end, we assume the sending vehicle to move towards the receiving vehicle. This approach is called *dynamicRE* in the following.

C. Encounter Probability

The *encounter probability* approach, created for the dissemination of parking space information, also assumes static movement extrapolation to predict the shortest distance Δd between a receiving vehicle and a remote entity [11]. Additionally the time Δt when Δd is reached is taken into account:

$$EP = \frac{1}{\alpha \cdot \min(\Delta d, \Delta d_{\text{max}}) + \beta \cdot \min(\Delta t, \Delta t_{\text{max}}) + 1}. \tag{3}$$

Δd and Δt are combined with weights α and β, and are limited to Δd_{max} and Δt_{max}. We call this approach *encounterRE*.



(a) TAPASCologne scenario. (b) Rural curve scenario.

Fig. 1. Simulated road scenarios.

III. SIMULATION ENVIRONMENT

VANET simulations usually cover the three domains traffic, channel/communication and application simulation [16]. We describe these modules in the following sections.

A. Traffic Simulation

Position and movement of vehicles are fundamental input variables for the analysis of VANETs. We use the traffic simulation tool SUMO due to its efficient implementation [17]. Its configuration requires three steps: 1) select a road map section; 2) define vehicle types; 3) define vehicle sources and sinks.

We configured two different traffic scenarios to compare the REMs in general traffic as well as in an application-dependent scenario. The simulation scenario TAPASCologne defines a whole day's traffic for the German city of Cologne [18]. This scenario allows for the collection of statistics for day-to-day traffic. The second scenario is an example for an application-specific scenario. The underlying road map is a wide curve, representing rural situations for applications like oncoming traffic collision avoidance [19]. Both scenarios are depicted in Figure 1.

B. Channel Simulation

We use a statistical channel model which we derived in [7] to simulate the message transmission. This approach allows for an efficient implementation. It also eliminates a huge set of parameters compared to discrete-event network simulators. To simplify the application of this statistical channel model we assume line-of-sight between senders and receivers.

C. Application Simulation

On the application layer the sending vehicle triggers CAMs according to its current movement pattern [3]. On the receiving side, the REMs are implemented. In addition the collection of data for the statistics is deployed here.

IV. Comparison of Message Orders

In this section we first introduce the observed relevance value as ground truth. Then, we provide methods to illustrate and quantify the differences between message orders induced by the ground truth and the considered REMs. Finally, we discuss our results.

A. Observed Relevance Values as Ground Truth

We define the *observed relevance value* to be the maximum value calculated by Equation (2), using the real distance of sending and receiving vehicle for discrete time steps between t_0 and $t_0 + t_{\text{max}}$ by taking the real movement of those two vehicles into account. This allows us to compare the estimated relevance value calculated at t_0 with the observed relevance value determined in the time interval $[t_0, t_0 + t_{\text{max}}]$.

We assume that the observed relevance values based on the real vehicle traces reflect the real relevance of CAMs in practice. The REMs make assumptions about vehicle traces which causes that their relevance values for a CAMs may differ from corresponding observed relevance values. Therefore, we use the observed relevance values based on Equation (2) as ground truth for comparison. The outcome of the comparison depends on the definition of the ground truth and may be adapted for other applications.

B. Message Order Comparison

To compare the effects of different REMs, we sort all messages by their observed and estimated relevance values. This yields the message orders τ_o and τ_e, respectively. $\tau_o(i)$ and $\tau_e(i)$ are the ranks of the ith message. The message orders are normalized to the range $[0, 1]$.

In Figures 2(a)–2(c) we plotted the message orders for the TAPASCologne scenario. We omit the figures for the curve scenario due to the lack of space. The dashed black 45° line indicates the points where the rank induced by both estimation and observation are the same. Points below that line correspond to messages which were ranked erroneously too high, points above that line correspond to messages which were ranked too low.

Figure 2(a) shows the results for the distance-based relevance estimation. We observe that most messages have a correct ranking while a certain amount of messages are ranked too high (values lie in the upper left part). This is in accordance with how this mechanism works. If two vehicles approach each other after the message transmission, the estimated relevance is too low. If they move apart, the observed relevance equals the estimated relevance.

