Privacy of Smart Traffic Lights Systems

Artur Hermann, Michael Wolf, Nataša Trkulja, Ines Ben-Jemaa, Anis Bkakria, Frank Kargl

Privacy of Smart Traffic Lights Systems

Artur Hermann*, Michael Wolf*, Nataša Trkulja*, Ines Ben Jemaa†, Anis Bkakria†, and Frank Kargl*
*Ulm University, Institute of Distributed Systems
†IRT System X

***Abstract*—Smart traffic lights systems (STLSs) are a promising approach to improve traffic efficiency at intersections. They rely on the information sent by vehicles via C2X communication (like in cooperative awareness messages (CAMs)) at the managed intersection. While there exists a large body of work on privacy-enhancing technologies (PETs) for cooperative Intelligent Transport Systems (cITS) in general, such PETs like changing pseudonyms often impact the performance of cITS applications. This paper analyzes the extent to which different PETs affect the performance of two types of STLSs, a phase-based and a reservation-based STLS. These are implemented in SUMO and combined with four different PETs. Through extensive simulations we then investigate the impact of those PETs on STLS performance metrics like time loss, waiting time, fuel consumption, and average velocity. Our analysis shows that the impact of PETs on performance varies greatly depending on the type of STLS. Finally, we propose a hybrid STLS which is a combination of the two STLS types as a potential solution for limiting the negative impact of PETs on performance.**

***Index Terms*—smart traffic lights systems, privacy, vehicular networking, vehicular edge computing**

I. Introduction

Due to the high number of vehicles, long traffic jams occur regularly in all major cities [1]. This leads to many negative consequences, such as high air pollution [2], long waiting times for the drivers [1], and even a less productive economy [3]. Therefore, there is a need to optimize the traffic flow in cities, and smart traffic lights systems (STLSs) are particularly promising in this respect [4]. An STLS is meant to replace conventional traffic lights and acts as a traffic management system that will dynamically optimize the traffic lights scheduling to generate a better traffic flow.

An STLS relies heavily on V2X communication to gather information about the traffic at the intersection it is managing. In this paper, we propose a vehicular edge computing (VEC) architecture where the nearby vehicles regularly send cooperative awareness messages (CAMs) over the V2X network to the road side unit (RSU) positioned at the intersection. Once the RSU has pre-processed the CAMs from the vehicles, it then forwards them to the cloud where the STLS algorithm is executed upon the data gathered from the CAMs. Such forwarding and central data processing is necessary to coordinate multiple intersections. But for the sake of simplicity, we only simulate one single intersection in this work.

The CAM contains several attributes, including vehicle-IDs that could allow tracking of all vehicles that have passed through the intersection. This obviously raises privacy concerns [5]. In order to better protect the privacy of drivers and passengers, different privacy-enhancing technologies (PETs) can be applied, which can, however, negatively affect the performance of the STLS on traffic efficiency.

Much research has been done on the development of new STLS algorithms to optimize traffic efficiency, and much research has also been done on the development of PETs to protect the privacy of the vehicular user. However, how various types of PETs affect STLS performance has not been analyzed in-depth in other works. To conduct an in-depth analysis, we analyze how different types of STLSs, different PETs, and different traffic situations affect the impact of privacy on the performance of STLSs. For this, we take a three-step approach:
1. A state-of-the-art traffic lights system (TLS) and two different STLSs were implemented to compare the performance of STLSs with a conventional TLS.
2. Pseudonymization-, perturbation-, generalisation- and spoofing-based PETs were integrated to investigate how different technologies impact the performance.
3. Various SUMO[1]-based simulations were conducted to analyze the performance of STLSs with different traffic volumes and traffic distributions.

The remainder of this paper is organized as follows: Section II discusses existing impact analyses of PETs on cITS-related applications. In Section III, we describe the system architecture, the TLS algorithms, and the PETs used in this architecture. Section IV describes the simulation setup and the results of the simulations. These results are discussed in Section V. Finally, in Section VI we draw conclusions and discuss future work.

II. State of the Art

The integration of edge computing technologies to improve vehicular services receives significant attention in the research and industrial community. One research direction here investigates how efficient and suitable PETs can be applied to reduce potential privacy risks for such services [6], [7]. The majority of the existing works on PETs for vehicular systems and services is focusing on privacy alone and the evaluation is mainly analyzing privacy metrics to quantify the impact of the used technologies on privacy [8], [9]. Only few works tackle the impact of the used PET on service quality, and most works are not related to complex cooperative services such as STLSs.

This research is partly accomplished within the project SecForCARs (grant number 16KIS0795). We acknowledge the financial support for the project by the Federal Ministry of Education and Research of Germany (BMBF).

[1]https://www.eclipse.org/sumo/

Some works such as [10] and [11] use differential privacy to protect the vehicular data in an edge context where vehicles send their kinematic data continuously to the edge servers. However, they only evaluate the service utility using generic metrics such as the average, the absolute or the relative errors.

Emara et al. [12] evaluate the impact of several pseudonym change strategies on the service quality in a V2X communication context. The work introduces safety metrics related to two safety critical applications, namely a Forward Collision Warning application and a Lane Change Warning application. As a conclusion, the evaluation demonstrates that some pseudonym change strategies are able to achieve a trade-off between a good privacy level and a reasonable quality of service.

