Trustworthiness Estimation of Entities within Collective Perception

Christoph Allig*, Tim Leinmüller*, Prachi Mittal*, and Gerd Wanielik§
*DENSO AUTOMOTIVE Deutschland GmbH, Germany,
c.allig@denso-auto.de, t.leinmueller@denso-auto.de, p.mittal@denso-auto.de
§Chemnitz University of Technology, Department of Communication Engineering, gerd.wanielik@etit.tu-chemnitz.de

***Abstract*—The idea behind collective perception is to improve vehicles' awareness about their surroundings. Every vehicle shares information describing its perceived environment by means of V2X communication. Similar to other information shared using V2X communication, collective perception information is potentially safety relevant, which means there is a need to assess the reliability and quality of received information before further processing. Transmitted information may have been forged by attackers or contain inconsistencies e.g. caused by malfunctions.**

This paper introduces a novel approach for estimating a belief that a pair of entities, e.g. two remote vehicles or the host vehicle and a remote vehicle, within a Vehicular ad hoc Network (VANET) are both trustworthy. The method updates the belief based on the consistency of the data that both entities provide. The evaluation shows that the proposed method is able to identify forged information.

***Index Terms*—Collective perception, situational awareness, sensor fusion, V2X communication, Bayes filter, misbehavior detection, data consistency.**

I. Introduction

Intelligent Transport Systems (ITS) aim at enabling safe, efficient and coordinated transportation. Cooperative-ITS (C-ITS) complements ITS with Vehicle-to-X (V2X) communication, short for Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication. V2X communication can provide a variety of benefits for assisted and automated driving, including enhanced environmental awareness beyond the capabilities of traditional exteroceptive sensors. This is achieved through periodic exchange of kinematic information, i.e. vehicle position, speed, and heading, and can be even further improved through collective perception. Collective perception means that each partner shares information describing its ego state, sensing capabilities and sensed/perceived environment via V2X in so-called Collective Perception Messages (CPMs). As a result, apart from capturing several hundred meters of surroundings, sensing behind obstacles or curves is enabled.

Exteroceptive on-board sensors provide high reliability and the overall system is aware of its performance. The overall system knows that specific sensors may perform inferior in certain scenarios, e.g. camera performance is worse than radar in rain, fog or darkness. Similarly, entities within the Vehicular ad hoc Network (VANET) have different and changing performance. Reduced performance might be caused by weather, different equipment grades, malfunction, or a malicious entity initiating cyber attacks. Applications relying on such erroneous information received via V2X communication would be severely interfered. In particular, safety applications such as adaptive cruise control, collision avoidance system, or intersection assistant may endanger road users if they rely on erroneous information. Thus, the detection of erroneous data is indispensable.

There are basically two types of mechanisms to prevent harm caused by erroneous data, cryptographic mechanisms and misbehavior detection (and mitigation). Cryptographic mechanisms, which in C-ITS are realized using pseudonym certificates issued by public key infrastructures (PKIs), assure that only authorized entities such as vehicles and infrastructure devices can prove being a legitimate VANET participant. This increases the effort for potential attackers, as they first have to obtain valid credentials to appear as a legitimate network participant. Nevertheless, authorized entities may exist that transmit faulty information or exhibit malicious behavior. For detecting this insider misbehavior detection mechanisms are applied. Misbehavior detection assesses the behavior and trustworthiness of entities, as well as the consistency and plausibility of the information they transmit. Misbehavior detection can be carried out by individual entities, or by multiple entities in collaboration. While the collaborative approach has higher potential for misbehavior detection, it disadvantageously relies on a honest majority assumption.

In this paper, we present a novel approach for estimating a belief that a pair of entities within a VANET are both trustworthy. The underlying idea is to enhance the approach we present in [1] through probabilistic modeling. We apply the Bayes filter for probabilistic modeling, which updates the belief based on the consistency of the data that both entities provide.

The paper is structured as follows. The next section summarizes related work, followed by a section that describes background regarding collective perception and feasible malicious attacks. The approach for estimating the trustworthiness, i.e. a belief that indicates if the exchanged data is trustworthy, is described in Section IV. Finally, our approach is evaluated in Section V, and Section VI concludes the paper.

II. Related Work

The basic security concepts for VANETs have been introduced for instance in [2], including the aforementioned

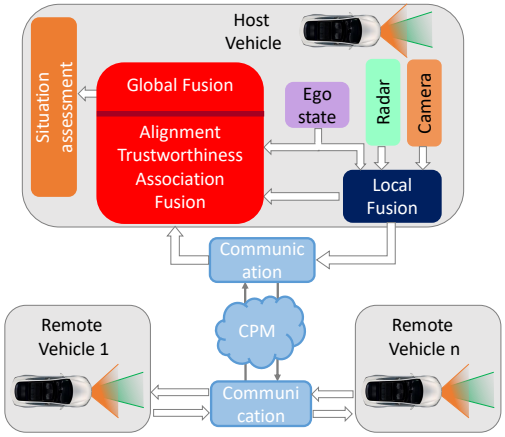

Fig. 1: Architecture for Collective Perception.

cryptographic mechanisms and misbehavior detection. Detailed surveys on misbehavior detection in VANETs, including position verification, are presented in e.g. [3]–[5].

Multiple mechanisms for position verification are proposed in [6]. They include checking whether the sending entity is located within the maximum communication range, or whether it pretends to move at impossible speed. The mechanisms also include checks taking into account that only a limited number of vehicles can be located within an area, map-based verification for identifying unlikely positions, e.g. off the streets, or in houses, and position claim overhearing. These position verification mechanisms are complemented by cooperative approaches in [7], exchanging neighbor tables or reactive position verification upon demand.