In Figure 2(b) there is a shift towards too low ranks because of the position prediction. We assume vehicular movement changes to be nearly identically distributed, so the effects of overestimation and underestimation are similar. This is expressed by the graph.

Figure 2(c) shows the results for the dynamic movement approach. We observe that the set of too low ranks decreases to less deviating values. In contrast, the number of overestimated relevance values is higher. That corresponds to this approach's concept which assumes worst-case movement changes towards higher relevance values.

Mechanism	**Kendall's τ rank correlation coefficient**	
	TAPASCologne	Curve
distanceRE	0.966	0.627
staticRE	0.989	0.962
dynamicRE	0.975	0.982
encounterRE	0.893	0.908

TABLE II
MESSAGE RANK CORRELATIONS.

For the encounter probability approach, Figure 2(d) shows that this mechanism yields a broad area of values. Apparently this mechanism tends to underestimate values. Other parameter sets lead to similar effects. This approach's concept obviously does not fit to the chosen ground truth.

The results show that using this graphical representation of message orders the effects of REMs can be analyzed and undesirable phenomena can be detected.

C. Message Rank Correlation

To quantify the similarity between message orders induced by observed and estimated relevance values, we propose to use *Kendall's τ rank correlation coefficient* as metric [20]. Kendall's τ can be determined for tuples $(\tau_e(i), \tau_o(i))$ of each message i by calculating

$$\tau = 2 \cdot \frac{n_{\text{concordant pairs}} - n_{\text{discordant pairs}}}{n \cdot (n-1)},$$

where $n_{\text{concordant pairs}}$ is the number of all concordant pairs, i.e., all pairs of tuples $(\tau_e(r), \tau_o(r))$, $(\tau_e(s), \tau_o(s))$ where the relative order is the same ($\tau_e(r) > \tau_e(s)$, $\tau_o(r) > \tau_o(s)$) and $n_{\text{discordant pairs}}$ is the number of all discordant pairs, i.e., all pairs of tuples where the relative order is different ($\tau_e(r) < \tau_e(s)$, $\tau_o(r) > \tau_o(s)$).

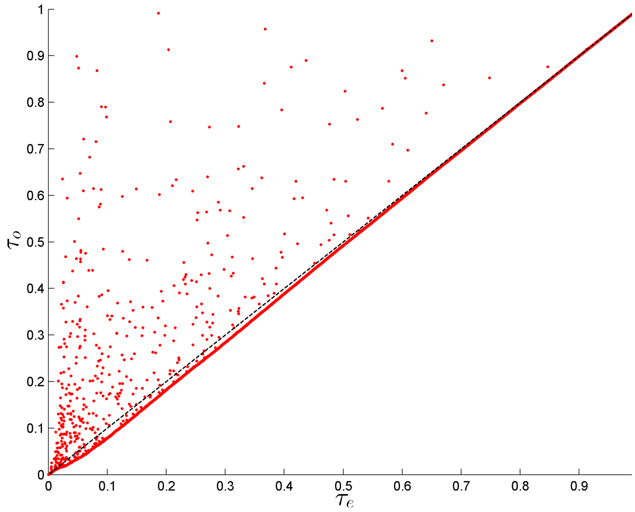
Table II lists the results for the TAPASCologne and curve scenarios. The encounter probability approach yields values of around 0.9. For the TAPASCologne scenario, the distance-based relevance estimation yields a high rank correlation coefficient of 0.966. The static-movement-based approach leads to an even better result near to the optimum value. In contrast, the dynamic relevance estimation results in between.

The curve scenario yields a bad performance for the distance-based approach, while the dynamic movement approach and the static movement approach switch places. Obviously the dynamic movement approach performs better in this scenario.