Some works also focus on applications in the area of traffic management. For example, Zhang et al. [13] propose a traffic monitoring system based on several PETs. Their architecture is quite similar to ours and is based on RSUs and a backend Traffic Monitoring Center (TMC). The RSUs and the TMC are considered as non-fully trusted entities. The RSUs that are located at the intersections are in charge of aggregating the individual driving information transmitted by the vehicles and to send them to the TMC. The vehicles use homomorphic encryption to encrypt their data and send them to the RSUs. To avoid privacy statistical attacks, RSUs use differential privacy on the aggregated data before sending them to the TMC. Compared to our work, the authors experimentation was only evaluating the computation time and the communication overhead. How PETs affect the output and, thus, the quality of the provided service was not analyzed in this work.

There also exist some works that have conducted a performance analysis of STLSs when using PETs. For example, Roth et al. [14] applied a pseudonym-based PET on an STLS and analyzed its performance by comparing it with a fixed time traffic lights systems (FTTLSs) and an actuated traffic lights system (ATLS) using inductive loops.

Ying et al. [15] used additive secret sharing to protect the privacy of the vehicles. Here, each vehicle encrypts its data with two different sharings of a secret and sends the data to two different RSUs. The RSUs then process the data to generate the output for the traffic light. To analyze the performance of the STLS, the average waiting time of the STLS was compared with an FTTLS.

In the last two described works, the performance of the STLS was analyzed using one PET and was then compared with other TLSs. However, an analysis of an STLS with and without PETs or with different PETs was not conducted. Thus, to the best of our knowledge, our work is the first work that conducts an in-depth analysis of the impact of several PETs on the performance of STLSs, by analyzing multiple TLSs, PETs, and traffic situations.

III. System Model

In this section, the system architecture used in this paper and the TLSs and PETs that were integrated into this architecture are described.

A. System Architecture

To realise our STLSs, a vehicular edge computing (VEC) architecture is used in this paper. We assume that such an architecture with edge servers and a central cloud will be common in the V2X environment in the future because services in autonomous driving or intelligent traffic management are very computationally intensive. Thus, a central cloud with high computing power and edge servers for preprocessing are necessary to provide a fast response to the vehicles. In addition, a cloud-based solution allows to synchronize traffic lights over multiple intersections. However, as synchronized traffic lights control can become highly complex, we only analyze a single intersection in this paper.

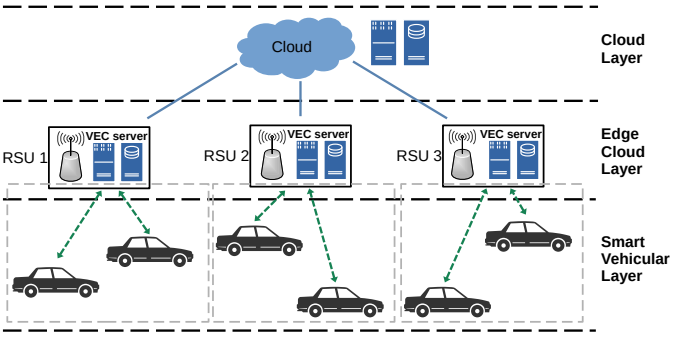

Fig. 1. VEC architecture with multiple RSUs and their area of responsibility (based on [6])

The VEC architecture used in this paper is shown in Figure 1. The smart vehicles periodically broadcast CAMs received by an RSU. The RSUs are located in the edge cloud layer and forward the messages to the cloud layer.

In our system, we assume that the cloud and possibly also the RSUs are operated by an honest-but-curious type of attacker that would not actively attack the system but would use any accessible data to try and track vehicles itineraries.

PETs can therefore be applied to the CAMs either inside the vehicle before they are broadcasted or in a trustworthy RSU before being forwarded to the cloud where the STLS algorithm is executed. When applying privacy mechanisms in the RSU, a trustworthy RSU or a trustworthy environment within the RSU is necessary, which could maybe be realised through a Trusted Execution Environment (TEE).

A more obvious approach would be to apply PETs to CAMs in the vehicle before sending them out. However, this would result in surrounding vehicles receiving CAMs with modified data as PETs were applied to them. For example, location accuracy could have been reduced.

As the surrounding vehicles use the CAMs also for safety-related services such as cooperative adaptive cruise control (CACC) [16], this could also have a negative effect on safety, which should be avoided. Therefore, in the following we assume that PETs are applied in a trustworthy TEE inside the RSU and that our attacker does not have access to the original data from the vehicle before PETs were applied. TEEs have already been used in the automotive sector, e.g. in [17].

B. TLS algorithms

The TLS algorithms were implemented for an isolated intersection within a city with right-hand traffic, shown in Figure 2. For this intersection three TLSs were implemented, one state-of-the-art TLS which provides a reference value for the two STLSs. In the two STLSs different approaches were implemented to analyze whether the impact of privacy mechanisms differs depending on the approach.

The *FTTLS* represents a TLS used at many intersections today. This TLS acts independently of the traffic situation at the intersection, since the duration of the green, yellow and red phases and the sequence of when which traffic light is green are statically defined. To calculate the duration of the green phases, Webster's Method [18] was used, assuming here a medium traffic volume of 1500 vec/h, evenly distributed among all driveways. For the calculation of the yellow and red phases, the approach of Beeber et al. [19] was used.