In [8] and [9], the claimed position is compared to a corresponding vehicle tracker, which estimates the predicted position based on previously received information utilizing a Kalman filter.

Data from different types of information sources such as messages, radar sensor and map data are aggregated to evaluate the trustworthiness of neighboring vehicles in [10]. A Particle filter is applied, which increases or decreases the particle weights depending on the consistency of the claimed position.

In [11], multi-object tracking is proposed to verify the consistency of positions in messages by means of already available existence estimates.

To the best of our knowledge, there has been no work using probabilistic modeling for estimating a trustworthiness in collective perception in VANETs.

III. Background

This section provides background information concerning the assumed basic system architecture for collective perception, how to use collective perception for position verification, and the attacker assumptions.

A. Architecture for Collective Perception

The basic architecture is shown in Figure 1. The main components are the local fusion module, the communication module, and the global fusion module.

EIS	ID	t	c	$\boldsymbol{d}$	dynamic state $[\varphi \ \lambda \ h \ \psi \ v \ a \ \dot{\psi}]$

Fig. 2: Simplified Cooperative Awareness Message (CAM).

EIS	ID	t	c	$\boldsymbol{d}$	dynamic state $[\varphi \ \lambda \ h \ \psi \ v \ a \ \dot{\psi}]$	
FOVs	x	y	r_{min}	r_{max}	α_{min}	α_{max}
PDOs	ID	$\boldsymbol{c}$	$\boldsymbol{d}$	$p(\exists)$	dynamic state $[x \ y \ g \ v]$	

Fig. 3: Simplified Collective Perception Message (CPM).

The local fusion module fuses measurements from the exteroceptive on-board sensors. The environmental awareness of a vehicle is typically gathered by exteroceptive on-board sensors such as radar, camera or lidar sensors. The sensor observations are fused within the local fusion module by means of multi-object tracking (MOT) [12] and multi-sensor fusion algorithms [13], [14]. The local fusion module provides as output a list of tracked objects. A unique object id is assigned to each object. In addition to the dynamic state, the existence probability, classification and vehicle extent are estimated. The dynamic state typically includes Cartesian position, velocity and heading information.

The communication module is responsible for disseminating messages over the VANET, e.g. Cooperative Awareness Messages (CAMs) and CPMs. V2X communication can partially be considered as an additional sensor that provides redundancy and an increased perception range, enabling the perception of occluded objects. Messages are exchanged periodically via wireless communication, either using IEEE802.11p [15], or using LTE-V2X. For exchanging information various message formats have already been standardized. The CAM [16] in Europe and the Basic Safety Message (BSM) [17] in the US enable the exchange of information that describes the ego state of the remote vehicle. Standardization for collective perception has recently been started at the European Telecommunications Standard Institute (ETSI) [18], [19]. The basic content of the CAM is shown in Figure 2 and the content for CPM in Figure 3. Both message types contain information that describe the Ego Information State (EIS) of the remote vehicle. The CPM contains additional information describing the Field of Views (FOVs) of the exteroceptive sensors the remote vehicle is equipped with and the perceived dynamic objects (PDOs) by the remote vehicle. The EIS provides a unique remote vehicle ID, the type c, dimensions $\mathbf{d}$ and dynamic state $\begin{bmatrix}\varphi & \lambda & h & \psi & v & a & \dot{\psi}\end{bmatrix}^T$ of the remote vehicle. The dynamic remote vehicle state is described by the latitude φ, longitude λ, altitude h, yaw angle ψ, velocity v, acceleration a and yaw rate $\dot{\psi}$. The provided time stamp t is valid for the total message. The FOV for each sensor might be given by the sensor position x, y, range r_{min}, r_{max} and opening angle α_{min}, α_{max}. Each PDO is identifiable by a unique object ID

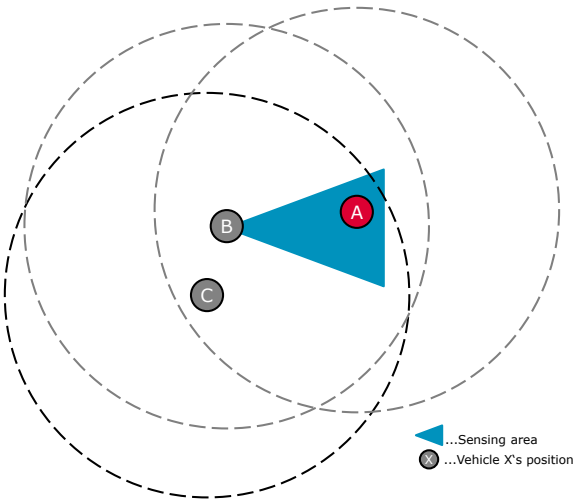

Fig. 4: Example 1: Positive Verification of Third Party Position Claim (Concept 1).