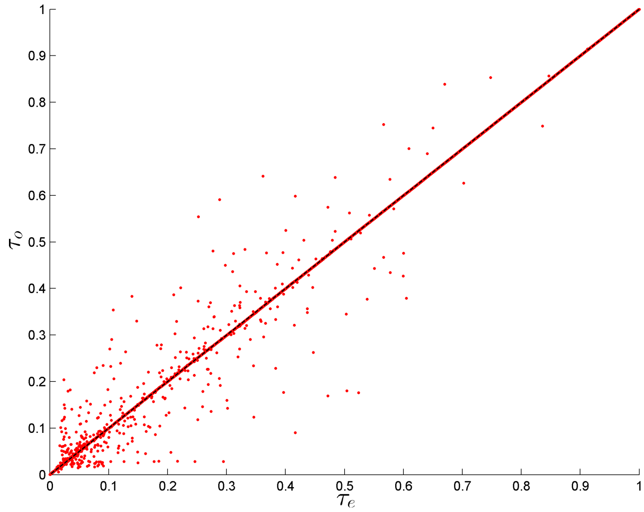
We conclude that Kendall's τ rank correlation coefficient can be used for the comparison of REMs in different scenarios using scalar values. However, for a deeper analysis of the causes leading to these values, the graphical representation from Section IV-B gives more information.

D. Erroneously Discarded Messages

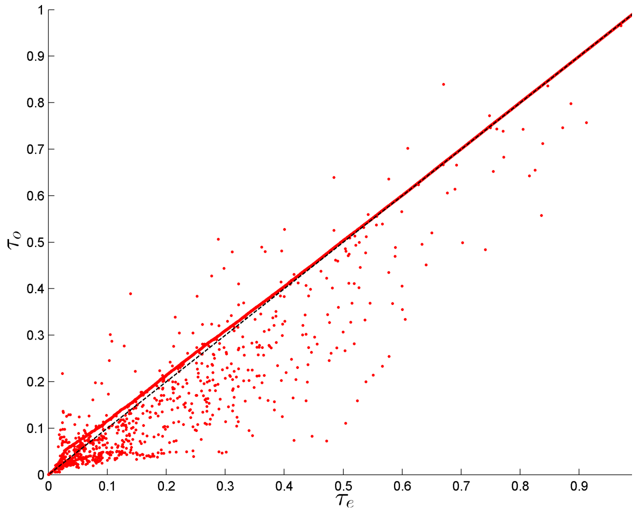
When messages are filtered by their relevance values, a certain amount of messages may be discarded because the relevance estimation was not precise enough. We aim for a comparative method for this phenomenon in this section.



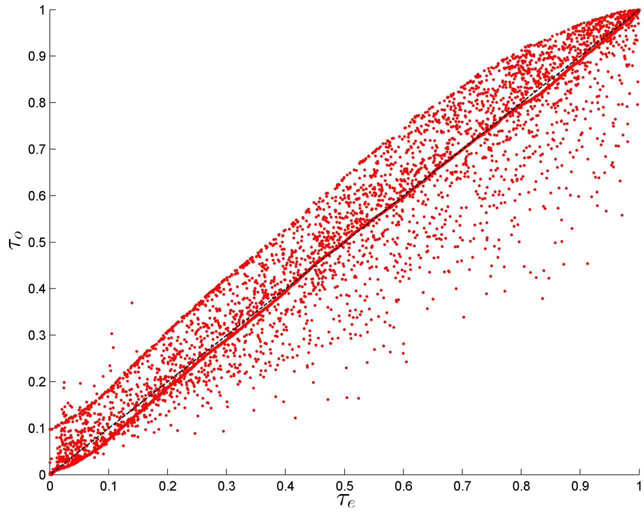
(a) Distance-based relevance estimation.



(b) Relevance estimation based on static movement approach.



(c) Relevance estimation based on dynamic movement approach.



(d) Encounter probability approach.

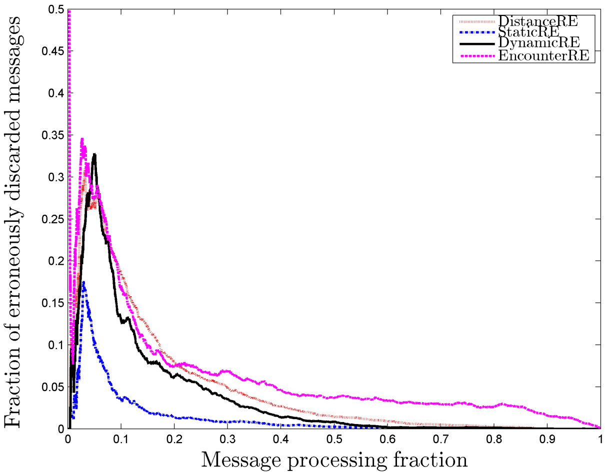
Fig. 2. Graphical representation of the message orders τ_e and τ_o induced by estimated and observed relevance values for the TAPASCologne scenario. The graphs show 1% of all data to avoid visual cluttering.