The *unidirectional smart traffic lights system (USTLS)* is a modified version of [20]. The STLS is located in the cloud and regularly receives the data of the vehicles from the RSU. Based on the position and turning direction of the vehicles, the vehicle density is calculated for each lane. For this purpose, 200 meter long capture zones were defined for each driveway, positioned like the velocity adaptation zones shown in Figure 2. The vehicles in these areas are taken into account in the calculation of the vehicle density, since only these vehicles can pass through the intersection in the upcoming green phase. The two non-conflicting lanes (e.g. left-turn lanes coming from east and west) with the highest vehicle density are set to green for a fixed period of time. After the end of the green phase, the vehicle density is determined again and the corresponding lanes are set to green. Since the green phases are determined dynamically depending on the traffic situation, the USTLS is called a phase-based approach. The yellow and red phase duration are the same as in the FTTLS, since the requirements are the same in both cases.

The USTLS and the ATLS have a similar mode of operation. The difference is, that the ATLS uses sensors, such as induction loops, to determine the traffic situation [21]. Since the USTLS has more detailed information about the traffic situation, we assume that the performance of the USTLS is much better than that of the ATLS. Therefore, the ATLS was not analyzed separately in this paper.

The *bidirectional smart traffic lights system (BSTLS)* uses the same VEC architecture as the USTLS. To schedule the traffic at the intersection, the vehicle-ID, position, length and turning direction of the vehicles are necessary. Based on this data, three steps are performed. In the first step, all vehicles in the 200 m long task capture zone (red area in Figure 2) are registered. Here, in each driveway to the intersection, left-turning vehicles are combined into one task and the straight or right-turning vehicles are combined into one task. This task generation is performed when an unregistered vehicle leaves the task capture zone. In this way, the number of tasks is kept small. In the second step, a time slot is assigned to each task in which all vehicles of the task can pass the intersection. For this purpose, all possible task permutations (but a maximum of $900'000$ permutations) are analyzed, each permutation describing a different sequence in which the tasks pass the intersection. For each sequence, time slots are assigned to these tasks, whereas the length of the time slot is referred to as the processing time in the following. Finally, the sequence with the lowest total processing time is selected. In the third step, the computed time slots of the tasks are sent to the corresponding vehicles so that they can adjust their velocity in the 290 m long velocity adaptation zone (green area in Figure 2) to arrive at the intersection exactly when their time slot starts. Since time slots are reserved for the vehicles, a reservation-based approach is used here.

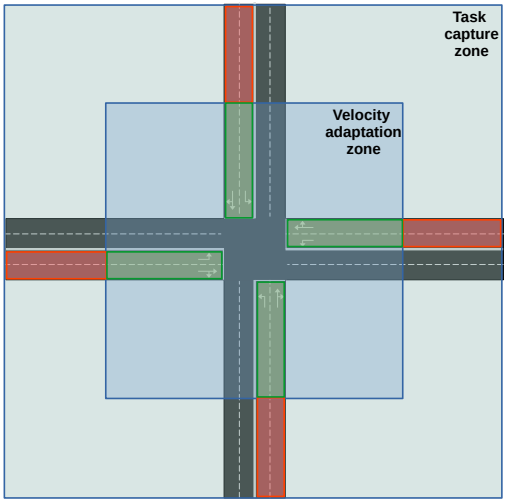

Fig. 2. Intersection with task capture zones (red) and velocity adaptation zones (green) used in the BSTLS.

C. Privacy-enhancing technologies

As described above, in the USTLS the position and turning direction of the vehicle are used. In addition, the vehicle-ID and the vehicle length are used in the BSTLS. Especially the vehicle-ID limits the privacy, because with this attribute several CAMs can be assigned to a vehicle. Therefore, the use of pseudonyms instead of the real vehicle-ID is already part of the standard. But since it can be argued that pseudonyms alone are not sufficient to ensure privacy [22], we investigate additional PETs that obfuscate position and length of the vehicle which hinders tracking attacks as discussed in [23]. For example, the vehicle length is different for different vehicles. Therefore, this attribute could be used as a pseudo-identifier to identify a vehicle even if pseudonyms are used.

1) Vehicle-IDs: To protect the vehicle-ID, the RSU generates a new pseudonym for each CAM so that the cloud cannot link the messages to the same vehicle. In the standard, pseudonyms are changed only at certain intervals, which does not fully protect the vehicles from tracking attacks.

Due to the message-based pseudonyms, adjustments in the BSTLS were necessary. In the original version, the task

generation was conducted when an unregistered vehicle leaves the task capture zone. But when using pseudonyms, it is no longer possible to determine which vehicles have already been registered. Therefore, we modified the task generation to generate tasks every 14.4 sec, which is the time it takes for a vehicle to pass the task capture zone at a speed of 50 km/h.

2) Position and vehicle length: Precise position data could be used to track a vehicle, while the vehicle length could be used as a quasi-identifier of the vehicle. Thus, these two parameters should also be protected. For this purpose, three different approaches were used to analyze how different approaches affect the performance of the STLSs.

The first approach is a *perturbation-based method* where noise is added to the data. Regarding the position, geo-indistinguishability [24] was used, which is a form of differential privacy. Here, noise is added to the position based on a two-dimensional Laplace distribution, so that depending on the distance between two positions, they are indistinguishable with a certain probability. This probability depends on an ϵ-value. Based on this value, the Laplace distribution and thus the average amount of noise added to the original position is different. In our simulations, we used an ϵ-value of 0.5, which means that on average 4 m of noise is added to the original position. For different ϵ-values, the performance results are similar, so that only one ϵ-value was used. The added noise was limited to a maximum of 15 m, which is only slightly above the 13 m width of each driveway. In this way, enough noise can be added that it is indistinguishable in which lane a vehicle is driving. More noise could be problematic because new vehicles might otherwise be too far away from the task capture zone and thus would not be registered in the BSTLS.