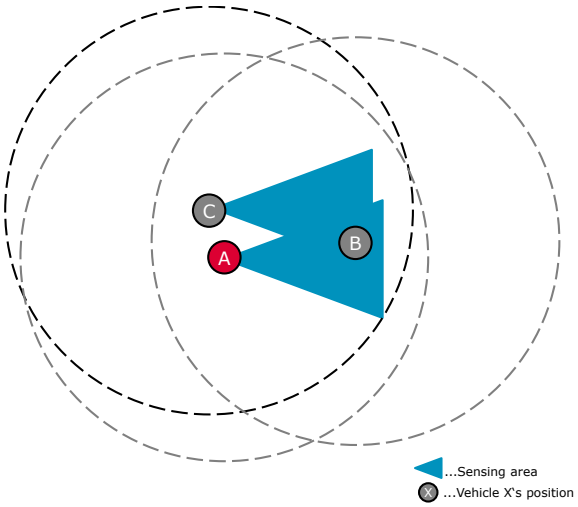

Fig. 5: Example 2: Positive Verification of CPM Sender Position (Concept 2).

per remote vehicle. Next to the dynamic state $\begin{bmatrix} x & y & g & v \end{bmatrix}^T$, the type $\mathbf{c}$, dimension $\mathbf{d}$ and existence $p\,(\exists)$ estimate might be transmitted. CAMs are typically transmitted with a frequency between 1 to 10Hz conditional on the scenario. CPMs are envisioned to be transmitted at similar but generally lower frequencies.

The global fusion module merges the received communication data with the local fusion output resulting in an enhanced environmental awareness. Firstly, the received data must be aligned in time and space [20]. We suggest that the next step is to inspect for misbehavior in order to assign a trustworthiness to the received data. Subsequently, the data is associated. In [21] a comparison is made between the nearest neighbor, modified auction, and suboptimal joint probabilistic data association. Finally, the data is fused. For the dynamic state, track-to-track fusion methods such as covariance intersection [22] or information matrix fusion [23] are appropriate. For merging existence probabilities, in [24] an appropriate method for track-to-track fusion is proposed.

B. Collective Perception for Position Verification

Beyond improving environmental awareness, information obtained through collective perception can be used to verify position claims in VANETs. Our previous work [1] presents two concepts for position verification in VANETs that make use of collective perception.

The first concept uses information from CPMs to verify position claims of third party vehicles, e.g. a vehicle C uses CPMs from a vehicle B to verify position claims of a vehicle A. If, as shown in the example in Figure 4, B's sensors cover A's claimed location, and if B's CPMs contain the same position, then C can use this as a positive confirmation of A's position (assuming that C trusts B).

The second concept uses CPMs to verify position claims of the CPM sending vehicle, e.g. a vehicle C uses CPMs from a vehicle A to verify vehicle A's position claims. The second concept relies on C having the possibility to determine what A should include in its CPMs if A is at its claimed location. A corresponding example is shown in Figure 5. C is aware of B's location through C's exteroceptive on-board sensors and B's CAMs. C also knows that A should include B in A's CPMs as B is within A's sensing area. If A does accordingly, this is a positive verifications af A's position.

C. Attacker Model

In this work, we assume that the attacker is an insider, i.e. he has valid key material and he is capable of signing his messages with valid signatures to appear like for instance a normal vehicle. The attacker is forging the message content. There are various conceivable forgery possibilities, e.g. modifying the ego state or of one or several PDOs. Modification of a single state variable as well as multiple state variables is conceivable. Depending on the intention, the attacker might fake for instance an acceleration or turning maneuver, or position information. Alternatively, adding a "ghost" PDO, though also omission of PDOs is realistic. However, the capabilities of the attacker are limited to the assumption acting non-cooperatively.

IV. Trustworthiness Estimation

This section presents our novel approach for estimating the belief that a pair of entities within a VANET are both trustworthy. The basic idea is to estimate the trustworthiness by assessing data consistency taking into account likelihoods, i.e. to quantify the confidence in the correctness of the data.

Firstly, we explain why we estimate the probability that a pair of entities are trustworthy. Data consistency assessment requires that at least two redundant data sets are available, which are checked against each other.

A straightforward approach would be to assess data consistency for each entity against the truth. However, the truth is unknown. Accordingly, a reference is mandatory. One option

may be to utilize the global awareness of the host vehicle as reference. This requires ensuring the prevention of infiltration of forged or erroneous data into the global awareness. Otherwise, the attacker might for instance forge an object located in an area that has been unobservable by other entities so far. Thus, with no inconsistency being detectable, the forged object will be assumed to be existent and the attacker will be classified as trustworthy. Additionally, a trustworthy entity might consequently be classified untrustworthy due to inconsistency.

That is why we chose to assess the incoming data sets directly amongst each other. However, multiple data sets complicate the source identification of inconsistencies. Consequently, trustworthiness is estimated in each case for two entities, i.e. two remote vehicles or the host vehicle and a remote vehicle. The concept is that each pair of entities is assigned a belief representing that both entities are trustworthy. The belief is expected to increase or decrease depending on the consistency of the environmental awareness of entity 1 compared to entity 2. In Section IV-A our novel approach for estimating trustworthiness probability is described. For estimating the trustworthiness probability, we propose to apply the well-known Bayes theorem. Section IV-B discusses applied parameters, the detection probability and the not forged object parameter.

A. Bayes Filter

Initially, we define the problem of trustworthiness estimation. In trustworthiness estimation the state consists of just a single, binary hypothesis, i.e. either entity 1 and entity 2 are both trustworthy ($\mathcal{T}^{e_1 \wedge e_2}$) or at least one of both is untrustworthy ($\not{\mathcal{T}}^{e_1 \vee e_2}$). For short notation the abbreviations ($\mathcal{T}$) and ($\not{\mathcal{T}}$) are introduced. This problem is analogous to the one we have in object existence estimation [25]. The probability that both entities are trustworthy at time k conditioned on all information about the tracks lists of both entities up to time k is denoted $p\left(\mathcal{T}_k^{e_1 \wedge e_2} | T_{e_1}^k, T_{e_2}^k\right)$. The information about both track lists is hereafter abbreviated by T^k. The probability that at least one of both is untrustworthy is given by the complement.