The message processing capabilities of VANET systems may differ significantly. Therefore, our approach does not aim for an analysis of a specific message processing rate but rather uses a relative representation, which we call the *message processing fraction* $p_{\text{processed messages}}$. The absolute rate of received messages corresponds to the system design and environmental assumptions [8], [10]. We compare $p_{\text{processed messages}}$ to the *fraction of erroneously discarded messages* $p_{\text{erroneously discarded}}$.

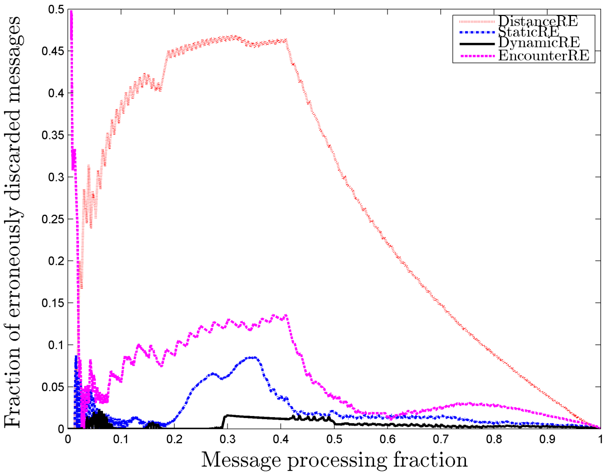
We use an example for clarification. Let the current rate of received messages in a VANET be 500 messages per second. A single system's processing capabilities allow only for a processing rate of 100 messages per second, so $p_{\text{processed messages}} = \frac{100}{500} = 0.2$ holds. Let the fraction of erroneously discarded messages $p_{\text{erroneously discarded}}$ be 30%. This means that 30 out of 100 processed messages per second are being filtered erroneously in this situation.

Figure 3(a) shows the results for the TAPASCologne scenario. We observe a steep increase for a low message processing fraction. The relevance estimation based on static movement reaches a maximum of $p_{\text{erroneously discarded}} \approx 18\%$ for $p_{\text{processed messages}} \approx 3\%$. This means that 18% of the messages did pass the filter erroneously while the same number of messages was discarded erroneously. We also observe that the relevance estimation based on dynamic movement performs similarly to the distance-based and encounter probability approaches which reach almost $p_{\text{erroneously discarded}} = 35\%$.

The situation is completely different for the curve scenario in Figure 3(b). The distance-based approach performs very badly, resulting in a $p_{\text{erroneously discarded}}$ of partially over 45%. The static movement-based approach reaches almost 10% for a message processing fraction of approximately 35%. In contrast to the TAPASCologne scenario, the dynamic movement-based approach yields the lowest $p_{\text{erroneously discarded}}$ values with a maximum of around 2%. We notice in this plot that low message processing fractions can also have low fractions of erroneously discarded messages. These results confirm the



(a) TAPASCologne scenario.



(b) Curve scenario.

Fig. 3. Erroneously discarded messages for both scenarios.

observation from the above sections. The performance of REMs has to be evaluated for a variety of scenarios, which correspond to the implemented applications.

V. Conclusion and Future Work

Relevance estimation mechanisms (REMs) are important for the selection of Cooperative Awareness Messages in VANETs. In this work we compared the quality of their results with regard to a ground truth that we defined on the basis of real vehicle traces. Both the ground truth and a REM provide relevance values for CAMs and induce an order on them. We compared these orders by illustrating their difference in a diagram and condensing it to a scalar value. Furthermore, we quantified the fraction of erroneously discarded CAMs, if only a certain percentage of most important CAMs can be processed.

We found that the simple static relevance estimation provides good results in a large city scenario, but when looking at a specifically demanding curve scenario, the more complex dynamic relevance estimation outperforms competitor REMs.