To protect the vehicle length, a one-dimensional Laplace distribution was used, based on which noise is added to the vehicle length. Here, a scale parameter has to be specified which is comparable to the ϵ-value described above. In this paper, a scale parameter of 1.0 was used, resulting in an average noise of 1 m. Other scale parameters showed no significant difference in the performance, so that only one scale parameter was used. The noise was limited to half of the vehicle length. In this way, more variations of the vehicle length are possible than when using a fixed maximum noise, since a fixed maximum noise must be very low so that the length of smaller vehicles is not reduced to zero.

Because of the added noise, the USTLS and BSTLS had to be slightly adjusted. The task capture zone was extended in the width by 15 m so that all vehicles driving towards the intersection are definitely registered. Similar adjustments were necessary in the USTLS. However, this leads to the problem that vehicles in the outgoing lanes that are not driving towards the intersection are also registered. To address this problem, we created an optimized version of the BSTLS where we reduced the number of vehicles in each task by a flat 33% (rounded up). This value was determined by empirical testing and showed the best performance. In the USTLS, such an optimization does not make sense, since the determined vehicle densities of all lanes entering the intersection are compared.

The second approach is a *generalization-based method*. To protect the position, spatial cloaking [25] and k-anonymity were used. Here, the STLS only receives an area where the vehicle is located. This area is dynamically defined, so that there are at least k vehicles in it. The maximum area has the size of the task capture zone of a driveway. The minimum area has the same width and a length of 12.5 m. By running several tests with regard to having both, an acceptable performance and privacy gain, a k-value of 3 was determined.

The length of each vehicle was generalized by rounding the actual length to a factor of 3 (e.g. 4.2 m $\rightarrow$ 3 m). This factor was derived from the simulation environment where the length of cars ranges between three and six meters. Thus, the car lengths are evenly distributed between two values. For trucks, which also exist in the simulation environment and are longer than vehicles, a cruder generalization would be appropriate. However, for the sake of simplicity, a uniform generalization was conducted for cars and trucks.

Our last approach is a *spoofing-based method*, which is an extension of the generalization-based method. For the position data, the individual lanes were divided into segments of 12.5 m length and 3.125 m width (equals to the lane width), resulting in a total of 625 segments. To indicate the position of the vehicle, the STLS is provided with the actual segment where the vehicle is located and additional random segments $(1-3)$ where it is claimed that the vehicle would also be located. The number of these additional dummy segments is referred to below as the k-value. In addition, segments are provided where it is stated that the vehicle is **not** located there, so that the total number of segments is always 50. Tests have shown that a lower or higher number of provided segments reduces performance or makes no difference. By summarizing all vehicles on each segment and knowing the k-value, the STLS can estimate the number of vehicles in each segment [26].

To protect the vehicle length, the same approach was used as in the generalization-based method, since the spoofing-based method is an extension of the generalization-based method. Therefore, the adjustments in the STLS algorithms are also the same as those described above.

3) Turning direction: The turning direction was in all PETs not protected, although it does contain sensitive data. The application of privacy mechanisms is difficult here, since this attribute can only take three discrete values. Thus, if PETs were used here, the STLSs would not know in which direction the vehicle wants to go. As a result, no meaningful vehicle density could be calculated in the USTLS and also no meaningful time slot could be calculated in the BSTLS.

IV. Performance analysis

In order to investigate performance impacts of the described TLSs and PETs, several simulations were conducted. The simulation setup, performance metrics and results of the simulations are described in this section.

A. Simulation Setup

The simulation setup is based on the tool SUMO, with which we simulate vehicle traffic on an isolated intersection

with four driveways (see Figure 2). To simulate the communication between the vehicles and the VEC architecture, the Traffic Control Interface (TraCI) provided by SUMO was used, which allows to obtain information about the vehicle states.

Using this simulation setup, for each STLS and each PET, simulations were conducted with traffic volumes ranging from 500 veh/h up to 2500 veh/h in 500 increments. This range of traffic volumes has been observed at real intersections with similar layouts and is therefore assumed to be realistic [27]. However, once PETs were used in the simulations, the traffic volume was largely limited to 1500 veh/h, since the inefficiencies caused by the PETs often made it impossible to realize higher traffic volumes. In addition, different traffic distributions to the lanes were used ($1/2$ to $1/2$, $1/4$ to $3/4$ and $1/8$ to $7/8$). For example, in the third case, $1/8$ of the traffic volume is originating from the driveways north and south and $7/8$ from the driveways east and west. The generated vehicles have different randomly generated lengths to simulate different kinds of vehicles (cars and trucks) and are driving towards the intersection at a maximum speed of 50 km/h, choosing the outgoing driveways at an equal ratio ($1/3$). These vehicles are assumed to be autonomous vehicles. Therefore, a perfect driving behavior of the vehicles is simulated, so that all traffic rules are observed and the vehicles drive as fast as it is allowed and comfortable for the passengers.

For each STLS, each PET, each traffic volume, and each traffic distribution, multiple simulation runs were performed. In each simulation run 2000 seconds of traffic were simulated, where each simulation run was repeated 40 times with different seed values. Finally, the results of each simulation run were used to calculate statistics over the performance metrics.