In order to solve the problem of trustworthiness estimation we propose to apply a Bayes filter, which recursively estimates the object state, i.e. the probability whether both entities are trustworthy or not, for each time step k. The state might change over time, which is modeled by the system model that predicts the state from the previous time $k-1$ to the current time k, which is described in Section IV-A1. Generally, there is an observable, which is related to the state according to an observation model. Section IV-A2 describes the update step that incorporates the observable, which is the result of the data consistency assessment.

1) Prediction: The prior trustworthiness probability is predicted using the state transition probabilities $p_{c\mathcal{T}}$ and $p_{\bar{c}\not{\mathcal{T}}}$ according to:

$$p\left(\mathcal{T}_k | T^{k-1}\right) = p_{c\mathcal{T}} p\left(\mathcal{T}_{k-1} | T^{k-1}\right) + p_{\bar{c}\not{\mathcal{T}}} p\left(\not{\mathcal{T}}_{k-1} | T^{k-1}\right), \tag{1}$$

$$\begin{aligned} &p\left(\not{\mathcal{T}}_k | T^{k-1}\right) = \\ &\quad [1 - p_{c\mathcal{T}}]\, p\left(\mathcal{T}_{k-1} | T^{k-1}\right) + [1 - p_{\bar{c}\bar{\not{\mathcal{T}}}}]\, p\left(\not{\mathcal{T}}_{k-1} | T^{k-1}\right), \end{aligned} \tag{2}$$

where $p_{c\mathcal{T}}$ is the probability that both entities continue to be trustworthy. The parameter $p_{\bar{c}\not{\mathcal{T}}}$ models the probability that one or both entities does not continue to be untrustworthy, i.e. both entities are trustworthy at time k, but at the previous time $k-1$ at least one of both was untrustworthy. The parameter $p_{c\mathcal{T}}$ is chosen slightly smaller 1 to model the assumption that both trustworthy entities likely continue to be trustworthy, but with the probability $1 - p_{c\mathcal{T}}$ at least one changes into an untrustworthy entity. The parameter $p_{\bar{c}\not{\mathcal{T}}}$ is chosen slightly greater than 0 to model the assumption that with a small chance all untrustworthy entities might become trustworthy, but likely the untrustworthy entity will continue to be untrustworthy.

2) Update: The trustworthiness is updated through checking consistency between the tracks lists T_{e_1} and T_{e_2} of entity 1 and entity 2, respectively. In section IV-A2a and IV-A2b probabilities for the different observable events are modeled. In Section IV-A2c we suggest determining a compound event probability to reduce computational effort, before finally the trustworthiness is updated using the compound probability, which is the result of the consistency check. It is to be noted that each list includes the perceived object tracks as well as the ego track. Accordingly, CAMs might also be used for assessing the consistency.

a) Association: The probability that a track t_j is available in the tracks list of entity 1 that corresponds to a track t_i in the tracks list of entity 2 conditioned on that entity 1 and entity 2 are trustworthy is $p\left(t_j \in T_{e_1} \sim t_i \in T_{e_2} | \mathcal{T}_k^{e_1 \wedge e_2}\right)$ and hereafter abbreviated by $p\left(t | \mathcal{T}_k\right)$. It depends on the object existence probability and the detection probability that the object is detectable by the sensors of entity 1:

$$p\left(t | \mathcal{T}_k\right) = p_{d_{e_1}}\left(t_i\right) p_{\exists_{e_2}}\left(t_i\right). \tag{3}$$

As discussed before, a reference is needed, for which entity 2 is used. Hence, the estimation of entity 2 is used for the probability of existence.

The probability that the track t_j is available in the tracks list of entity 1 that corresponds to a track t_i in the tracks list of entity 2 conditioned on that at least one of the entities is untrustworthy is $p\left(t_j \in T_{e_1} \sim t_i \in T_{e_2} | \not{\mathcal{T}}_k^{e_1 \vee e_2}\right)$ and hereafter abbreviated by $p\left(t | \not{\mathcal{T}}_k\right)$. The object is expected to be present with the aforementioned likelihood lowered by the parameter $\bar{p}_f$ that estimates the portion of not forged objects. However, the object is at least present with a clutter probability p_c that models the likelihood of a forged object being accidentally assignable to a true object:

$$\begin{aligned} &p\left(t | \not{\mathcal{T}}_k\right) \\ &\quad = \begin{cases} p_c, & \bar{p}_f p_{d_{e_1}}\left(t_i\right) p_{\exists_{e_2}}\left(t_i\right) < p_c \\ \bar{p}_f p_{d_{e_1}}\left(t_i\right) p_{\exists_{e_2}}\left(t_i\right), & \text{else} \end{cases} \end{aligned} \tag{4}$$

The clutter probability p_c is modeled as a constant slightly greater than 0.

b) No Association: The probability that no track t_j is available in the tracks list of entity 1 that is assignable to a track t_i in the tracks list of entity 2 conditioned on that entity 1 and entity 2 are trustworthy is $p\left(t_j \notin T_{e_1} \sim t_i \in T_{e_2} | \mathcal{T}_k^{e_1 \wedge e_2}\right)$ and hereafter abbreviated by $p\left(\bar{t}|\mathcal{T}_k\right)$. It is the complement of (3):

$$p\left(\bar{t}|\mathcal{T}_k\right) = 1 - p\left(t|\mathcal{T}_k\right). \tag{5}$$