Future work includes the impact of actual hardware modules with limited processing capacity. Simulated traffic data can be replaced by data from real road traffic, e.g. by utilizing data from field tests like sim^{TD} [21]. The definition of the ground truth may be improved and adapted for specific applications, the same holds for REMs. Finally, the impact of discarded CAMs on applications has to be examined.

References

[1] M. Sichitiu and M. Kihl, “Inter-vehicle Communication Systems: A Survey,” *IEEE Communications Surveys & Tutorials*, vol. 10, no. 2, pp. 88–105, 2008.

[2] D. Jiang, Q. Chen, and L. Delgrossi, “Optimal data rate selection for vehicle safety communications,” p. 30.

[3] European Telecommunications Standards Institute, “Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service,” 2013-08.

[4] N. Lyamin, A. Vinel, and M. Jonsson, “Poster: On the performance of ETSI EN 302 637-2 CAM generation frequency management,” in *2014 IEEE Vehicular Networking Conference (VNC)*, pp. 107–108.

[5] John Harding, Gregory Powell, Rebecca Yoon, Joshua Fikentscher, Charlene Doyle, Dana Sade, Mike Lukuc, Jim Simons, and Jing Wang, “Vehicle-to-Vehicle Communications: Readiness of V2V Technology for Application: DOT HS 812 014,” Washington, DC, August, 2014.

[6] Car 2 Car Communication Consortium, “Memorandum of Understanding for OEMs within the CAR 2 CAR Communication Consortium on Deployment Strategy for Cooperative ITS in Europe,” 27.7.2011.

[7] J. Breu and M. Menth, “An Improved Relevance Estimation Function for Cooperative Awareness Messages in VANETs,” vol. 8435, pp. 43–56.

[8] J. Breu, A. Brakemeier, and M. Menth, “A quantitative study of Cooperative Awareness Messages in Production VANETs,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, p. 98, 2014.

[9] J. Breu and M. Menth, “Relevance Estimation of Cooperative Awareness Messages in VANETs,” in *Proceedings of the 5th International Symposium on Wireless Vehicular Communications 2013*, pp. 1–5.

[10] J. Breu, A. Brakemeier, and M. Menth, “Analysis of Cooperative Awareness Message rates in VANETs,” in *Proceedings of the 13th International Conference on ITS Telecommunications 2013*, pp. 8–13.

[11] N. Cenerario, T. Delot, and S. Ilarri, “A Content-Based Dissemination Protocol for VANETs: Exploiting the Encounter Probability,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 771–782, 2011.

[12] S. Singh and S. Agrawal, “VANET Routing Protocols: Issues and Challenges,” in *Recent Advances in Engineering and Computational Sciences 2014*, pp. 1–5.

[13] B. Defude, T. Delot, S. Ilarri, J.-L. Zechinelli, and N. Cenerario, “Data Aggregation in VANETs: The VESPA Approach,” *Proceedings of the 5th Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pp. 13:1–13:6, 2008.

[14] C. Adler, S. Eichler, T. Kosch, C. Schroth, and M. Strassberger, “Self-organized and Context-Adaptive Information Diffusion in Vehicular Ad Hoc Networks,” in *Proceedings of the 3rd International Symposium on Wireless Communication Systems 2006*, pp. 307–311.

[15] Z. Li and C. Chigan, “On Resource-Aware Message Verification in VANETs,” in *Proceedings of the International Conference on Communications 2010*, pp. 1–6.

[16] Car 2 Car Communication Consortium, *The Handbook for Vehicle-to-X Cooperative Systems Simulation*, 2011.

[17] SUMO - Simulation of Urban MObility. [Online]. Available: http://sumo.sourceforge.net

[18] TAPAS Cologne Scenario. [Online]. Available: http://sumo.dlr.de/wiki/Data/Scenarios/TAPASCologne

[19] DEKRA Automobil GmbH, “Road Safety Report 2013 Rural Roads,” Stuttgart, April 2013.

[20] M. G. Kendall and J. D. Gibbons, *Rank Correlation Methods*, 5th ed. London and New York, NY: E. Arnold and Oxford University Press, 1990.

[21] C. Weiß, “V2X Communication in Europe – From Research Projects Towards Standardization and Field Testing of Vehicle Communication Technology,” *Computer Networks*, vol. 55, no. 14, pp. 3103–3119, 2011.