B. Performance metrics

Four performance metrics were used. The first metric is the *time loss*, which is the additional time a vehicle needs to travel from its starting point to its destination point compared to the time the vehicle would need if there were no other vehicles at the intersection and the respective traffic light was green. The second metric is the *waiting time*, which is the duration the vehicle was standing still. The last two metrics are the *fuel consumption* and the *average velocity* which describe the consumed fuel and average velocity from the starting point to the destination point. The fuel consumption is calculated based on a standard consumption model from SUMO.

C. Results

The simulation results of the TLSs when using the different PETs are described in the following.

1) TLSs without PETs: The results of the three TLSs without PETs and with an equal traffic distribution are shown in Figure 3. The vertical lines in this figure show the minimum and maximum values and the data points show the average values of all simulation runs.

The FTTLS performs similar than the USTLS. The reason for the good performance of the FTTLS is that the traffic volume is equally distributed here on all driveways and the

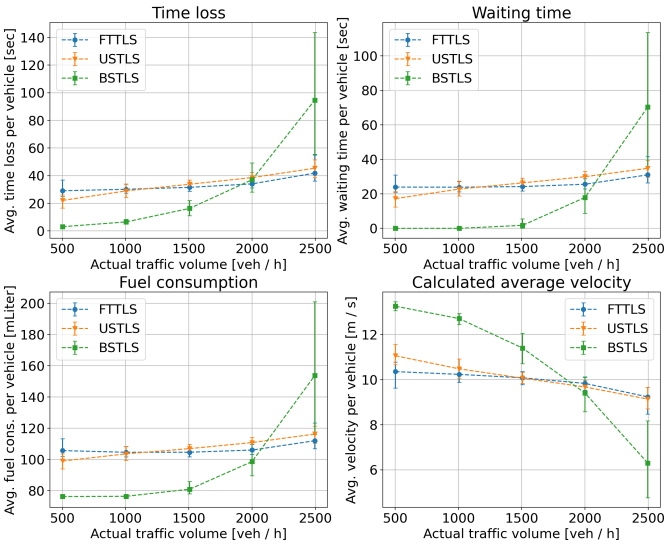

Fig. 3. Results of TLSs without PETs and a traffic distribution of 1/2 to 1/2

FTTLS is designed for this traffic situation. Since the FTTLS is a static implementation that cannot adjust its behavior to different traffic situations, the FTTLS performs much worse with an unequal traffic distribution, as can be seen in Figure 4. For the sake of simplicity and space, we only show the time loss from now on because the time loss correlates with the other attributes, as can be seen in the previous results. The bars of the FTTLS in Figure 4 are at different x-axis positions at higher traffic volumes than in the other two TLSs. The reason for this is that the FTTLS is so inefficient at high traffic volumes that higher traffic volumes could not be achieved.

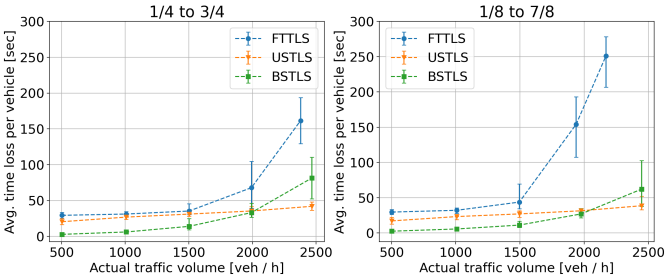

Fig. 4. Results of the TLSs without PETs and an unequal traffic distribution

The BSTLS performs significantly better than the other TLSs at low traffic volumes of up to 1500 veh/h. For example, with an equal traffic distribution, the waiting time here is almost zero, so that the average velocity is higher and the fuel consumption is around 20% lower than in the other TLSs. But the performance of the BSTLS gets worse at higher traffic volumes of 2000 veh/h and above. With unequal traffic distributions, the performance of the BSTLS is similar because it can adjust its behavior to different traffic situations. The same applies to the USTLS.

After analyzing the performance of the three TLSs without PETs, in the next step PETs were applied in the BSTLS and

USTLS. Since the FTTLS is not relying on information from vehicles, no PETs were applied here.

2) PETs in BSTLS: Several PETs were applied here. The first PET is the pseudonymization of the vehicle-IDs.

a) Vehicle-IDs: The results when using pseudonyms are shown in Figure 5. Because the results of both unequal traffic distributions are similar, only $1/8$ to $7/8$ is shown. With an equal traffic distribution, the performance difference is very small, whereas with an unequal traffic distribution and a high traffic volume, the performance difference is higher. For example, with a traffic distribution of $1/8$ to $7/8$ and a traffic volume of 2500 veh/h, the average time loss is around 25 seconds higher when using changing pseudonyms. In this case, there is a very high traffic volume in some driveways. Therefore, vehicles in these driveways can only drive slowly through the task capture zone and thus are in the capture zone for a long time. Due to pseudonym changes, vehicles in the capture zone will be counted multiple times and are registered in multiple tasks, so that the traffic load is overestimated. This leads to overly long time slots, which worsens performance. Depending on how often the pseudonyms are changed and whether they are changed when the vehicle is in the task capture zone, the number of vehicles registered in multiple tasks varies. In this work, we have assumed the worst case, namely that the pseudonyms are changed in every CAM.