Accordingly, the probability that no track t_j is available in the tracks list of entity 1 that is assignable to a track t_i in the tracks list of entity 2 conditioned on that at least one of the entities is untrustworthy is $p\left(t_j \notin T_{e_1} \sim t_i \in T_{e_2} | \not{\mathcal{T}}_k^{e_1 \vee e_2}\right)$ and hereafter abbreviated by $p\left(\bar{t}|\not{\mathcal{T}}_k\right)$. It is the complement of (4):

$$p\left(\bar{t}|\not{\mathcal{T}}_k\right) = 1 - p\left(t|\not{\mathcal{T}}_k\right). \tag{6}$$

c) Combining the Observable: Assuming that the individual events whether a track is available or not are conditionally independently distributed, the compound probability of the compound event conditioned on whether both entities are trustworthy or not can be determined as follows:

$$p\left(E|\mathcal{T}_k\right) = \prod_{n=0}^{N} p_n\left(t|\mathcal{T}_k\right) \prod_{m=0}^{M} p_m\left(\bar{t}|\mathcal{T}_k\right), \tag{7}$$

$$p\left(E|\not{\mathcal{T}}_k\right) = \prod_{n=0}^{N} p_n\left(t|\not{\mathcal{T}}_k\right) \prod_{m=0}^{M} p_m\left(\bar{t}|\not{\mathcal{T}}_k\right). \tag{8}$$

The number N states that N tracks have been assigned between tracks list T_{e_1} and T_{e_2} and vice versa. Accordingly, there are M tracks that have not been associated.

The probabilities that both entities are trustworthy and its complement that at least one entity is untrustworthy are finally updated according to:

$$p\left(\mathcal{T}|T^k\right) = \eta p\left(E|\mathcal{T}_k\right) p\left(\mathcal{T}|T^{k-1}\right), \tag{9}$$

$$p\left(\not{\mathcal{T}}|T^k\right) = \eta p\left(E|\not{\mathcal{T}}_k\right) p\left(\not{\mathcal{T}}|T^{k-1}\right), \tag{10}$$

where the posterior trustworthiness with information from the tracks lists of entity 1 and entity 2 up to time k is denoted by the probability that both entities are trustworthy $p\left(\mathcal{T}|T^k\right)$ and the probability that at least one is untrustworthy $p\left(\not{\mathcal{T}}|T^k\right)$. The prior trustworthiness with information from the tracks lists of entity 1 and entity 2 up to time $k-1$ is denoted by $p\left(\mathcal{T}|T^{k-1}\right)$ and $p\left(\not{\mathcal{T}}|T^{k-1}\right)$, respectively. The measurement information derived from the tracks lists at time k is given by $p\left(E|\mathcal{T}_k\right)$ and $p\left(E|\not{\mathcal{T}}_k\right)$. The normalizing factor η is defined as:

$$\eta = \frac{1}{p\left(E|\mathcal{T}_k\right) p\left(\mathcal{T}|T^{k-1}\right) + p\left(E|\not{\mathcal{T}}_k\right) p\left(\not{\mathcal{T}}|T^{k-1}\right)}. \tag{11}$$

B. Modeling the Parameters

The previously described method employs several parameters. While most parameters have already been sufficiently described, this section provides a more detailed description of the detection probability p_d and the parameter that estimates the portion of not forged objects $\bar{p}_f$.

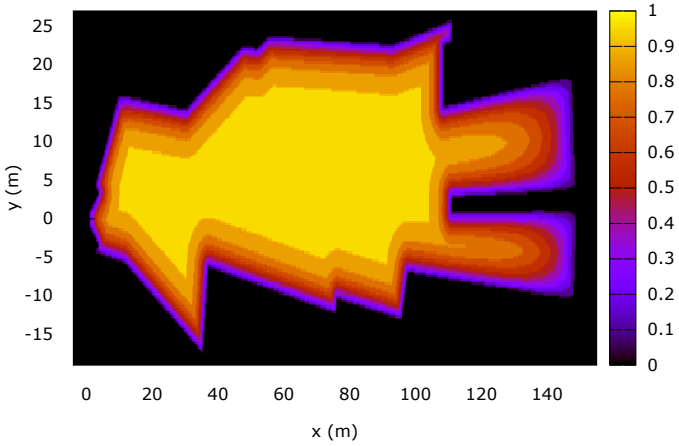

Fig. 6: Detection probabilities for an entity located at $x = y = 0$ that is equipped with forward facing sensors.

1) Detection Probability: The detection probability p_d represents the probability whether or not an entity is capable of providing valid information about an object. It is modeled using the FOVs of the sensors and obstacle occlusions, as described in [26] for the persistence probability. A joint FOV or more specifically a visibility polygon is specified based on the FOVs of the sensors and the surrounding obstacles. The FOV of each sensor is restricted by its minimum and maximum range and angle. The detection probability drops near the limits of the FOV, as described in [27]. The surrounding obstacles comprise only of confirmed tracked objects as well as static objects from map data. Based on the extent of the obstacles, the obscured area is modeled, in which the detection probability converges to 0 for line-of-sight sensors. If the entity is equipped with sensors that are able to detect occluded objects such as radar, a more sophisticated model is required. Figure 6 shows an exemplary visibility polygon including the detection probabilities for an entity located at $x = y = 0$ that is equipped with forward facing sensors. Whether the object is inside or outside the visibility polygon is checked by ray casting [28]. Depending on the distance to the closest border of the visibility polygon, the detection probability is additionally reduced to take into account likely missed detections near the obscured areas. Besides, the visibility polygon is supplemented by one's own awareness.

2) Not Forged Object Parameter: The parameter $\bar{p}_f$, which estimates the portion of not forged objects, is determined based on the detection probability and whether there exists an assignable track. The number of times the probability that entity 1 is able to detect track t_i exceeds a threshold t is counted for all objects in the tracks list of entity 2. Additionally, we count how often no track from entity 1 is assignable to a track from the tracks list of entity 2 conditioned on that entity 1 should be able to detect track t_i. The probability $\bar{p}_f$

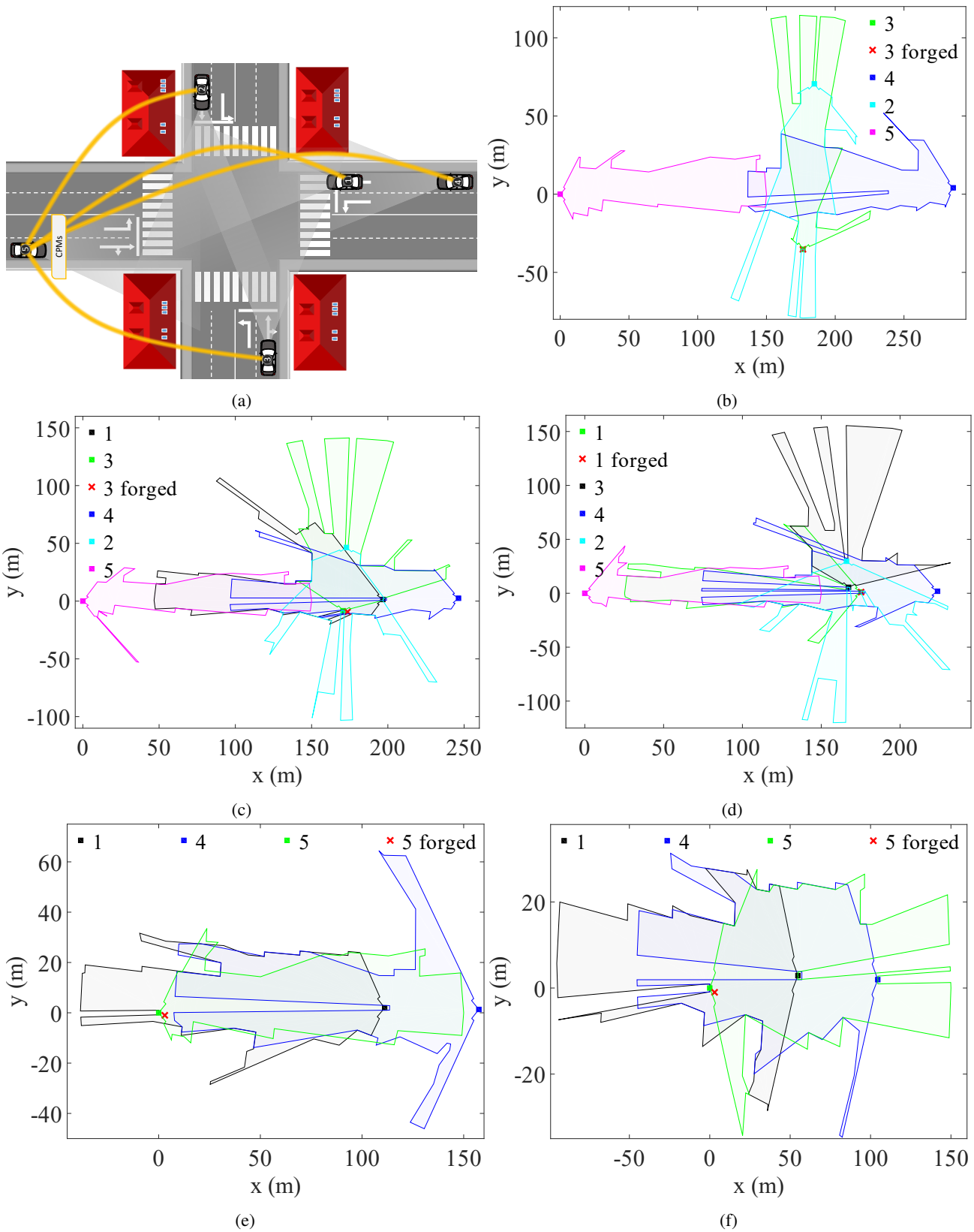

Fig. 7: Emulated attacks.

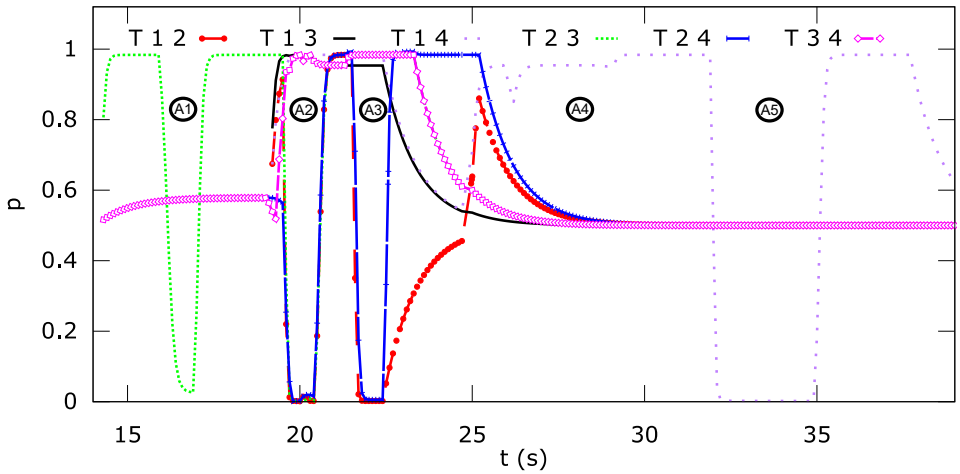

Fig. 8: Estimated probability that entity a and b are trustworthy.

is finally calculated by the following relation:

$$\bar{p}_f = \frac{\sum\limits_{i}^{\forall p_d > t} t_i \in T_{e_2} - \sum\limits_{i}^{\forall t_j \notin T_{e_1} \sim t_i \in T_{e_2} | p_d > t} t_i \in T_{e_2}}{\sum\limits_{i}^{\forall p_d > t} t_i \in T_{e_2}}. \quad (12)$$

In case the second term in the numerator or the denominator is zero, $\bar{p}_f$ is set to constant smaller one and greater zero.