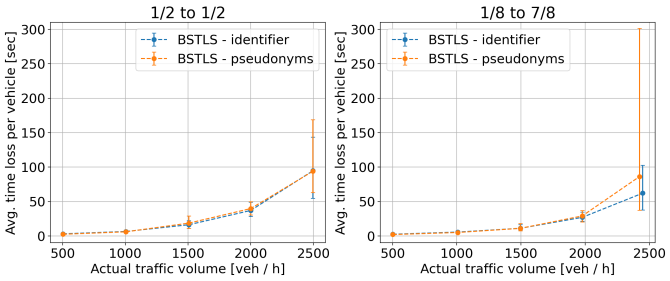

Fig. 5. Results of the BSTLS with the use of pseudonyms for the traffic distributions of 1/2 to 1/2 and 1/8 to 7/8

As mentioned above, based on pseudonyms, further privacy mechanisms for vehicle position and length were investigated in the next step.

b) Position and length: The results of the perturbation, generalization and spoofing-based PET are summarized for an equal traffic distribution in Figure 6, whereas for the other traffic distributions the results of the BSTLS are similar. For all three PETs at a traffic volume of 500 veh/h, the BSTLS performs better than the other TLSs. At a traffic volume of 1000 veh/h, the BSTLS performs slightly worse and at a traffic volume of 1500 veh/h the BSTLS performs significantly worse than the other TLSs.

In the perturbation and generalization-based PET, the optimized versions perform significantly better than the original versions because in the optimized versions estimations of the actual number of vehicles in a task are made. This is useful because due to the noisy positions it is unclear which vehicles are in the capture zone so that sometimes vehicles are erroneously registered in the tasks. Such an estimation is performed a-priori in the spoofing-based PET. When one dummy is generated in the spoofing-based PET, which is comparable in the privacy level to the first two PETs, the results are better than in the optimized versions of the other PETs. Thus, the spoofing-based PET has the lowest impact on the performance. The reason for this is that in the spoofing-based PET more accurate estimations of the actual number of vehicles in the tasks are conducted. When more dummies are generated in the spoofing-based PET, the privacy level increases but the performance decreases because the positions become more inaccurate.

3) PETs in USTLS: To protect the position, the same PETs were used in the USTLS than in the BSTLS. The results of these PETs are shown in Figure 7 for an equal traffic distribution, whereas the results of the USTLS for the other traffic distributions are similar.

In contrast to the BSTLS, the impact of PETs is different in the USTLS, because here for all PETs the impact is very small and independent of the traffic volume. The best results were achieved with the generalization-based PET. For example, at a traffic volume of 500 veh/h, the time loss here is only about two seconds higher than without the use of PETs. In the perturbation-based PET and the spoofing-based PET with one dummy, the time loss is at the same traffic volume about four seconds higher. When more dummies are generated in the spoofing-based PET, the time loss is up to eight seconds higher. So the impact of privacy mechanisms is lower in the USTLS than in the BSTLS, which is discussed in detail in the following section.

V. Discussion

The results show that the impact of PETs is different in the USTLS and BSTLS. In the BSTLS the impact of PETs is low at low traffic volumes and high at high traffic volumes. It is low at low traffic volumes because there is a lot of idle time. If the processing time of the tasks is calculated too long because there are many erroneously registered vehicles in the tasks due to the use of PETs, this will not delay later tasks. Thus, the impact is low at low traffic volumes. This situation is illustrated in Figure 8 on the left. However, at higher traffic volumes there is less idle time. Thus, if many erroneously registered vehicles are in a task and the processing time is therefore too long, this delays all subsequent tasks. This situation is illustrated in Figure 8 on the right. So the reason for the high impact of PETs at high traffic volumes is that a reservation-based approach is used here, where time slots are assigned to the vehicles. Thus, if the time slots are overestimated due to PETs, this leads to a high efficiency loss in the BSTLS.

In contrast to the BSTLS, the impact of PETs on the USTLS is only very small. The reason for this is that the impact of wrong calculations due to PETs is lower in the USTLS than in the BSTLS. If the actual and the calculated vehicle density differ because PETs are used and a lane is set to green even though another lane has a higher vehicle density, this is not

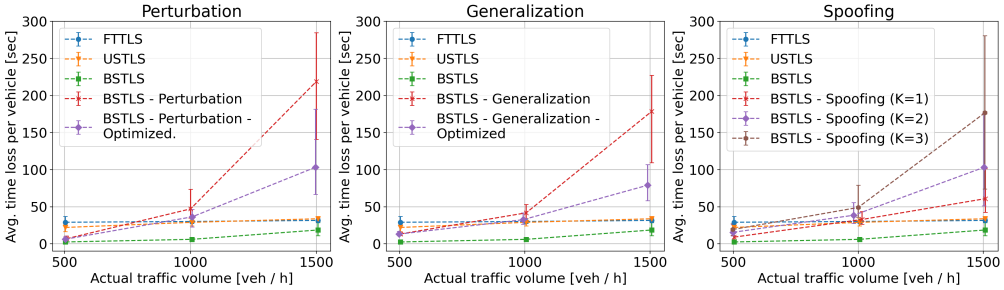

Fig. 6. Results of the BSTLS with the use of three different PETs for the traffic distribution of 1/2 to 1/2