V. Evaluation

Within the scope of the evaluation the trustworthiness estimation method described in Section IV is evaluated. The evaluation is based on simulation [29] that provides mobility data and sensor measurements. The simulation scenario includes five vehicles which are exchanging CPMs, i.e. four remote vehicles denoted hereafter as entity 1 to 4 and one host vehicle denoted hereafter as entity 5. All entities are crossing the same intersection. Entity 4 follows entity 1 driving northwards. Entity 2, 3 and 5 drive westwards, eastwards and southwards, respectively.

For the purpose of assessing the proposed method, five attacks are emulated, which are outlined in Figure 7. Positions and FOVs of the host vehicle as well as of remote vehicles that are currently transmitting CPMs are displayed in the host vehicle body frame. The sub-figures are arranged chronologically so that the movement of vehicles is deducible. During $16\text{s} < t < 17\text{s}$ entity 2 attacks the VANET by transmitting forged heading information about entity 3. Simultaneously, entity 3 is itself transmitting CPMs that include ego state information, which is why the host vehicle receives conflicting information. However, it is not possible to deduce which entity is executing the attack. The subsequent attack at $19.5\text{s} < t < 20.5\text{s}$ involves once again forging information about entity 3 by entity 2, though position data this time. Since there is correct information from entity 1 and 4 in addition to from entity 3 describing entity 3, the contradiction should be retraceable to attacker entity 2. During $21.5\text{s} < t < 22.5\text{s}$ forged velocity information about entity 1 is transmitted by entity 2. The host vehicle receives further information from entity 1 and 4 describing entity 1. The subsequent two attacks at $28\text{s} < t < 29\text{s}$ and $32\text{s} < t < 35\text{s}$ involve the transmission of forged position data about entity 5. Inititally, entity 1 is executing the attack, which is not detectable due to absence of redundancy. The fifth attack is performed by entity 4, while entity 1 provides contradictory, though correct information. It is to be noted that during evaluation data provided by the host vehicle is ignored.

In Figure 8 the result of the trustworthiness estimation is shown for any pair of the four remote vehicles. It shows the estimated probability that both entities are trustworthy. The attacks are indicated by A1 ... A5. Initially, the probability that entity 2 and 3 are trustworthy increases and converges to 1. The probability that entity 2 and 4 as well as 3 and 4 are trustworthy increases slightly and keeps constant, since their FOVs are partly overlapping but without a common object. During the first security attack the probability that entity 2 and 3 are trustworthy converges to 0. Once an object enters the FOV of both entities, its probability increases or slightly decreases depending on the promptness that the object is confirmed. The amount of decrease highly depends on the detection probability model. During the second and third security attack, the probability that entity 1 and 2 as well as entity 2 and 4 are trustworthy converge to 0. The same applies for entity 2 and 3 for the second security attack. Once, no object is present inside an entity's FOV, its corresponding trustworthiness probability converges to 0.5. As expected, the fourth security attack is not detected, while the fifth security attack results in that the probability of entity 1 and 4 being trustworthy converges to 0.

VI. Conclusions

Security in VANETs remains challenging, even after more than a decade of R&D in the domain. Through research, numerous concepts have been developed that increase security, be it by introducing (pseudonym-)certificates to make if diffi-

cult for attackers to appear as legitimate network participant, or by developing approaches for misbehavior detection / trustworthiness estimation.

In this paper, we present a novel approach for quantifying the confidence in the correctness of data that is transmitted within VANETs for the purpose of collective perception. The underlying idea is to assess consistency of the received data in order to recursively estimate the trustworthiness of transmitting entities. We utilize Bayes theory for recursive estimation and we provide observation models. The observation model expresses the relation between the result of the data consistency assessment and the trustworthiness state.

The stimulative investigation of our approach shows that it is appropriate for trustworthiness estimation. Contradictions between received data caused by an attack result in low trustworthiness belief. Furthermore, the approach allows the attacker to be identified if at least triple redundancy exists for the contradiction.

Opportunities for future work include real-world testing, particularly to examine the detection probability model. Also of interest is the investigation of an approach that combines the result of multiple misbehavior detection methods using recursive estimation in order to obtain a common trustworthiness belief. We further think that the result of the trustworthiness estimation might be applied to improve the robustness of the global fusion.

ACKNOWLEDGMENT

This work was partially funded by the European Union's Horizon 2020 research and innovation program under grant agreement No. 688900.

REFERENCES

[1] T. Leinmüller, P. Mittal, and C. Allig, "Using Collective Perception for Position Verification in VANETs," in *To appear in Proceedings of 26th ITS World Congress*, 2019.

[2] T. Leinmüller, E. Schoch, and C. Maihöfer, "Security issues and solution concepts in vehicular ad hoc networks," in *Proceedings of the Fourth Annual Conference on Wireless On demand Network Systems and Services (WONS 2007)*, (Obergurgl, Austria), Jan. 2007.

[3] R. W. Van Der Heijden, S. Dietzel, T. Leinmüller, and F. Kargl, "Survey on Misbehavior Detection in Cooperative Intelligent Transportation Systems," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 1, pp. 779–811, 2018.

[4] H. Criuckshank, N. Ahmad, M. Khalid, M. Arshad, Y. Cao, and Z. Ullah, "A Survey of Local/Cooperative-Based Malicious Information Detection Techniques in VANETs," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, no. 1, pp. 1–17, 2018.

[5] F. Sakiz and S. Sen, "A Survey of Attacks and Detection Mechanisms on Intelligent Transportation Systems : VANETs and IoV," *Ad Hoc Networks*, vol. 61, pp. 33–50, 2017.

[6] T. Leinmüller, C. Maihöfer, E. Schoch, and F. Kargl, "Improved Security in Geographic Ad hoc Routing through Autonomous Position Verification," in *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks*, pp. 57–66, 2006.

[7] T. Leinmüller, E. Schoch, F. Kargl, and C. Maihöfer, "Dezentralized Position Verification in Geographic Ad hoc Routing," *Security and Communication Networks*, vol. 3, no. 4, pp. 289–302, 2010.

[8] H. Stubing, A. Jaeger, C. Schmidt, and S. A. Huss, "Verifying Mobility Data Under Privacy Considerations in Car-to-X Communication," in *17th ITS World Congress*, 2010.

[9] A. Jaeger, N. Bißmeyer, H. Stübing, and S. A. Huss, "A Novel Framework for Efficient Mobility Data Verification in Vehicular Ad-hoc Networks," *International Journal of Intelligent Transportation Systems Research*, vol. 10, no. 1, pp. 11–21, 2012.

[10] N. Bismeyer, S. Mauthofer, K. M. Bayarou, and F. Kargl, "Assessment of Node Trustworthiness in VANETs using Data Plausibility Checks with Particle Filters," in *IEEE Vehicular Networking Conference*, pp. 78–85, 2012.

[11] M. Obst, L. Hobert, and P. Reisdorf, "Multi-Sensor Data Fusion for Checking Plausibility of V2V Communications by Vision-based Multiple-Object Tracking," in *IEEE Vehicular Networking Conference*, pp. 143–150, 2014.

[12] B.-n. Vo, M. Mallick, Y. Bar-shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget Tracking," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–15, 2015.

[13] H. F. Durrant-Whyte and T. C. Henderson, "Multisensor Data Fusion," *Handbook of Robotics*, pp. 585–610, 2008.

[14] D. Smith and S. Singh, "Approaches to Multisensor Data Fusion in Target Tracking: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 12, pp. 1696–1710, 2006.

[15] IEEE 802.11 p Working Group, "IEEE Standard for Information Technology – Local and Metropolitan Area Networks– Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments," *IEEE Std; IEEE: Piscataway, NJ, USA*, 2010.

[16] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2 : Specification of Cooperative Awareness Basic Service," *EN 302 637-2 - V1.3.2*, 2014.

[17] SAE International, "Dedicated Short Range Communications (DSRC) Message Set Dictionary," *SAE J2735*, 2016.

[18] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Specification of the Collective Perception Service," *Draft TS 103 324 V0.0.12*, 2017.

[19] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Analysis of the Collective Perception Service (CPS)," *Draft TR 103 562 V0.0.16*, 2019.

[20] C. Allig and G. Wanielik, "Alignment of Perception Information for Cooperative Perception," in *IEEE Intelligent Vehicles Symposium*, pp. 1639–1644, 2019.

[21] F. Seeliger, *Fahrzeugübergreifende Informationsfusion für ein Kreuzungsassistenzsystem*. PhD thesis, Universität Ulm, 2017.

[22] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," *Proceedings of the 1997 American Control Conference*, vol. 4, pp. 2369–2373, 1997.

[23] B. Belkin, S. L. Anderson, and K. M. Sommar, "The Pseudo-Measurement Approach to Track-to-Track Data Fusion," in *Joint Service Data Fusion Symposium*, pp. 519–538, 1993.

[24] M. Aeberhard, S. Paul, N. Kaempchen, and T. Bertram, "Object Existence Probability Fusion using Dempster-Shafer Theory in a High-Level Sensor Data Fusion Architecture," in *IEEE Intelligent Vehicles Symposium*, pp. 770–775, 2011.

[25] R. Altendorfer and S. Matzka, "A Confidence Measure for Vehicle Tracking based on a Generalization of Bayes Estimation," in *IEEE Intelligent Vehicles Symposium*, pp. 766–772, 2010.

[26] M. Maehlisch, W. Ritter, and K. Dietmayer, "De-cluttering with Integrated Probabilistic Data Association for Multisensor Multitarget ACC Vehicle Tracking," *IEEE Intelligent Vehicles Symposium*, pp. 178–183, 2007.

[27] M. Aeberhard, *Object-Level Fusion for Surround Environment Perception in Automated Driving Applications*. PhD thesis, Universität Ulm, 2017.

[28] M. Shimrat, "Algorithm 112: Position of Point Relative to Polygon," *Communications of the ACM*, vol. 5, no. 8, p. 434, 1962.

[29] C. Allig and G. Wanielik, "Extending the Vehicular Network Simulator Artery in order to Generate Synthetic Data for Collective Perception," *Advances in Radio Science*, vol. 17, pp. 189–196, 2019.