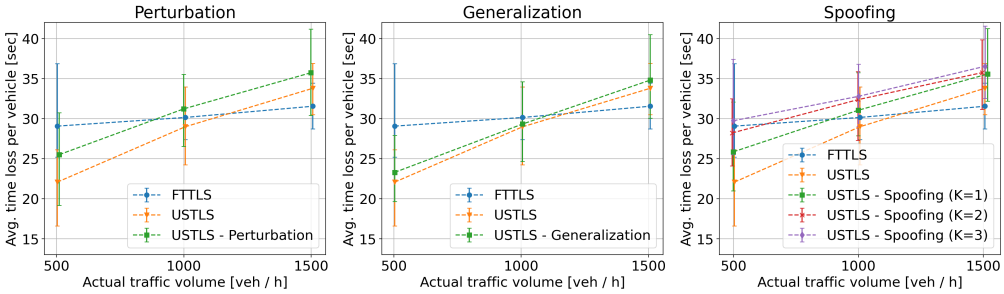

Fig. 7. Results of the USTLS with the use of three different PETs for the traffic distribution of 1/2 to 1/2

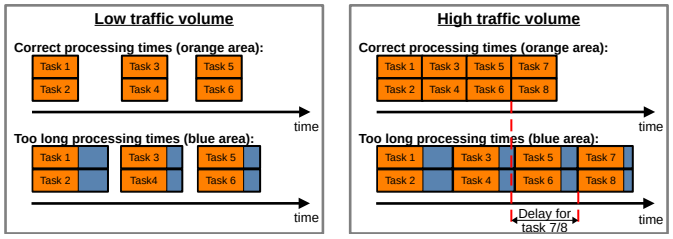

Fig. 8. Effect of too long processing times on subsequent tasks in the BSTLS

optimal but does not lead to significant delays. The reason for this is that vehicles in the lane that was set to green can still pass through the intersection, as no time slots are assigned to the vehicles here. Therefore, in the USTLS wrong calculations do not result in a high efficiency loss.

According to our results, it can be concluded that the impact of PETs on the STLS depends on the specific STLS approach. A phase-based approach is more robust against PETs than a reservation-based approach, because the impact of incorrect estimates is lower in a phase-based approach than in a reservation-based approach. Furthermore, the type of specific PETs and desired level of privacy also affect the performance. As the accuracy of the position data is different depending on the privacy mechanism and privacy level, this results in a different number of erroneously registered vehicles and thus a different impact on the performance. Finally, at least in a reservation-based approach, the traffic volume also affects how much PETs impact the performance.

So overall, we have shown that the impact of PETs on the STLS can differ substantially and care should be taken when selecting a combination of STLSs and PETs.

Although the impact of PETs depends on several parameters, our results indicate that a reservation-based STLS is expected to perform better than a phase-based STLS at low traffic volumes. On the other hand, a phase-based STLS is assumed to provide a better performance at higher traffic volumes compared to a reservation-based STLS. Therefore, a hybrid STLS using a reservation-based approach for lower traffic volumes and a phase-based approach for higher traffic volumes is expected to perform best under varying traffic conditions.

Since the USTLS implements a phase-based approach and the BSTLS implements a reservation-based approach, the described hybrid STLS could be built based on these two STLSs. At low traffic volumes, the performance of the BSTLS with PETs is better than the performance of the other TLSs. At higher traffic volumes, the performance of the USTLSs with PETs is better than the performance of the BSTLSs but often worse than that of the FTTLSs. However, this only applies to the case of equal traffic distributions, where the FTTLS performs best. With unequal traffic distributions, the performance of the FTTLS decreases sharply so that the USTLS with PETs performs better than the FTTLS. Furthermore, a rather simple approach was used in the USTLS. Therefore, we assume that with more advanced phase-based approaches, the performance of the USTLS could become significantly better, while the impact of PETs remains small, since a phase-based approach is used. In this way, an STLS can be created that performs significantly better than the FTTLS and thus as a state-of-the-art TLSs for all traffic situations. So the STLS could still enhance traffic performance even if PETs are applied.

VI. Conclusion and Future Work

In this paper, an in-depth analysis of the impact of privacy-enhancing technologies (PETs) on the performance of smart traffic lights systems (STLSs) was conducted. For this purpose, we implemented a state-of-the-art traffic lights system (TLS), fixed time traffic lights system (FTTLS), and two smart traffic lights systems (STLSs), unidirectional STLS (USTLS) and bidirectional STLS (BSTLS). In the next step, a SUMO-based simulation was created. Then, PETs were applied to the data transmitted by the vehicles, and the resulting data was eventually forwarded to TLSs. Several simulations were run with and without the use of PETs to analyze the performance of the considered TLSs, as well as the performance degradation when PETs are used. Without PETs, the BSTLS performs very well, resulting in almost no waiting times for the vehicles and fuel savings of about 20% compared to the FTTLS. Only at very high traffic volumes the BSTLS performs worse than the other TLSs, rendering the benefit of reservation-based STLSs, such as BSTLS, questionable at high traffic volumes. In contrast, the USTLS performs relatively consistent at all traffic volumes. When PETs are used, several aspects impact the performance of STLSs. The impact mainly depends on whether a reservation-based approach (BSTLS) or a phase-based approach (USTLS) is used, the type of the PET, the desired level of privacy, and the traffic volume. For example, the impact of PETs on the BSTLS is low at a low traffic volume and higher at a high traffic volume. In contrast, the impact of PETs in the USTLS is consistently low.

Based on these results, we propose a hybrid STLS that uses a reservation-based approach for low traffic volumes and a phase-based approach for high traffic volumes. This hybrid STLS should perform well in all traffic conditions, even if PETs are applied. The hybrid STLS could be realized based on the USTLS and BSTLS and would predominantly perform better than the FTTLS. Such a hybrid STLS could be analyzed and optimized in future works. Furthermore, more complex intersections with more lanes could be analyzed in future works. Since the STLS is then more complex, the impact of PETs could be different. Another interesting aspect for future work is the analysis of PETs in the context of several interconnected intersections. In this case, the STLS also becomes much more complex, so that the impact of PETs could again be different.

References

[1] TomTom. (2022) Tomtom traffic index - ranking 2021. [Online]. Available: https://www.tomtom.com/en_gb/traffic-index/ranking/

[2] K. Zhang and S. Batterman, "Air pollution and health risks due to vehicle traffic," *Science of The Total Environment*, pp. 307–316, 04 2013.

[3] S. Fleming. (2019) Traffic congestion cost the us economy nearly 87 billion dollar in 2018. [Online]. Available: https://www.weforum.org/agenda/2019/03/traffic-congestion-cost-the-us-economy-nearly-87-billion-in-2018/

[4] X.-F. Xie, S. F. Smith, L. Lu, and G. J. Barlow, "Schedule-driven intersection control," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 168–189, 2012.

[5] J. Cui, J. Wen, S. Han, and H. Zhong, "Efficient privacy-preserving scheme for real-time location data in vehicular ad-hoc network," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3491–3498, 2018.

[6] D. Liu, Z. Yan, W. Ding, and M. Atiquzzaman, "A survey on secure data analytics in edge computing," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4946–4967, 2019.

[7] J. Ni, A. Zhang, X. Lin, and X. S. Shen, "Security, privacy, and fairness in fog-based vehicular crowdsensing," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 146–152, 2017.

[8] Y. Zhao and I. Wagner, "On the strength of privacy metrics for vehicular communication," *IEEE Transactions on Mobile Computing*, vol. 18, no. 2, pp. 390–403, 2019.

[9] G. P. Corser, H. Fu, and A. Banihani, "Evaluating location privacy in vehicular communications and applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2658–2667, 2016.

[10] S. Ghane Ezabadi, A. Jolfaei, L. Kulik, and R. Kotagiri, "Differentially private streaming to untrusted edge servers in intelligent transportation system," in *2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications*, 2019, pp. 781–786.

[11] S. Ghane, A. Jolfaei, L. Kulik, K. Ramamohanarao, and D. Puthal, "Preserving privacy in the internet of connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, pp. 5018–5027, 2021.

[12] K. Emara, "Safety-aware location privacy in vanet: Evaluation and comparison," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10 718–10 731, 2017.

[13] C. Zhang, L. Zhu, J. Ni, C. Huang, and X. Shen, "Verifiable and privacy-preserving traffic flow statistics for advanced traffic management systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10 336–10 347, 2020.

[14] C. Roth, M. Nitschke, M. Hörmann, and D. Kesdoğan, "itlm: A privacy friendly crowdsourcing architecture for intelligent traffic light management," in *Proceedings of the 9th International Conference on Data Science, Technology and Applications*, 2020, pp. 252–259.

[15] Z. Ying, S. Cao, S. Xu, X. Liu, and M. Ma, "Privacy-preserving intelligent traffic light control," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[16] M. Wolf, A. Willecke, J.-C. Müller, K. Garlichs, T. Griebel, L. Wolf, M. Buchholz, K. Dietmayer, R. W. van der Heijden, and F. Kargl, "Securing cacc: Strategies for mitigating data injection attacks," in *2020 IEEE Vehicular Networking Conference (VNC)*, 2020, pp. 1–7.

[17] P. Boos and M. Lacoste, "Networks of trusted execution environments for data protection in cooperative vehicular systems," in *Vehicular Ad-hoc Networks for Smart Cities*. Springer Singapore, 2020, pp. 99–109.

[18] F. Webster, "Traffic signal settings," *Department of Scientific and Industrial Research*, 1958.

[19] J. Beeber, "An explanation of mats järlström's extended kinematic equation," *ITE journal*, vol. 90, no. 3, pp. 34–38, 2020.

[20] V. Astarita, V. P. Giofrè, G. Guido, and A. Vitale, "The use of adaptive traffic signal systems based on floating car data," *Wireless Communications and Mobile Computing*, vol. 2017, 2017.

[21] M. Eom and B.-I. Kim, "The traffic signal control problem for intersections: a review," *European transport research review*, vol. 12, pp. 1–20, 2020.

[22] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 760–776, 2019.

[23] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in inter-vehicular networks: Why simple pseudonym change is not enough," in *2010 Seventh International Conference on Wireless On-demand Network Systems and Services (WONS)*, 2010, pp. 176–183.

[24] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, 2013, pp. 901–914.

[25] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '03, New York, NY, USA, 2003, p. 31–42.

[26] D. Quercia, I. Leontiadis, L. McNamara, C. Mascolo, and J. Crowcroft, "Spotme if you can: Randomized responses for location obfuscation on mobile phones," in *2011 31st International Conference on Distributed Computing Systems*, 2011, pp. 363–372.

[27] E. Macioszek and A. Kurek, "Extracting road traffic volume in the city before and during covid-19 through video remote sensing," *Remote Sensing*, vol. 13, no. 12, 2021.